

---

# TRANSMISSION AND SWITCHING: CORNERSTONES OF A NETWORK

## 4.1 TRANSMISSION AND SWITCHING DEFINED

The IEEE defines *transmission* as the propagation of a signal, message, or other form of intelligence by any means such as optical fiber, wire, or visual means. Our definition is not so broad. Transmission provides the transport of a signal from an end-user source to the destination such that the signal quality at the destination meets certain performance criteria.

*Switching* selects the route to the desired destination that the transmitted signal travels by the closing of switches in either the space domain or the time domain or some combination(s) of the two.

Prior to about 1985, transmission and switching were separate disciplines in telecommunication with a firm dividing line between the two. Switching engineers knew little about transmission, and transmission engineers knew little about switching. As we mentioned in Chapter 1, that dividing line today is hazy at best. *Signaling* develops and carries the control information for switches. If a transmission path becomes impaired, signaling becomes ineffectual and the distant-end switch either will not operate or will not function correctly, misrouting the connectivity. Timing, which is so vital for the digital transmission path, derives from the connected switches.

## 4.2 TRAFFIC INTENSITY DEFINES THE SIZE OF SWITCHES AND THE CAPACITY OF TRANSMISSION LINKS

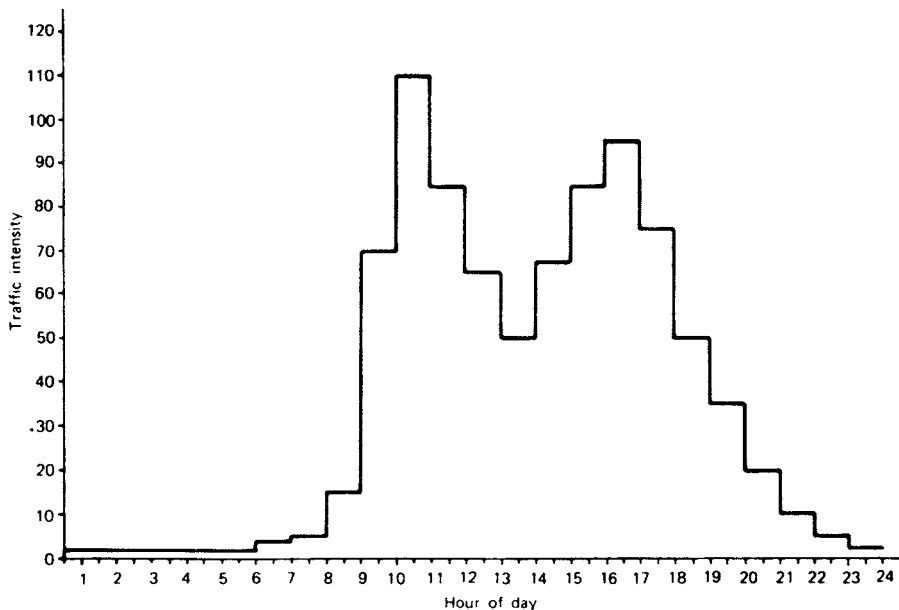
### 4.2.1 Traffic Studies

As we have already mentioned, telephone exchanges (switches) are connected by trunks or junctions.<sup>1</sup> The number of trunks connecting exchange X with exchange Y is the number of voice pairs or their equivalent used in the connection. One of the most important steps in telecommunication system design is to determine the number of trunks required on a

<sup>1</sup>The term *junction* means a trunk in the local area. It is a British term. *Trunk* is used universally in the long-distance plant.

route or connection between exchanges. We could say we are *dimensioning* the route. To dimension the route correctly we must have some idea of its usage—that is, how many people will wish to talk at once over the route. The usage of a transmission route or switch brings us into the realm of traffic engineering; and usage may be defined by two parameters: (1) *calling rate*, or the number of times a route or traffic path is used per unit time period; or more properly defined, “the call intensity per traffic path during the busy hour (BH)”; and (2) *holding time*, or “the average duration of occupancy of one or more paths by calls.” A *traffic path* is a “channel, time slot, frequency band, line, trunk switch, or circuit over which individual communications pass in sequence.” *Carried traffic* is the volume of traffic actually carried by a switch, and *offered traffic* is the volume of traffic offered to a switch. Offered traffic minus carried traffic equals *lost calls*. A lost call is one that does not make it through a switch. A call is “lost” usually because it meets congestion or blockage at that switch.

