Source: STANDARD HANDBOOK OF ELECTRONIC ENGINEERING

CHAPTER 17.2
SWITCHING SYSTEMS

Amos E. Joel, Jr.

A telecommunication service that includes directing a message from any input to one or more selected outputs requires a switching system. The terminals are connected to the switching system by *loops*, which together with the terminals are known as *lines*. The switching systems at nodes of a network are connected to each other by channels called *trunks*. This section deals primarily with systems that provide *circuit switching*, i.e., provision of a channel that is assigned for the duration of a call. Other forms of switching are noted later. Switching systems find application throughout a communication network. They range from small and simple manual key telephone systems or PBXs to the largest automatic local and toll switching systems.

SWITCHING FUNCTIONS

Introduction

A switching system performs certain basic functions plus others that depend on the type of services being rendered. Generally switching systems are designed to act on each message or call, although there are some switches that perform less often, e.g., to switch spare or alternate facilities. Each function is described briefly here and in greater detail in specific paragraphs devoted to each function.

A basic function of a circuit telecommunication switching system is connected by the *switching fabric*,* the transfer of communication from a source to a selected destination. Vital to this basic function are the additional functions of *signaling* and *control* (call processing) (Fig. 17.2.1). Other functions are required to *operate, administer*, and *maintain* the system.

Signaling

Automatic switching is remote-controlled switching. Transfer of control information from the user to the switching office and between offices requires electrical technology and a format. This is known as signaling, and it is usually a special form of data communication. Voice recognition is also used.

*The term *switching fabric* will be used in these paragraphs to identify the implementation of the connection function within a switching system. The term *communications* or *switched network* will refer to the collection of switching systems and transmission systems that constitute a communications system.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

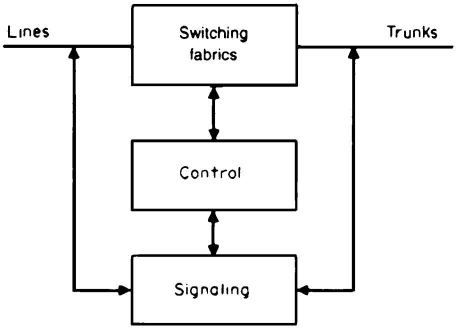

FIGURE 17.2.1 Basic switching functions in circuit switching.

Originally, signaling was developed to accommodate the type of switching technology used for local switching. Most of these systems used dc electric signals. Later, with the advent of longer distances, signaling using single- and multiple-frequency tones in the voice band was developed. Most recently, signaling between offices using digital signals has been introduced over dedicated networks, distinct from the talking channels.

As dialing between more distant countries became feasible, specific international signaling standards were set. These standards were necessarily different from national signaling standards since it was necessary to provide for differences in calling devices (dials) and call handling such as requests for language assistance or restrictions in routing.

Control

Control of switching systems and their application is called *system control*, the overall technique by which a system receives and interprets signals to take the required actions and to direct the switching fabric to carry them out.

In the past, the control of switching systems was accomplished by logic circuits. Virtually all systems now employ stored-program control (SPC). By changing and adding to a program one can modify the behavior of a switching system faster and more efficiently than with wired logic control.

Switching Fabrics

The switching fabric provides the function of connecting channels within a circuit-switching system. Store-and-forward or packet-switching systems do not need complex switching fabrics but do require connecting networks such as a bus structure.

Switching systems have generally derived their names or titles from the type of switching technology used in the switching fabric, e.g., step-by-step, panel, and crossbar. These devices constitute the principal connective elements of switching fabrics. Two-state devices that change in electrical impedance are known as *crosspoints*. Typically electromechanical crosspoints are metallic and go from almost infinite to zero impedance; electronic crosspoints change impedance by several orders of magnitude. The off-to-on impedance ratio must be great enough to keep intelligible signals from passing into other paths in the network (crosstalk). A plurality of crosspoints accessible to or from a common path or *link* is known as a *switch* or, for a rectangular array, a *switch matrix*. A crosspoint may contain more than one gate or contact. The number depends on the information switched and the technology.

Generally a number of stages of switches are used to provide a network in order to conserve the total number of crosspoints required. For connecting 100 inputs to 100 outputs, a single switch matrix requires $100 \times 100 = 10{,}000$ crosspoints. A two-stage fabric requires only 2000 crosspoints when formed with twenty 10×10 matrices. In a two-stage fabric, an output of each first-stage switch is connected to an input of a second-stage switch via a link. There is a connectable path for each and every input to each and every output. Since each input has access to every output, the network is characterized as having *full* access. However, two paths may not simultaneously exist between two inputs on the same first-stage switch and two outputs of a single-output-stage switch. (There is only one link between any first- and second-stage switch.) A second call cannot be placed, and this network is said to be a *blocking network*. By making the switches larger and adding links to provide parallel paths, the chance of incurring a blocking condition is reduced or eliminated. A three-stage Clos *nonblocking fabric* can be designed requiring only 5700 crosspoints. Even fewer crosspoints are needed if existing internal paths can be *rearranged* to accommodate a new connection that would otherwise encounter blocking. The design of most practical switching fabrics includes a modest degree of blocking in order to provide an economical design.

Large central-office switching networks may have more than 100,000 lines and trunks to be interconnected and provide tens of thousands of simultaneous connections. Such networks typically require six to eight stages of switches and are built to carry loads, which result in less than 2 percent of call attempts in the peak traffic period being blocked.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

Network Control

While the switching system as a whole requires a control, the control required for a switching fabric may be separated in part or in its entirety from the system control function. The most general network control accepts the address of the input(s) and output(s) for which an interconnection is required and performs all the logic and decision functions associated with the process of establishing (and later releasing) connections. The control for some networks may be common to many switches or individual to each switch.

Self-routing is also used in fabrics where terminal addresses are transmitted through and acted on by the switches.

Some form of memory is involved with all networks. It may be intimately associated with the crosspoint device employed, e.g., to hold it operated, or it may be separated in a bulk memory. The memory keeps a record of the device in use and of the associated switch path. (In some electronic switching systems it may also designate a path reserved for future use.)

Operation, Administration, and Maintenance (OAM)

When switching systems are to be used by the public, a high-quality continuous service day in and day out over every 24-h period is required.

A system providing such reliable service requires additional functions and features. Examples are continuity of service in the presence of device or component failure and capability for growth while the system is in service.

Separate maintenance and administrative functions are introduced into systems to monitor, test, and record and to provide human control of the service-affecting conditions of the system. These functions together with a human input/output (I/O) interface constitute the basic maintenance functions needed to detect, locate, and repair system and component faults.

In addition to specific maintenance functions, *redundancy* in the switching system is usually necessary to provide the desired quality of service. Complete duplication of an active system with a standby system will protect against one or more failures in one system but presents severe recovery problems in the event of a simultaneous failure of both systems. Judicious subdivision of the system into parts that can be reconfigured (e.g., either of a pair of central processors may work with either of a pair of program memories) can greatly increase the ability of the system to continue operation in the presence of multiple faults.