To dimension a traffic path or size a telephone exchange, we must know the traffic intensity representative of the normal busy season. There are weekly and daily variations in traffic within the busy season. Traffic is random in nature. However, there is a certain consistency we can look for. For one thing, there is usually more traffic on Mondays and Fridays, and there is a lower volume on Wednesdays. A certain consistency can also be found in the normal workday variation. Across a typical day the variation is such that a one-hour period shows greater usage than any other one-hour period. From the hour of the day with least traffic intensity to the hour of greatest traffic, the variation can exceed 100:1. Figure 4.1 shows a typical hour-to-hour traffic variation for a serving switch in the United States. It can be seen that the busiest period, the *busy hour* (BH), is between 10 A.M. and 11 A.M. (The busy hour from the viewpoint of grade of service was introduced in Section 1.3.4). From one workday to the next, originating BH calls can vary as much as 25%. To these fairly “regular” variations, there are also unpredictable peaks caused by stock market or money market activity, weather, natural disaster, international events, sporting events, and so on. Normal traffic growth must also be taken into account.



**Figure 4.1** Bar chart of traffic intensity over a typical working day. (US, mixed business and residential).

Nevertheless, suitable forecasts of BH traffic can be made. However, before proceeding further in this discussion, consider the following definitions of the busy hour.

1. *Busy Hour*. The busy hour refers to the traffic volume or number of call attempts, and is that continuous one-hour period being wholly in the time interval concerned for which this quantity (i.e., traffic volume or call attempts) is greatest.
2. *The Average Busy Season Busy Hour (ABSBH)*. This is used for trunk groups and always has a grade of service<sup>2</sup> criterion applied. For example, for the ABSBH load, a call requiring a circuit in a trunk group should encounter *all trunks busy* (ATB) no more than 1% of the time.

Other definitions of the busy hour may be found in Ref. 1.

When dimensioning telephone exchanges and transmission routes, we shall be working with BH traffic levels and care must be used in the definition of the busy hour.

Peak traffic loads are of greater concern than average loads for the system planner when dimensioning switching equipment.

Another concern in modern digital switching systems is *call attempts*. We could say that call attempts is synonymous with offered traffic. Even though a call is not *carried* and is turned away, the switch's processor or computer is still exercised. In many instances a switch's capability to route traffic is limited by the peak number of call attempts its processor can handle.

**4.2.1.1 Measurement of Telephone Traffic.** If we define *telephone traffic* as the aggregate of telephone calls over a group of circuits or trunks with regard to the duration of calls as well as their number, we can say that traffic flow ( $A$ ) is expressed as

$$A = C \times T, \quad (4.1)$$

where  $C$  designates the number of calls originated during the period of one hour, and  $T$  is the average *holding time*, usually given in hours.  $A$  is a dimensionless unit because we are multiplying calls/hour by hour/call.

Suppose that the average holding time is 2.5 minutes and the calling rate in the BH for a particular day is 237. The traffic flow ( $A$ ) would then be  $237 \times 2.5$ , or 592.5 call-minutes (Cm) or  $593.5/60$ , or about 9.87 call-hours (Ch).

The preferred unit of traffic intensity is the *erlang*, named after the Danish mathematician A.K. Erlang (Copenhagen Telephone Company, 1928). The erlang is a dimensionless unit. One erlang represents a circuit occupied for one hour. Considering a group of circuits, traffic intensity in erlangs is the number of call-seconds per second or the number of call-hours per hour. If we knew that a group of 10 circuits had a call intensity of 5 erlangs, we would expect half of the circuits to be busy at the time of measurement.