Where there are many switching entities in a telecommunications network and as systems have become more reliable and training more expensive, the centralization of maintenance has become a more efficient technique. It ensures better and more continuous use of training and can also provide access to more extensive automated data bases that benefit from more numerous experiences.

For public operation, a basic subset of administration and operation features has become accepted as required features. These include the collecting of traffic data, service-evaluation data, and data for call billing.

SWITCHING FABRICS

Three different aspects will be considered in the design of switching fabrics: (1) the types of switching fabrics, (2) the technology of the devices, and (3) the topology of their interconnection.

Types of Switching Fabrics

The three types of switching fabrics are known by the manner in which the message passes through the network.

In *space-division* fabrics analog or digital signals representing messages pass through a succession of operated crosspoints that are assigned to the call for all or most of its duration. In *virtual circuit-switching* systems previously assigned crosspoints are reoperated and released during successive message segments.

In *time-division* fabrics analog or digital signals representing periodically sampled message segments from a plurality of time multiplexed inputs are switched to the same number of outputs. Using equal length segments assigned in time to *time slots* identifies them for address purposes in the system control.

There are two kinds of time-division switching elements, referred to as *space* switches and *time* switches (or time-slot interchanges, TSI). The space switch (also known as a time-multiplexed switch, TMS), shown in

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

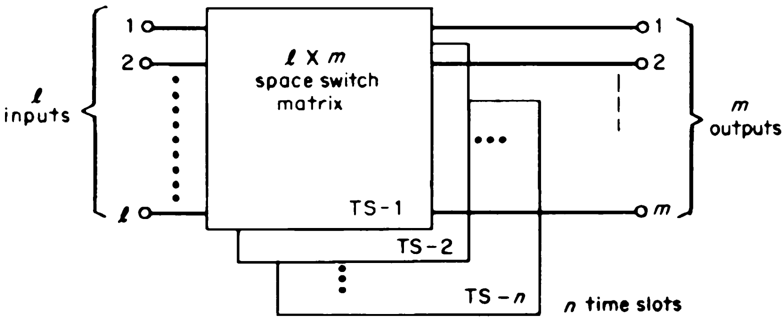

FIGURE 17.2.2 Time-multiplex switch (TMS): space switch.

Fig. 17.2.2, operates like the normal space-switch matrix but with each new time slot the electronic gates are reconfigured to provide a new set of input and output connections. The two-dimensional space switch now has an added third dimension of time.

The time-slot interchange uses a buffer memory, into which one frame of input information is stored. Under direction of the contents of the control memory, transfer logic recorders the sequence of information stored in the buffer, as shown in Fig. 17.2.3. To ensure the timely progression of signals through the TSI, two memories are used, one being loaded while the other is being read. The TSI necessarily creates delays in handling the information stream. Also with storage of message (voice) samples in T-stages, delay of at least one frame (e.g., 125 μs) is introduced into transmission by each switch through which the message passes.

Channels arriving at the switch in time-multiplexed form can be further multiplexed (and demultiplexed) into frames of greater (or lesser) capacity, i.e., at a higher rate and with more time slots. This function is generally used before using TSI so that channels from different input multiplexes can be interchanged.

Time-division switch fabrics are designated by the sequence of time and space stages through which the samples pass, e.g., TSST. The most popular general form of fabric is TST. The choice of others, e.g., STS, is dependent on the size of the fabric and growth patterns.

Analog samples can be switched in both directions through bilateral gates. An efficient and accurate transfer of the pulse is effected by a technique known as *resonant transfer*. For most analog and all digital time-division networks, however, the two directions of signals to be switched are separated. Therefore two reciprocal connections or what is known as four-wire connections (the equivalent of two wires in each direction) must be established in the network. When connections transmit in only one direction, amplification and other forms of signal processing can more readily be switched into the network.

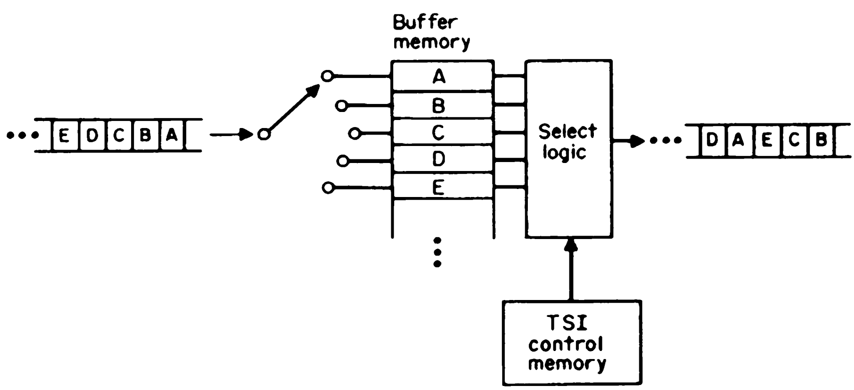

FIGURE 17.2.3 Time-slot interchange (TSI): time switch.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

If multiplexing is performed in such a way that samples from an incoming circuit can be assigned arbitrarily to any of a number of time slots on an outgoing circuit, time-slot interchange and multiplexing are effectively achieved in a single operation.

With the application of digital facilities throughout telecommunications and in particular with the digitalization of speech, *digital time-division fabrics* are currently the most popular form of switching found in public networks. Digital voice communication exists throughout the public network. The *ISDN* (see Chap. 17.4) becomes a reality with digital access line interfaces in the local central offices, completing the end-to-end digital capability. As a result, switched 64,000 b/s clear digital channels are now available not only for voice but also for data.

The number of time slots provided for in a time-division fabric depends on the speed employed. Typically in a voice-band fabric there may be from 32 to 1024 time slots. The coded information in digitized samples may be sent serially (typically 8 b per sample for voice signals), in parallel, or combinations. Extra bits are sometimes added as they pass through the switch for checking parity, for other signals, or to allow for timing adjustments. For digital transmission, the crosspoints used in S stages of a switching fabric need not be linear.

Figure 17.2.4 shows the block diagram of the switching fabric for a no. 4 ESS, a large digital time-division switching system presently being deployed mainly in North America. Incoming digital T carrier streams (five T1 lines with 24 channels each) are further multiplexed to frames of 120 (DS120). The information is buffered in registers to permit synchronization of all inputs. The TSIs on the right side of the figure reverse the order of selecting and buffering; selected input sequences driven by a control memory (not shown) and sequentially gated out of the buffer attain the desired interchange in time. Note that the fabric shown is unilateral (left to right); the complete fabric includes a second unilateral fabric to carry the right-to-left portion of the conversation. This fabric has a maximum of 107,000 input channels, which can accommodate over 47,000 simultaneous conversations with essentially no blocking.

When digital time division fabrics are designed to work with digital carrier (T carrier in the United States) systems either in the loop as pair gain systems, line concentrators, or as interoffice trunks, carrier multiplexed bit streams can be synchronized and applied directly to the switch fabrics requiring no demultiplexing. This represents a cost advantage synergy between switching and transmission.

Frequency Division. Since frequency-multiplex carrier has been used successfully for transmission, its use for switching has been proposed. Connections are established by assigning the same carrier frequency to the two terminals to be connected. Generally to achieve this requires a tunable modulator and a tunable demodulator to be associated with each terminal, and therefore frequency-division switching has had little practical application. *Wave Division* switching is a version of frequency division used in optical or photonic transmission and switching. Other forms of *photonic switching* use true space division to switch optical paths en masse in *free space* (see Hinton and Miller, 1992).

Switching Fabric Technology

Broadly speaking, basically three types of technology have been used to implement switching networks. (1) From the distant past comes the *manually operated switch*, where wires, generally with plug ends, can be moved within the reach of the operator. (2) *Electromechanical* switches can be remotely controlled. They may be *electromagnetically operated* or *power-driven*. Another classification is by the contact movement distance, *gross* motion and *fine* motion. Gross-motion switches inherently have limitations in their operating speeds and tend to provide noisy transmission paths. Consequently, they have seen little recent development. (3) The *electronic* switch is prevalent in modern design.

Electronic Crosspoints. Gross- and fine-motion switches can be used only in space-division systems. Electronic crosspoints achieve much higher operating speeds. Although they can be used in space-, time-, and frequency-division systems, they have the disadvantage of not having as high an open-to-closed impedance ratio as metallic contacts. Steps must therefore be taken to ensure that excessive transmission loss or crosstalk is not introduced into connections.

The crosspoint devices are either externally triggered or are self-latching diodes of the four-layer *pnpn* type. The external trigger may be an electric or optical pulse. The devices have a negative resistance characteristic

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

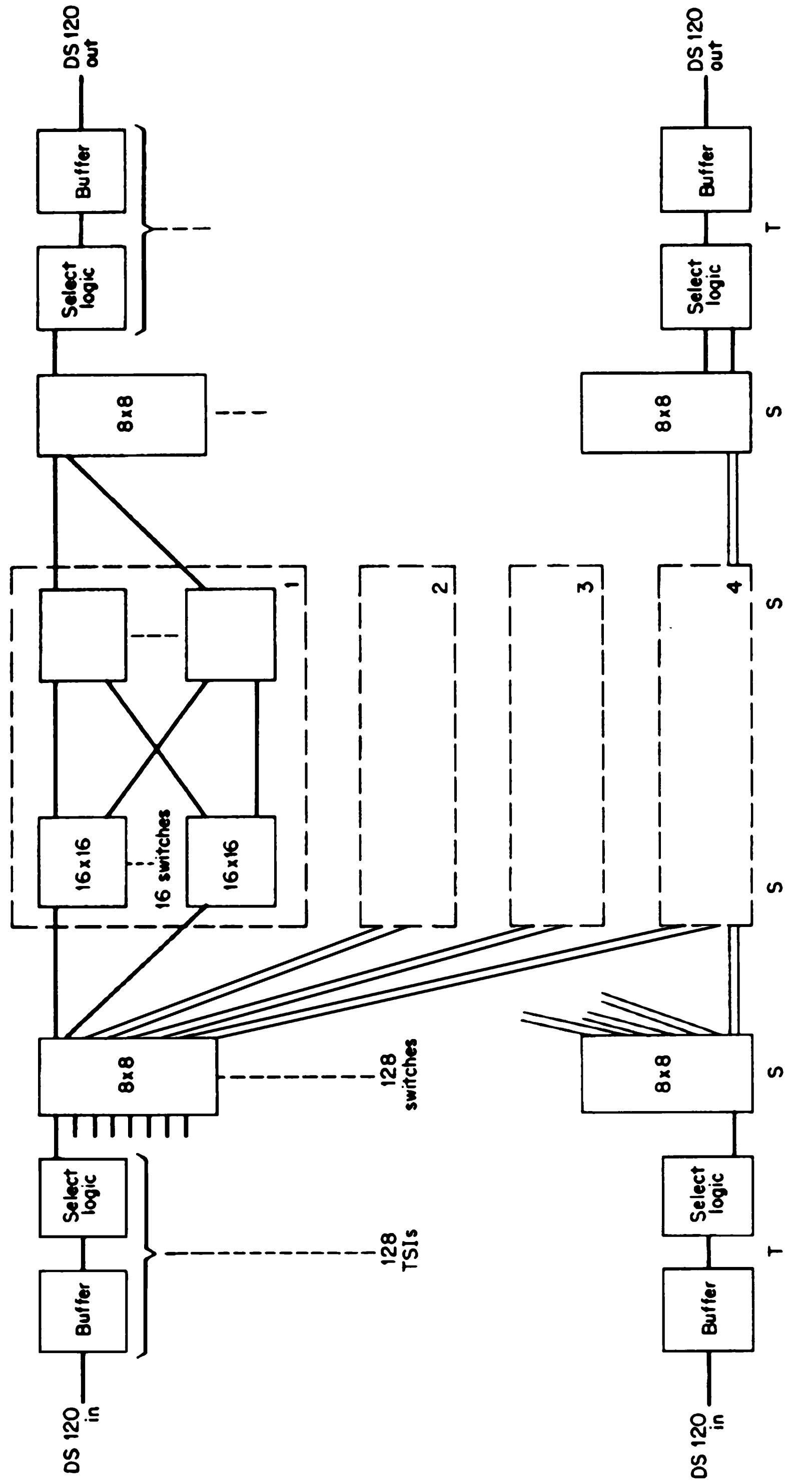

FIGURE 17.2.4 Block diagram of no. 4 ESS digital time-division fabric.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

and are operated in a linear region if they are to pass analog voice or wideband signals. For fixed-amplitude pulse transmission, as in PCM, the devices need not be operated over a linear region.

Electronic crosspoints are generally designed to pass low-level signals at high speed. Recently a new class of high-energy integrated-circuit crosspoints has been developed, which can pass signals used in telephone circuit switching such as ringing and coin control.

Switching Fabric Topology

Of all the switching functions, the topology and traffic aspects of fabrics have been most amenable to normal analytical treatment, although many less precise engineering and technology considerations are also involved.

The simplest fabric is one provided by a single-stage rectangular switch (or, equivalently, a TSI) so that any idle input can reach any idle output. If some of the contacts are omitted, *grading* has been introduced and not every input can reach every output. With the advent of electronic crosspoints and time division, grading has become less important and will not be pursued further here. When inputs to a rectangular switch exceed the outputs, *concentration* is achieved; the converse is *expansion*.

A switching fabric is usually arranged in stages. Input lines connect to a concentration stage, several stages of distribution follow, and a last expansion stage connects to trunks or other lines. Within the design of a switching system, provision is usually made for installation of switches in only the quantity required by the traffic and number of inputs and outputs of each particular application. To achieve this, the size of each stage and sometimes the number of stages is made adjustable. Consideration of control, wiring expense, transition methods (for rearranging the system during growth without stopping service), and technology leads to the configurations selected for each system.

In order to achieve acceptable blocking in networks that are smaller than their maximum designed size, more parallel paths are provided from one stage to the next. In this case, because the distribution need is also reduced, the connections between stages are rewired so that those switch inputs and outputs which are required for distribution in a large network are used for additional parallel paths instead.