In the United States the term *unit call* (UC), or its synonymous term, *hundred call-second*, abbreviated ccs,<sup>3</sup> generally is used. These terms express the sum of the number of busy circuits, provided that the busy trunks were observed once every 100 seconds (36 observations in 1 hour) (Ref. 2). The following simple relationship should be kept in mind: 1 erlang = 36 ccs, assuming a 1-hour time-unit interval.

<sup>2</sup>*Grade of service* refers to the planned value criterion of probability of blockage of an exchange. This is the point where an exchange just reaches its full capacity to *carry* traffic. This usually happens during the busy hour.

<sup>3</sup>The first letter c in ccs stands for the Roman number 100.

Extensive traffic measurements are made on switching systems because of their numerous traffic-sensitive components. Usual measurements for a component such as a service circuit include call attempts, calls carried, and usage. The typical holding time for a common-control element in a switch is considerably shorter than that for a trunk, and short sampling intervals (e.g., 10 seconds) or continuous monitoring are used to measure usage.

**4.2.1.2 Blockage, Lost Calls, and Grade of Service.** Let's assume that an isolated telephone exchange serves 5000 subscribers and that no more than 10% of the subscribers wish service simultaneously. Therefore, the exchange is dimensioned with sufficient equipment to complete 500 simultaneous connections. Each connection would be, of course, between any two of the 5000 subscribers. Now let subscriber 501 attempt to originate a call. She/he cannot complete the call because all the connecting equipment is busy, even though the line she/he wishes to reach may be idle. This call from subscriber 501 is termed a *lost call* or *blocked call*. She/he has met blockage. The probability of encountering blockage is an important parameter in traffic engineering of telecommunication systems. If congestion conditions are to be met in a telephone system, we can expect that those conditions will usually be encountered during the BH. A switch is dimensioned (sized) to handle the BH load. But how well? We could, indeed, far overdimension the switch such that it could handle any sort of traffic peaks. However, that is uneconomical. So with a well-designed switch, during the busiest of BHs we can expect moments of congestion such that additional call attempts will meet blockage. *Grade of service*<sup>4</sup> expresses the probability of meeting blockage during the BH and is commonly expressed by the letter  $p$ . A typical grade of service is  $p = 0.01$ . This means that an average of one call in 100 will be blocked or "lost" during the BH. Grade of service, a term in the Erlang formula, is more accurately defined as the *probability of blockage*. It is important to remember that lost calls (blocked calls) refer to calls that fail at *first trial*. We discuss attempts (at dialing) later—that is, the way blocked calls are handled.

We exemplify grade of service by the following problem. If we know that there are 345 seizures (i.e., lines connected for service) and 6 blocked calls (i.e., lost calls) during the BH, what is the grade of service?

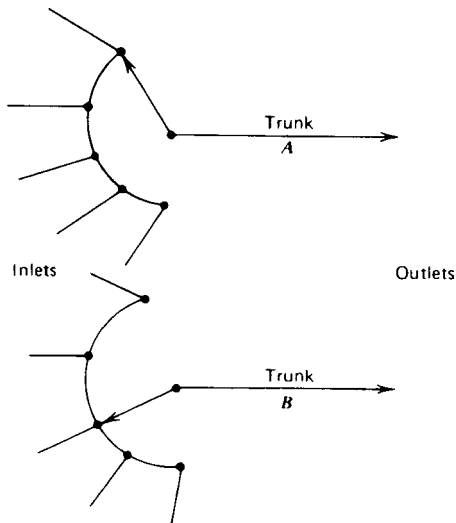
$$\begin{aligned} \text{Grade of service} &= \text{Number of lost calls} / \text{Number of offered calls} \\ &= 6 / (345 + 6) = 6 / 360 \\ p &\approx 0.017. \end{aligned} \tag{4.2}$$

The average grade of service for a network may be obtained by adding the grade of service provided by a particular group of trunks or circuits of specified size and carrying a specified traffic intensity. It is the probability that a call offered to the group will find available trunks already occupied on first attempt. This probability depends on a number of factors, the most important of which are (1) the distribution in time and duration of offered traffic (e.g., random or periodic arrival and constant or exponentially distributed holding time), (2) the number of traffic sources [limited or high (infinite)], (3) the availability of trunks in a group to traffic sources (full or restricted availability), and (4) the manner in which lost calls are "handled." Several new concepts are suggested in these four factors. These must be explained before continuing.