It is convenient to divide the fabric into groups of stages or subfabrics according to the direction of the connection. (Calls are considered as flowing from an originating circuit, associated with the request for a connection, to a terminating circuit.) Local interoffice telephone trunks, for example, are usually designed to carry traffic in only one direction. The trunk circuit appearances at a tandem office are then either originating (incoming) or terminating (outgoing). Figure 17.2.5*a* illustrates such an arrangement where the whole network is *unidirectional*.

Telephone lines are usually *bidirectional:* they can originate or terminate calls. For control or other design purposes, however, they can be served by unidirectional stages, as shown in Fig. 17.2.5*b*. Concentration and expansion are normally used with line switching to increase the internal network occupancy above that of lines. In smaller systems a bidirectional network can serve all terminal needs: lines, trunks, service circuits, and so forth (Fig. 17.2.5*c*). When interconnection between trunks, as in a combined local-tandem office, is required, line stages can be kept bidirectional while trunk stages are unidirectional (Fig. 17.2.5*d*). When the majority of trunks are bidirectional, as may occur in a toll office, a bidirectional switching fabric is used (Fig. 17.2.5*e*). Many other configurations are possible.

SYSTEM CONTROLS

Stored Program Control

As discussed earlier, most modern systems use some form of general-purpose stored-program control (SPC). Full SPC implies a flexibility of features, within the capability of existing hardware, by changes in the program.

SPC system controls generally include two memory sets, one for the program and other semipermanent memory requirements and one for information that changes on a real-time basis, such as progress of telephone calls, or the busy-idle status of lines, trunks, or paths in the switching network. These latter writable memories are *call stores* or *scratch-pad memories*. The two memories may be in the same storage medium, in which case there is a need for a nonvolatile backup store such as disc or tape. Sometimes the less frequently used programs are also retrieved from this type of bulk storage when needed.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

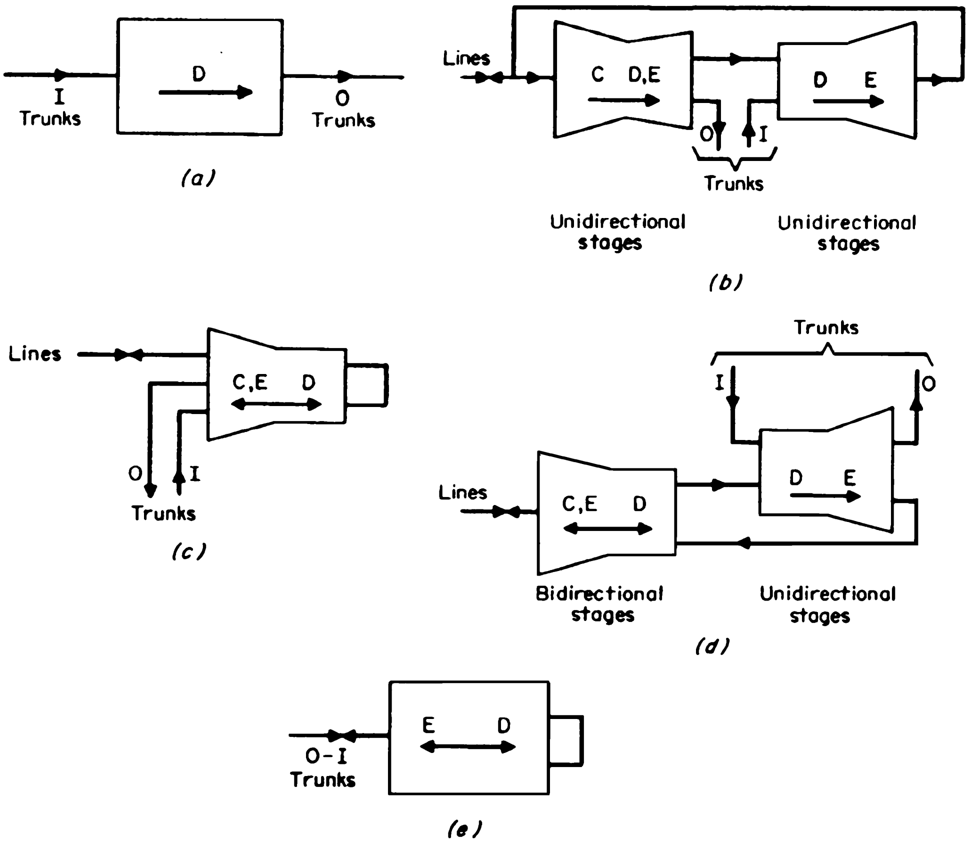

FIGURE 17.2.5 Switching fabrics: (*a*) unidirectional network (tandem); (*b*) unidirectional network (local); (*c*) bidirectional network (local); (*d*) combined network (local-toll); (*e*) bidirectional network; line stages: C = concentration, E = expansion, D = distribution; trunk stages: O = outgoing, I = incoming. Arrows indicate direction of progress of setup of call.

Nonprogram semipermanent memory is required for such data as *parameters* and *translations*. A switching system is generally designed to cover a range of applications; memory, fabric, and other equipment modules are provided in quantities needed for each particular switching office. Parameters define for the program the actual number of these modules in a particular installation. The translation data base provides relations between signal addresses and physical addresses as well as other service and feature class identification information.

Central Control

Single Active. The simplest system control in concept is the *common* or *centralized control*. Before the advent of electronics, the control of a large telephone switching system generally required up to 10 or more central controls. The application of electronics has made it possible for a system to be fully serviced by a single control. This has the advantage of greatly simplifying the access circuits between the control and the remainder of a switching system. It also presents a single point in the system for introducing additional service capabilities. It has the disadvantage that a complete control must be provided regardless of system size, and complete redundancy is often required so that the system can continue to operate in the presence of a single trouble or while changes are being made in the control.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

Redundancy is usually provided by a duplicate central control that is idle and available as a standby to replace the active unit (the unit actually in control) if it has a hardware failure. The duplicate may carry out each program step in synchronism with the active unit; matching circuits observe both units and almost instantaneously detect the occurrence of a fault in either unit. Otherwise central-control faults can be detected by additional redundant self-checking logic built into each central control unit or by software checks. In these latter modes of operation the central controls may be designed to operate independently and share the workload. *Load sharing* allows the two (or more) central controls to handle more calls per unit time than would be possible with a single unit. However, in the event of a failure of one unit, the remaining unit(s) must carry on with reduced system capacity.

Load-sharing represents *independent multiprocessing* where two or more processors may have full capability of handling calls and do not depend on each other. At least part of the call store memory (containing, for example, the busy-idle indications of lines) must be accessible, either directly or through another processor, to more than one processor in order to avoid conflicting actions among processors.

A small office would require less than the maximum number of processors, so that the control cost is lower for that office; as the office grows, more processors can be added. Increasing the number of processors results in decreasing added capacity per processor. Conflicts on processor access to memory and other equipment modules, with accompanying delays, accelerate with the number of processors. Independent multiprocessing or load sharing rapidly reaches its practical limit.

Functional Multiprocessing. Another way to allocate central-control workload is to assign different functions to different processors. Each carries out its task; together they are responsible for total capability of the switching system. This functional or *dependent* multiprocessing arrangement can also evolve from a single central control. A small office can start with the entire program in one processor. When one or more functional processing units are added, the software is modified and apportioned on a functional basis. As in load sharing, the mutually dependent processors must communicate with each other directly or through common memory stores.

In handling calls, each processor may process a portion and hand the next step to a succeeding processor, as in a factory assembly line. This *sequential multiprocessing* has been used in wired-logic switching systems. Virtually all SPC-dependent multiprocessing arrangements are *hierarchical*. A master processor assigns the more routine tasks to subsidiary processors and maintains control of system.

The one or more subsidiary processors may be *centralized* or *distributed*. If the subsidiary processors are centralized, they have full access to network and other peripheral equipment. Distributed controls are dedicated to segments of the switching network and associated signaling circuits. As network and associated signaling equipment modules are added, the control capability is correspondingly enlarged. Most newer switching systems use distributed controls.

TYPES OF SWITCHING SYSTEMS

In the preceding paragraphs the various switching functions were described. A variety of switching systems can be assembled using these functions. The choice of system type depends on the environment and the quantity of the services the system is required to provide. Combining the various types of systems within one embodiment is also possible.

Circuit Switching

Circuit switching is generally used where visual, data, or voice messages must be delivered with imperceptible delay (<0.050 s) and are relatively long. For these applications connections are established through a switch that will pass the required bandwidth. The Nyquist criterion is used in time-division networks by choosing the sample rate to be at least twice the desired maximum bandwidth.

Circuit switching usually (but not always) implies message dialogue, i.e., reciprocal communication in both directions.

Circuit switching has the advantage of being able to switch a very broad spectrum of signal rates. For example, analog or digital video signals can be switched using either metallic or nonmetallic crosspoints.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

Systems Other than Circuit Switching

Services that provide one-way transmission and can accept deferred delivery of data messages have been available in both public and private versions for many years; the switching involved is sometimes called *message switching*. Messages are stored in the switching system until a transmission channel is available, providing efficient use of channels, or until it is convenient to deliver the message to the recipient's station. Further, messages can be provided from the source to the communications system in bulk, or over a single high-volume *port*, which permits efficiencies in the nature and operation of such a source. Semipermanent record storage and message numbering can be provided by the switching system. Message switching lends itself particularly well to *multiple-address messages*.

The implementation of message switching service has evolved to systems that are computer-based, employing a large electronic local telephone switching system augmented with large disk stores, special input-output equipment, and special programs. The airlines industry has long been a user of computer-based electronic message switching for applications ranging from operations traffic to agent traffic.

Message switching, like circuit switching, provides for full duplex operation. A one-way service is known as *file transfer*.

Long data messages or bursts of real-time generated data are divided into packets and transmitted over networks especially arranged to take advantage of these message bursts. The packets are usually of uniform maximum length and prefixed with an address header. Switching takes place to select available network facilities over which to transmit the packets. Packets are reassembled at the receiving terminal or switching node. Retransmission of missing or defective packets are implemented at the switch.

This *packet switching* of data is used in a wide variety of private and public networks; local, campus, metropolitan, wide-area, and so forth, most with their own standard digital protocols. Analog data generally uses voice circuit switching. Packet switching may be enhanced by the use of virtual circuit switching.

Fast packet switching uses locally generated address headers that interpret incoming packet addresses so that they may pass through unique high-speed self-routing switching fabrics. As a result, packets from different inputs destined for the same output route are switched to the same fabric output(s).

By assigning packets to time slots in digital transmission systems and eliminating the checks and retransmission required by packet protocols a technique known as *frame relay* is popular. A switched version is becoming available.

To take advantage of higher speed, broadband digital transmission, such as SONET digitized visual, data, or voice services may be reduced to a fixed size (53 bytes) packet standard called a "cell." This method provides for the delivery of several or multimedia services intermixed over the same transmission facilities. Like packets, cells include headers that are interpreted by the switch in several ways depending on the type of service represented by the information contained in each cell. Networks functioning with these cells are said to employ the *Asynchronous Transfer Mode (ATM)* See McDysan and Spohn (1995).

ATM or *Cell Switching* enables ATM to become a ubiquitous switched services network or broadband ISDN for both private and public applications. Like fast packet switches, switch fabrics for ATM cells are designed to forward cells to the route indicated by the address. Cells are self-routed through space stage fabrics or by a time stage using a shared memory. Paths and channels through the network are identified to enable switches to act as both a combined virtual circuit and packet switch depending on the adapted service represented by particular cells. To reduce the probability of blocking in the switch, cell buffers are needed.

ATM switching is also useful in joining together backbone and other networks that use different protocols internally, such as LANs.

SIGNALING

Basic Purposes

Signaling has three basic purposes: *supervising* the call or message, *addressing* the call or message, and conveying *supplementary* information relative to the call. Supervision consists of indicating a call origination, answer or end, and station alerting. The application of these signals may be between the stations and the switch, in which case it is known as *station signaling*, or between switching offices, when it is known as

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

interoffice signaling. There are two basic approaches to signaling, *per channel* and *common channel*. Per channel signaling may be in or out of the bandwidth used to carry the messages. In the case of digital time-division multiplex transmission, signals may occur in each channel time slot (*in slot*) or in a separate signaling channel (*out of slot*). Common-channel signaling, on the other hand, is carried over one or more channels dedicated to signaling which usually serve many trunks. The electrical characteristics of signals are divided into three categories, *dc*, *ac*, and *digital signals*.

Standards and Compatibility

Only by following standards for the transmission and use of signaling can the various elements of a telecommunication network function together. Signaling standards for public networks operating between nations are generally established by the CCITT, now known as the "International Telecommunications Union-Telecommunications Standardization Sector (ITU-TSS)." The deliberations leading to the setting of these standards involve telecommunication administrations and recognized private operating agencies with the assistance of scientific and industrial organizations. Over the years, as new technology and requirements have appeared, new and revised signaling standards have followed.

National systems are more dependent on the type of signaling required by the local switching systems, and conditions vary widely throughout the world. In the United States the "Alliance for Telecommunications Industry Solutions (ATIS)" is responsible for devising national standards.

Station Signaling

There are few varieties of station signaling in public networks. One is the universal dc loop supervision and dial pulsing. Other signals sent over the loop are those for ringing, coin control, party identification, toll denial (signals indicating limitation of access to the DDD network of a particular station), metering, and so forth. Worldwide they vary in voltage, frequency, and how ground is used as a conductor. An important attribute is the distance or electrical range over which each signal functions satisfactorily.

A standard ac station-address signaling using two-out-of-eight frequencies (twice one-out-of-four) is known as *dual-tone multifrequency* (DTMF) (or "TOUCH TONE" as trademarked by AT&T). For digital access ISDN national standards have been issued by ATIS. *Voice recognition* is also being introduced as a means for addressing telephone calls. For purposes of interconnection to the United States network an Electronic Industries Association standard has been issued on station signaling as applied to PBXs.