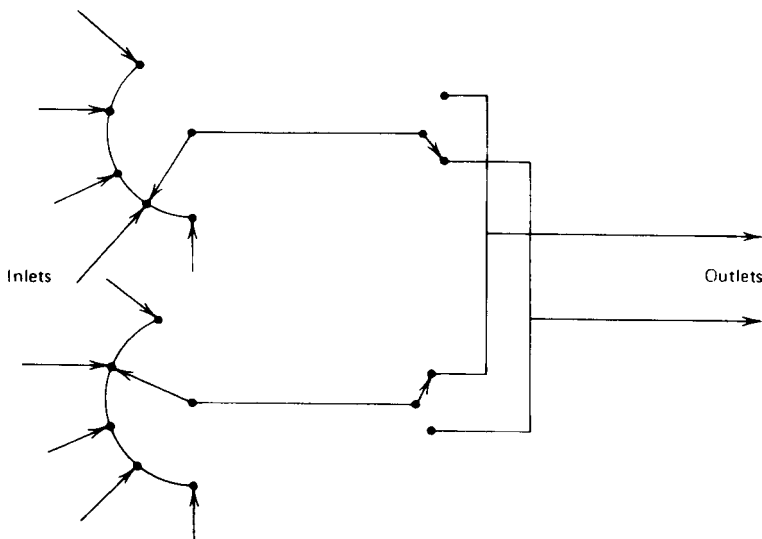
<sup>4</sup>Grade of service was introduced in Section 1.3.4.

**4.2.1.2.1 Availability.** Switches were previously discussed as devices with lines and trunks, but better terms for describing a switch are *inlets* and *outlets*. When a switch has *full availability*, each inlet has access to *any* outlet. When not all the free outlets in a switching system can be reached by inlets, the switching system is referred to as one with *limited availability*. Examples of switches with limited and full availability are shown in Figures 4.2a and 4.2b

Of course, full availability switching is more desirable than limited availability, but is more expensive for larger switches. Thus full availability switching is generally found only in small switching configurations and in many new digital switches (see Chapter 6). *Grading* is one method of improving the traffic-handling capabilities of switching configurations with limited availability. Grading is a scheme for interconnecting switching subgroups to make the switching load more uniform.



**Figure 4.2a** An example of a switch with limited availability.



**Figure 4.2b** An example of a switch with full availability.

4.2.1.2.2 “*Handling*” of Lost Calls. In conventional telephone traffic theory, three methods are considered for the handling or dispensing of lost calls:

1. Lost calls held (LCH)
2. Lost calls cleared (LCC)
3. Lost calls delayed (LCD)

The LCH concept assumes that the telephone user will immediately reattempt the call on receipt of a congestion signal and will continue to redial. The user hopes to seize connection equipment or a trunk as soon as switching equipment becomes available for the call to be handled. It is the assumption in the LCH concept that lost calls are held or waiting at the user’s telephone. This concept further assumes that such lost calls extend the average holding time theoretically, and in this case the average holding time is zero, and all the time is waiting time. The principal traffic formula (for conventional analog space division switching) in North America is based on the LCH concept.

The LCC concept, which is primarily used in Europe or those countries that have adopted European practice, assumes that the user will hang up and wait some time interval before reattempting if the user hears the congestion signal on the first attempt. Such calls, it is assumed, disappear from the system. A reattempt (after the delay) is considered as initiating a new call. The Erlang B formula is based on this criterion.

The LCD concept assumes that the user is automatically put in queue (a waiting line or pool). For example, this is done, of course, when an operator is dialed. It is also done on all modern digital switching systems. Such switches are computer-based for the “brains” of the control functions and are called switches with *stored program control* (SPC). The LCD category may be broken down into three subcategories, depending on how the queue or pool of waiting calls is handled. The waiting calls may be handled *last in first out*, *first in first out*, or *at random*.

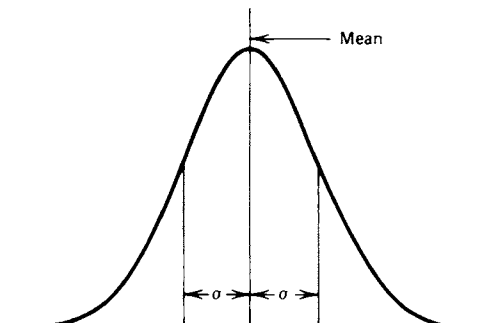
4.2.1.2.3 *Infinite and Finite Traffic Sources*. We can assume that traffic sources are either infinite or finite. For the infinite-traffic-sources case the probability of call arrival is constant and does not depend on the occupancy of the system. It also implies an infinite number of call arrivals, each with an infinitely small holding time. An example of finite traffic sources is when the number of sources offering traffic to a group of trunks is comparatively small in comparison to the number of circuits. We can also say that with a finite number of sources the arrival rate is proportional to the number of sources that are not already engaged in sending a call

4.2.1.2.4 *Probability-Distribution Curves*. Telephone-call originations in any particular area are random in nature. We find that originating calls or call arrivals at an exchange closely fit a family of probability-distribution curves following a Poisson<sup>5</sup> distribution. The Poisson distribution is fundamental in traffic theory

Most probability-distribution curves are two-parameter curves; that is, they may be described by two parameters: mean and variance. The *mean* is a point on the probability-distribution curve where an equal number of events occur to the right of the point as to the left of the point. Mean is synonymous with average. Consult Figure 4.3.

The second parameter used to describe a distribution curve is the *dispersion*, which tells us how the values or population are dispersed about the center or mean of the curve. There are several measures of dispersion. One is the familiar *standard deviation*.

<sup>5</sup>S. D. Poisson was a nineteenth-century French mathematician/physicist specializing in “randomness.”



**Figure 4.3** A normal distribution curve showing the mean and standard deviation,  $\sigma$ .

The standard deviation is usually expressed by the Greek letter sigma ( $\sigma$ ). For example,  $1\sigma$  either side of the mean in Figure 4.3 will contain about 68% of the population or measurements,  $2\sigma$  will contain about 95% of the measurements, and  $3\sigma$  will contain around 99% of the subject, population, or whatever is being measured. The curve shown in Figure 4.3 is a *normal distribution curve*.

#### 4.2.2 Discussion of the Erlang and Poisson Traffic Formulas

When dimensioning a route, we want to find the optimum number of circuits to serve the route. There are several formulas at our disposal to determine that number of circuits based on the BH traffic load. In Section 4.2.1.2, four factors were discussed that will help us to determine which traffic formula to use given a particular set of circumstances. These factors primarily dealt with (1) call arrivals and holding-time distributions, (2) number of traffic sources, (3) availability (full or limited), and (4) handling of lost calls.

The Erlang B loss formula was/is very widely used outside of the United States. Loss in this context means the probability of encountering blockage at the switch due to congestion or to “all trunks busy” (ATB). The formula expresses grade of service or the probability of finding  $x$  channels busy. The other two factors in the Erlang B formula are the mean of the *offered traffic* and the number of trunks or servicing channels available. The formula assumes the following:

- Traffic originates from an infinite number of sources.
- Lost calls are cleared assuming a zero holding time.
- The number of trunks or servicing channels is limited.
- Full availability exists.

The actual Erlang B formula is out of the scope of this text. For more detailed information, it is recommended that the reader consult Ref. 3, Section 1. It is far less involved to use traffic tables as found in Table 4.1, which gives trunk-dimensioning information for some specific grades of service, from 0.001 to 0.05 and from 1 to 49 trunks. The table uses traffic-intensity units UC (unit call) and TU (traffic unit), where TU is in erlangs assuming BH conditions and UC is in ccs (cent-call-seconds). Remember that 1 erlang = 36 ccs (based on a 1-hour time interval).