Distributed Switching

The range of dc signaling has been particularly important in station signaling since it has determined the location and the number of wire centers required to serve a given area. Various signaling arrangements were developed to extend the dc loop and ringing range of central switching systems. The lower cost of electronics for small transmission and switching systems has made it economical to extend the range further and to decentralize switching.

Interoffice Signaling

Many forms of dc interoffice signaling have been developed and used. They were designed to accommodate specific electromechanical switching systems.

Signaling systems designed for use between offices, particularly over long distances and between countries, must be designed to be transmitted over carrier systems either analog or digital. For analog carrier trunks international standards have been set involving codes of two-out-of-six frequencies (multifrequency code, MFC) for address signaling and one or two frequencies for supervisory signaling. In digital carrier systems a means of carrying encoded supervisory signals in the bit stream is provided to be used with the MFC address signaling.

The per channel signaling systems have a limited number of signals that can be transmitted and usually lengthen the time of use of the transmission paths. To overcome these limitations and provide other advantages,

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

common channel signaling (CCS) systems have been introduced. Common-channel signaling systems consist of a full-time data link between two signaling points and the necessary terminal equipment. By appropriately encoding the data stream, a variety of signals needed for both address and supervisory as well as service enhancements can be transmitted at higher speed between SPC switches. If the CCS link carries the signals for the group of trunks between two offices only, it is said to be operating in the *associated* mode. To signal between two offices with too few trunks to justify an associated link, it is possible to signal over two or more signaling links in tandem via intermediate signaling points. These act as packet switches, routing messages between offices depending on label information contained in each packet. The links carry the signals for more than one trunk group, are not associated with any one, and are said to be operating in the *non-associated* mode.

For common-channel signaling, two international standards were adopted, one optimized for analog transmission (CCITT signaling system 6) and one optimized for use with digital data links (CCITT signaling system 7). With digital transmission becoming dominant, signaling system 7 is gradually being introduced into all networks.

Signaling Fabrics

With common channel signaling becoming the basic signaling method, its signaling links form a separate *signaling network*. The switching systems appear as users on this network.

Each office eventually will rely on this network for all its interoffice or internodal signaling needs. *Signal transfer points* (STPs) serve offices in a geographical *signal region*. An associated signaling link between two offices at any point in the signaling network, it can be added where justified by traffic and reliability considerations. To ensure service reliability, each office has access to two or more STPs in its region (see Fig. 17.2.6). Alternate signal messages from served offices are sent over A links to a different STP in the same region. Signaling networks interconnect not only signaling regions of a particular carrier but also between carriers. The

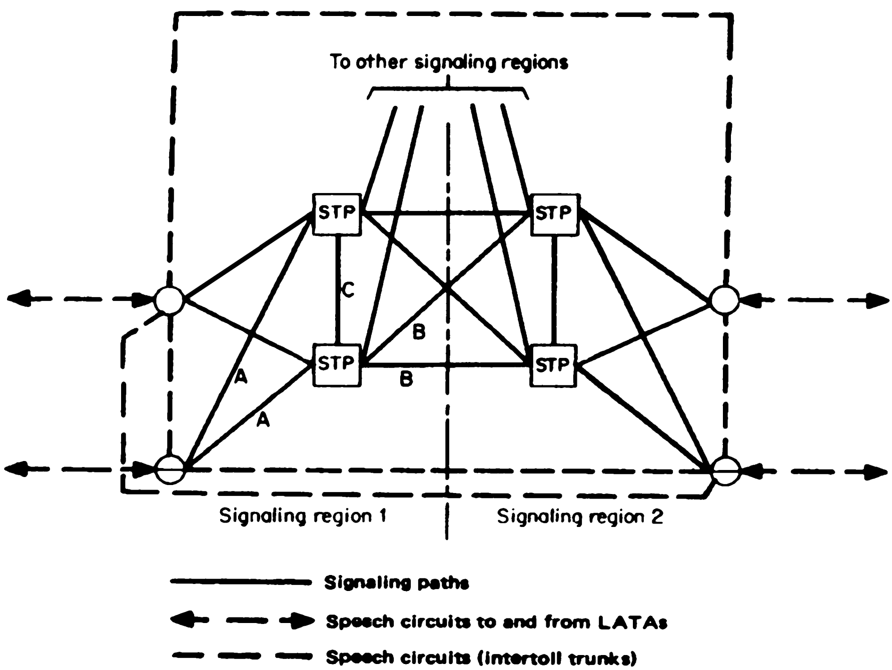

FIGURE 17.2.6 Signaling-network concept. (*Adapted from R. C. Nance and B. Kaskey, ISS Conf. Rec., 1976, p. 413. Copyright 1976 Inst. of Elec. and Comm. Engrs. of Japan. Used by permission*)

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

message switches that serve as STPs employ redundancy of their own and achieve the same quality of service continuity as the switching offices.

Separate data links can also be used for signaling from customer premises, e.g., for passing the calling PBX station identification to the central office in what is known as automatic identified outward dialing (AIOD).

Tones and Announcements

Switching systems need to inform the users of progress being made in serving a particular call. Tones and verbal announcements are used for this purpose. To enable tones to be automatically detected by call originators and to prevent interference with DTMF signals, a precise (frequency) tone plan has been adopted in the United States.

The tones range from *dial tone*, which prompts the caller to start dialing, to *busy tone*, which indicates that a called line is unavailable. Similarly, where more than symbolic information is needed, verbal announcements are given to the caller.

Generally tones and announcements reach the caller through a regular switching network connection. For tones, one or more network terminals are used. If the same source supplies all terminations, low impedance is employed to prevent crosstalk between calls simultaneously reaching the same tone source.

Announcements may be composed from a recorded vocabulary; e.g., intercept systems announce number changes by concatenating a limited vocabulary of phrases and digits.

SERVICES AND OAM FEATURES

A service is what the user perceives as being delivered by a switching system. OAM *features* of central-office telephone switching systems are those uses of system functions which are needed to operate, administer, and maintain the switching, customer terminal, and transmission equipment to provide services.

For example, to provide telephone service (see below) a directory number or address is assigned to a line. Because directory numbers are not permanently associated with switching-network terminations, an administrative feature is required to store this association as a translation in system-control memory.

Telephone Service

The switching functions required to establish a telephone call were listed earlier. Other services also use these functions, plus such additional functions as may be needed.

Telephone service includes the ability to place calls in a telephone network. Users reach other users' telephones by dialing a sequence of digits and can similarly be reached. To respond and alert users, switching systems in a network provide the appropriate signals: dial tone, ringing, busy tone, and so on.

In many networks, operators provide additional services: assisting in placing toll calls (person-to-person), calculating charges and collecting coins in certain calls originating from public coin telephones, providing telephone numbers for users given a name and address (directory assistance), informing users whenever they have dialed a number that has been changed or is not in service (intercept), and so on. These operator services may be fully or partially automated with SPC systems connected to the network. Depending on the operating entity and the service provided, users may be billed for the use of such operator services.

Additional services are provided (again depending on the country and/or the operating entity) for additional one-time or periodic charges. These include DTMF dialing, key telephone or PBX services, or private user networks embedded within the public switched communication network, special telephone sets, and special calling services.

A user wishing to purchase or subscribe to services such as the above makes appropriate arrangements with the business office of the telephone administration or operating company. Some of the other services may be activated, used, and deactivated by dialing directly into an SPC switching system. As the connectivity of SPC systems pervades a communication network, more and more services become available to meet user needs and the trend is toward further automation of the direct control by the user in obtaining such services.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

Call waiting is an example of a calling service available in the North American network where SPC switching systems are in service. When a user who has call-waiting service is engaged in a telephone conversation, a distinctive tone will be received if a new call arrives. The user can be connected to the new call by "flashing," or briefly operating the switch hook. The system responds by placing the existing connection in a hold state and reconnects the subscribing user to the new call. By subsequent flashes of the switch hook the subscribing user can alternate connections between the two calls as needed.

New Services

With the introduction of SPC and associated bulk memory, many new services have been devised for switching offices. Usually they can be implemented with little or no additional hardware except for memory to store additional programs. Switching systems generally include a dynamic set of service capabilities.

Because of their location in networks switching systems provide for control implementation of most services. They serve party and mobile telephones, private branch exchanges, and many different rate categories (flat rate or measured service, coin, and so on.). Some central office systems include capabilities known as "CENTREX" service for serving intramural business needs directly, avoiding the need for key and PBX systems.

Just as the use of SPC in central offices has increased the number and flexibility of the services offered, the use of common channel signaling has created new opportunities to provide new services and features in public and private networks. A popular service made possible by the deployment of the common channel signaling networks is *caller identification*. Except where the caller blocks its application on specific calls or on all calls, the calling telephone number appears on a device at the called telephone after the first ring. Eventually it is expected that this service will be extended nationwide. (See Walker, 1991.) By adding centralized data bases, known as network control points (NCP) or service control points (SCP), to some of the STPs of a signaling network, it is possible to offer new network-based services. Some of those currently in service include translation of called addresses when there are special charge treatments, credit-card calling, and virtual private networks using public network facilities.

By adding *triggers* to the call processing software in each local central office, calls requiring special treatment are directed to an SCP that is able to continue call processing and addressing that is tailored to the needs of specific subscribers or organizations within an area or nationwide. This concept is known as the *Advanced Intelligent Network (AIN). See* Majeti and Prasad, 1993. Services changes are implemented centrally through a *service management system (SMS)* using *service creation environment (SCE)* software.

Customer premises equipment has been at the forefront of offering new services. General business users need both voice and data services. Most of the PBXs designed since 1980 have included provision for switching data. With the growth of such data services have come new switching system architectures that use the same switching networks for connecting digitized voice and packetized data terminations. Many can also serve local area networks (LANs) through separate gateways or with LAN rings built into PBXs.

OAM Features

Network Management. Automatic and manual routing control is provided in large modern networks to route calls around portions of the network that are temporarily congested or where disasters or other types of problems have reduced or eliminated the ability to reach or receive traffic from portions of the network. This is known as *network management*. Network management also includes turning back calls to reduce congestion to a particular destination and the temporary augmentation of facilities to serve overloads. Further, it includes threshold measurements and alerting when offered calls do not appear to be reaching their intended specific destinations. The application of self-healing loops of digital optical facilities for trunks is improving network management and reliability.

Traffic Measurement. To observe how the switching system is operating, an administration needs indications of what loads are being carried by the system components. Typical are counts of calls, measurements of call delays, counts of call dispositions, duration of all circuits busy, and circuit use. The last measurement is best reported in terms of call hours per hour, or *erlangs*. It is often made by periodic counting of busy circuits

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

SWITCHING SYSTEMS

and reporting the average or by summing the actual hours of circuit use for all circuits in a group. Output of data may be printed out directly at a switching system or may be transmitted to a centralized support system.

Billing. Broad allocation of costs to various users of a network can be determined by traffic measurements. In public service systems, more detailed information is needed. Two basic methods are used, *bulk billing* and *detailed billing*. In bulk billing, charge units are allocated to call setup, duration of call, and distance called. In Europe, the pulse metering system is used, where the local office generates pulses on a per line basis. In North America detailed billing is used for toll calls. Call details are recorded, either centrally or locally, and charges are later computed in centralized data processing centers where bills are prepared. This process, called *automatic message accounting* (AMA), has the advantage of more flexibility and full reporting of charges to the customer at the expense of more data processing than is required of bulk billing.

For coin telephones, with pulse metering systems, the charge pulses can be used to control coin collection directly. With detailed billing systems, operators or centralized charge calculation and control capabilities are required to quote charges and handle the collection of coins.

Calls in public data networks are usually billed on the basis of duration of call or number of packets plus a network access charge. The rates may change with the time of day and priority required.

Maintenance

The place of the switching system in a communication network puts unusually severe requirements on its maintenance. Loss of service for any but very short periods is unacceptable. Detection, recovery, diagnosis, and repair of trouble must be carried out while the system continues to process calls. Central processor design therefore incorporates a considerable amount of attention to internal trouble detection and the ability to reassign faulty system elements automatically so that processing can continue without loss of calls. Tests diagnosing the nature of the trouble may also be automatic or subject to request by maintenance personnel.

Alarms and diagnosis results must be generated; most indications appear as lighted lamps or typewritten characters. A standard *user-machine language* is being adopted by the ITV-TSS.

Contents of translation and other semipermanently stored data bases change daily in a large public system. Provision is made to change this information locally or remotely. Care is needed to ensure that changes are free from error and that the data bases will not be lost. In some systems changes are made in two steps. The information is first stored in a temporary *recent change* location and later relocated to a regular data base address.

Another class of feature deals with the cutover of a new system, recovery of a failed system, and change of hardware or software in a working system.

In addition to the tests internal to the switching system, separate test sets, consoles, display boards, and test access switchboards are used to varying degrees and constitute maintenance features. The switching system is also used for connecting test circuits to remote switching and transmission systems.

In the design of systems, a critical factor is the objective in-service time. For public switching systems for two-way voice service, an objective of 2 h in 40 years downtime has been used.

Centralized Operation Support. In the past, operational features have been largely contained within the design of the switching system. There is a trend toward providing these features, as well as maintenance and administrative facilities, at a centralized point where they can service a number of switching systems. The centralized systems are known as *operations support systems* (OSS). Operations includes maintenance and administration as well as operation. The design of switching system hardware and/or software includes features needed to provide and interact with OSSs.

In some centralized support systems, programs and other information for infrequently used real-time call processing can be accessed at a centralized call processor or OSS designed with the required reliability objective.

Applications

This section gives the reader an understanding of typical switching systems. To determine whether a given switching system meets a specific application need one must first know the basic dimensions and capacity requirements.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

Dimensions

Limiting the size of a system are the number of physical terminations for lines and trunks, the call-carrying capacity of the control, the traffic capability of the network(s), and the address range of the memories.