To exemplify the use of Table 4.1, suppose a route carried 16.68 erlangs of traffic with a desired grade of service of 0.001; then 30 trunks would be required. If the grade of service were reduced to 0.05, the 30 trunks could carry 24.80 erlangs of traffic. When sizing a route for trunks or an exchange, we often come up with a fractional number

**Table 4.1 Trunk-Loading Capacity, Based on Erlang B Formula, Full Availability**

Trunks	Grade of Service 1 in 1000		Grade of Service 1 in 500		Grade of Service 1 in 200		Grade of Service 1 in 100		Grade of Service 1 in 50		Grade of Service 1 in 20	
	UC	TU	UC	TU	UC	TU	UC	TU	UC	TU	UC	TU
1	0.04	0.001	0.07	0.002	0.2	0.005	0.4	0.01	0.7	0.02	1.8	0.05
2	1.8	0.05	2.5	0.07	4	0.11	5.4	0.15	7.9	0.22	14	0.38
3	6.8	0.19	9	0.25	13	0.35	17	0.46	22	0.60	32	0.90
4	16	0.44	19	0.53	25	0.70	31	0.87	39	1.09	55	1.52
5	27	0.76	32	0.90	41	1.13	49	1.36	60	1.66	80	2.22
6	41	1.15	48	1.33	58	1.62	69	1.91	82	2.28	107	2.96
7	57	1.58	65	1.80	78	2.16	90	2.50	106	2.94	135	3.74
8	74	2.05	83	2.31	98	2.73	113	3.13	131	3.63	163	4.54
9	92	2.56	103	2.85	120	3.33	136	3.78	156	4.34	193	5.37
10	111	3.09	123	3.43	143	3.96	161	4.46	183	5.08	224	6.22
11	131	3.65	145	4.02	166	4.61	186	5.16	210	5.84	255	7.08
12	152	4.23	167	4.64	190	5.28	212	5.88	238	6.62	286	7.95
13	174	4.83	190	5.27	215	5.96	238	6.61	267	7.41	318	8.83
14	196	5.45	213	5.92	240	6.66	265	7.35	295	8.20	350	9.73
15	219	6.08	237	6.58	266	7.38	292	8.11	324	9.01	383	10.63
16	242	6.72	261	7.26	292	8.10	319	8.87	354	9.83	415	11.54
17	266	7.38	286	7.95	318	8.83	347	9.65	384	10.66	449	12.46
18	290	8.05	311	8.64	345	9.58	376	10.44	414	11.49	482	13.38
19	314	8.72	337	9.35	372	10.33	404	11.23	444	12.33	515	14.31
20	339	9.41	363	10.07	399	11.09	433	12.03	474	13.18	549	15.25
21	364	10.11	388	10.79	427	11.86	462	12.84	505	14.04	583	16.19
22	389	10.81	415	11.53	455	12.63	491	13.65	536	14.90	617	17.13
23	415	11.52	442	12.27	483	13.42	521	14.47	567	15.76	651	18.08
24	441	12.24	468	13.01	511	14.20	550	15.29	599	16.63	685	19.03
25	467	12.97	495	13.76	540	15.00	580	16.12	630	17.50	720	19.99
26	493	13.70	523	14.52	569	15.80	611	16.96	662	18.38	754	20.94
27	520	14.44	550	15.28	598	16.60	641	17.80	693	19.26	788	21.90
28	546	15.18	578	16.05	627	17.41	671	18.64	725	20.15	823	22.87
29	573	15.93	606	16.83	656	18.22	702	19.49	757	21.04	858	23.83
30	600	16.68	634	17.61	685	19.03	732	20.34	789	21.93	893	24.80
31	628	17.44	662	18.39	715	19.85	763	21.19	822	22.83	928	25.77
32	655	18.20	690	19.18	744	20.68	794	22.05	854	23.73	963	26.75
33	683	18.97	719	19.97	774	21.51	825	22.91	887	24.63	998	27.72
34	711	19.74	747	20.76	804	22.34	856	23.77	919	25.53	1033	28.70
35	739	20.52	776	21.56	834	23.17	887	24.64	951	26.43	1068	29.68
36	767	21.30	805	22.36	864	24.01	918	25.51	984	27.34	1104	30.66
37	795	22.03	834	23.17	895	24.85	950	26.38	1017	28.25	1139	31.64
38	823	22.86	863	23.97	925	25.69	981	27.25	1050	29.17	1175	32.63
39	851	23.65	892	24.78	955	26.53	1013	28.13	1083	30.08	1210	33.61
40	880	24.44	922	25.60	986	27.38	1044	29.01	1116	31.00	1246	34.60
41	909	25.24	951	26.42	1016	28.23	1076	29.89	1149	31.92	1281	35.59
42	937	26.04	981	27.24	1047	29.08	1108	30.77	1182	32.84	1317	36.58
43	966	26.84	1010	28.06	1078	29.94	1140	31.66	1215	33.76	1353	37.57
44	995	27.64	1040	28.88	1109	30.80	1171	32.54	1248	34.68	1388	38.56
45	1024	28.45	1070	29.71	1140	31.66	1203	33.43	1282	35.61	1424	39.55
46	1053	29.26	1099	30.54	1171	32.52	1236	34.32	1315	36.53	1459	40.54
47	1083	30.07	1129	31.37	1202	33.38	1268	35.21	1349	37.46	1495	41.54
48	1111	30.88	1159	32.20	1233	34.25	1300	36.11	1382	38.39	1531	42.54
49	1141	31.69	1189	33.04	1264	35.11	1332	37.00	1415	39.32	1567	43.54