The *terminations* are those lines, trunks, service, and similar circuits, which are principal service inputs and outputs of the system. The service circuits do not extend out of the system but are used to provide a function such as ringing, or call signal receiving or transmitting.

Small data and PBX switching systems may serve tens or hundreds of terminations while large central-office voice-switching systems can have a termination capacity of over 150,000. Most switching systems are designed so that they can grow over a range of terminations, typically an order of magnitude or more. Ideally a switching system should be able to grow over a greater range, and new technology continues to expand this range.

System capacity depends on many factors, including not only the properties of the switching system but also customer traffic characteristics and customer expectation of grade of service. Grade-of-service criteria are usually defined for the average busy season busy hour (ABS) as well as for the 10 highest busy hours and for the highest busy hour. Examples of ABS criteria are 1.5 percent of calls delayed over 3 s in receiving dial tone and 2 percent of incoming calls blocked in the switching network. Capacity must be determined for each central office installation, as it depends on the amount of equipment provided.

Call Capacity. The limit to the maximum office size is frequently designed to be the capacity of the central processor(s). Processor capacity is usually stated in terms of busy hour originating plus incoming calls, although any particular office capacity calculation must consider the proportions of all types of calls. In determining engineered capacity, consideration must be given to the maximum capacity of the central processor. With the growth of services and features designed into the programs of SPCs, the more real time required to process all calls, and the lower the call capacity. Usually this corresponds to almost 100 percent use of the time available for call processing. An allowance must then be made for safety (typically 5 percent) on a high-day engineering basis, and beyond that for the ratio of high day to average business day. (A common ABS use is 70 percent of available processor time.) Then all delay criteria must be checked; if any are not met, the capacity must be lowered further.

Systems have been built with distributed microprocessors serving as many as 100,000 busy-hour call attempts (BHCA), and with a central SPC using multiprocessors and high-speed integrated-circuit technology of over 1,200,000 BHCA. Generally BHCAs are at least twice the number of successfully processed calls, owing to the large number of calls where a receiver–off-hook signal is detected without the completion or initiation of dialing.

Data-message or packet-switching capacity is measured in terms of maximum rate of *throughput*. The capacity of the control includes overhead as well as message handling. The control capacity is generally considered to be the major factor in throughput. Processing capacity can be limited by congestion delays within a system delivering or receiving calls for processing.

Switching-Fabric Capacity. The switching-fabric capacity is expressed in the number of erlangs that can be carried at an objective blocking. Fabrics for smaller offices are partially equipped; each arrangement will have its own capacity. If the average call duration is long, the erlang capacity may limit call capacity below that of the processor:

$$\text{Network calls/h} = \frac{\text{erlang capacity}}{\text{call holding time (h)}}$$

Because processors are not partially equipped and fabrics can be, designers choose to design fabrics with a smaller chance of being the limiting dimension of system capacity.

Memory Capacity. The total memory requirements depend on many components. An important factor is whether there is only one storage subsystem or separate storage subsystems are used for different storage needs. Typically, separate subsystems might be used for program and call data storage. In a message-switching system separate subsystems might be used for call processing and message storage.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

One limit on memory size is the amount of memory that can be directly addressed. By having larger program words (more bits per word), more memory can be accessed, but this means that larger, more expensive program memories are required.

As in most SPC systems, software techniques can be used to extend the address range at the expense of real time. One address can refer to a table containing another range of addresses.

Physical memory modules are used to form a storage subsystem. Only as many memory modules are used as are needed to provide for the memory requirements for a particular installation.

Host/Remote Systems

The dimensional limits represent the individual maxima in each case (call attempts per hour, erlangs, and terminations). Depending on the specific environment, any one of these may limit further growth of the system, and the remaining limits would be unattainable. These limits represent approximate values, which in themselves depend on assumptions of the system environment, e.g., the ratio of intraoffice to interoffice calls in a local central office. The limits change with time as additional features and services are added to a system or when improvements in hardware or software are introduced.

Two basic switching entities are used at points distant from the host central wire center. One is the *remote line concentrator* (RLC) and the other the *remote switching unit* (RSU). Both the RLC and RSU are used to concentrate traffic at a point closer to the lines they serve. The RLC provides only for remoting this network function. Generally, in the central office an equivalent expansion function is provided, and consequently each RLC line is given a central office appearance. Some RLC systems eliminate the need for this expansion by connecting RLC trunks to links within the switching network. Generally the line range is not extended with the use of RLCs. Generally digital subscriber line carrier (SLC) systems are used to connect the host with the RLCs and RSUs. This provides a range as much as 100 mi. Recently fiber-optic links with digital transmission have also been employed by extending internal system links as in the No. 5 ESS.

Should all the trunks between the RLC and the central office be busy or the facilities carrying the trunks be severed, calls to and from the RLC lines cannot be served. Since the lines served are in a confined area, they are subject to higher traffic variation. Concentrators are generally small, serving no more than a few hundred lines.

With the advent of microprocessors, not only the network but also the control can be remoted. This means that more intelligence can be designed into the remote switching. Also, the remote switch may be able to complete intra-RSU calls without using trunks to the host office. The RSUs have been developed with SPC and limited call-processing capability so that they can provide basic service if the link to the host is lost. This is known as *stand-alone* capability.

Trends in Switching Systems

Switching systems have made the transition from those relays and other forms of electromechanical switches to all-electronic fabrics and controls. More than half of the switches in service are digital time-division systems. Whether time or space division, the major ingredient to stimulate this transition has been stored-program control. SPC has provided the flexibility and power to add to and change switching system capabilities easily. As administrations make the transition to national and international networks of SPC systems, capabilities are being extended further. The advent of common-channel signaling has provided a separate fast communication network between the processors of SPC switching systems, further increasing the capability of telecommunication networks. New services concepts that are possible within an SPC office may be extended to the entire intelligent network.

ISDN will be in full blossom by the end of this century. Other digital loop technologies will further extend the end-to-end digital capabilities of the public network. The inherent high-speed capability of optical fiber for digital transmission will be used for *Broadband ISDN (B-ISDN)*. Initially ATM as part of a B-ISDN for data, file transfers and images will further extend private backbone networks to the public B-ISDN.

Switching provides access to the public network for wireless telephones and terminals. For cellular radio it also provides not only the means for access but follows mobiles as they move from one area to another.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.

As mobiles move from one system to another, Intelligent Network switches and databases will provide seamless connections. *Personal Communication System (PCS)* will provide wireless switched access for users who do not move significantly during a call.

As more telecommunications services vendors enter the switched public network many new switching solutions will be needed to deal with such things as *number portability*. The telephone number addresses used by switches in the worldwide public network are the immutable factor in extending telecommunications services. Competition generally places new requirements on public networks that are manifested and implemented in the switches.

Downloaded from Digital Engineering Library @ McGraw-Hill (www.digitalengineeringlibrary.com)
Copyright © 2004 The McGraw-Hill Companies. All rights reserved.
Any use is subject to the Terms of Use as given at the website.