of servicing channels or trunks. In this case we would opt for the next highest integer because we cannot install a fraction of a trunk. For instance, if calculations show that a trunk route should have 31.4 trunks, it would be designed for 32 trunks.

The Erlang B formula, based on lost calls cleared, has been standardized by the CCITT (CCITT Rec. Q.87) and has been generally accepted outside the United States. In the United States the Poisson formula is favored. This formula is often called the *Molina*

*formula*. It is based on the LCH concept. Table 4.2 provides trunking sizes for various grades of service deriving from the  $P$  formula; such tables are sometimes called “ $P$ ” tables (Poisson) and assume full availability. We must remember that the Poisson equation also assumes that traffic originates from a large (infinite) number of independent subscribers or sources (random traffic input), with a limited number of trunks or servicing channels and LCH (Ref. 3).

### 4.2.3 Waiting Systems (Queueing)

The North American PSTN became entirely digital by the year 2000. Nearly all digital switches operate under some form of queueing discipline, which many call *waiting systems* because an incoming call is placed in queue and waits its turn for service. These systems are based on our third assumption, namely, lost calls delayed (LCD). Of course, a queue in this case is a pool of callers waiting to be served by a switch. The term *servicing time* is the time a call takes to be served from the moment of arrival in the queue to the moment of being served by the switch. For traffic calculations in most telecommunication queueing systems, the mathematics is based on the assumption that call arrivals are random and Poissonian. The traffic engineer is given the parameters of offered traffic, the size of the queue, and a specified grade of service and will determine the number of serving circuits or trunks that are required.

The method by which a waiting call is selected to be served from the pool of waiting calls is called *queue discipline*. The most common discipline is the first-come, first-served discipline, where the call waiting longest in the queue is served first. This can turn out to be costly because of the equipment required to keep order in the queue. Another type is random selection, where the time a call has waited is disregarded and those waiting are selected in random order. There is also the last-come, first-served discipline and bulk service discipline, where batches of waiting calls are admitted, and there are also priority service disciplines, which can be preemptive and nonpreemptive. In queueing systems the grade of service may be defined as the probability of delay. This is expressed as  $P(t)$ , the probability that a call is not being immediately served and has to wait a period of time greater than  $t$ . The average delay on all calls is another parameter that can be used to express grade of service, and the length of queue is yet another.

The probability of delay, the most common index of grade of service for waiting systems when dealing with full availability and a Poissonian call arrival process (i.e., random arrivals), is calculated using the Erlang C formula, which assumes an infinitely long queue length. A more in-depth coverage of the Erlang C formula along with Erlang C traffic tables may be found in Ref. 3, Section 1.

### 4.2.4 Dimensioning and Efficiency

By definition, if we were to dimension a route or estimate the required number of servicing channel, where the number of trunks (or servicing channels) just equaled the erlang load, we would attain 100% efficiency. All trunks would be busy with calls all the time or at least for the entire BH. This would not even allow time for call setup (i.e., making the connection) or for switch processing time. In practice, if we sized our trunks, trunk routes, or switches this way, there would be many unhappy customers.

On the other hand, we do, indeed, want to dimension our routes (and switches) to have a high efficiency and still keep our customers relatively happy. The goal of our previous exercises in traffic engineering was just that. The grade of service is one measure of subscriber satisfaction. As an example, let us assume that between cities X and Y

**Table 4.2 Trunk-Loading Capacity, Based on Poisson Formula, Full Availability**

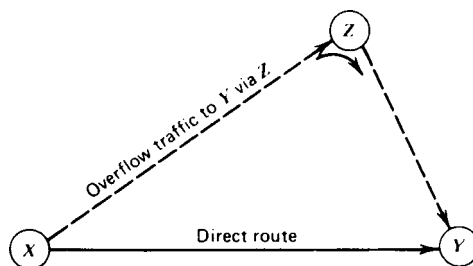
Trunks	Grade of Service 1 in 1000		Grade of Service 1 in 100		Grade of Service 1 in 50		Grade of Service 1 in 20		Grade of Service 1 in 10	
	UC	TU	UC	TU	UC	TU	UC	TU	UC	TU
1	0.1	0.003	0.4	0.01	0.7	0.02	1.9	0.05	3.8	0.10
2	1.6	0.05	5.4	0.15	7.9	0.20	12.9	0.35	19.1	0.55
3	6.9	0.20	16	0.45	20	0.55	29.4	0.80	39.6	1.10
4	15	0.40	30	0.85	37	1.05	49	1.35	63	1.75
5	27	0.75	46	1.30	56	1.55	71	1.95	88	2.45
6	40	1.10	64	1.80	76	2.10	94	2.60	113	3.15
7	55	1.55	84	2.35	97	2.70	118	3.25	140	3.90
8	71	1.95	105	2.90	119	3.30	143	3.95	168	4.65
9	88	2.45	126	3.50	142	3.95	169	4.70	195	5.40
10	107	2.95	149	4.15	166	4.60	195	5.40	224	6.20
11	126	3.50	172	4.80	191	5.30	222	6.15	253	7.05
12	145	4.05	195	5.40	216	6.00	249	6.90	282	7.85
13	166	4.60	220	6.10	241	6.70	277	7.70	311	8.65
14	187	5.20	244	6.80	267	7.40	305	8.45	341	9.45
15	208	5.80	269	7.45	293	8.15	333	9.25	370	10.30
16	231	6.40	294	8.15	320	8.90	362	10.05	401	11.15
17	253	7.05	320	8.90	347	9.65	390	10.85	431	11.95
18	276	7.65	346	9.60	374	10.40	419	11.65	462	12.85
19	299	8.30	373	10.35	401	11.15	448	12.45	492	13.65
20	323	8.95	399	11.10	429	11.90	477	13.25	523	14.55
21	346	9.60	426	11.85	458	12.70	507	14.10	554	15.40
22	370	10.30	453	12.60	486	13.50	536	14.90	585	16.25
23	395	10.95	480	13.35	514	14.30	566	15.70	616	17.10
24	419	11.65	507	14.10	542	15.05	596	16.55	647	17.95
25	444	12.35	535	14.85	572	15.90	626	17.40	678	18.85
26	469	13.05	562	15.60	599	16.65	656	18.20	710	19.70
27	495	13.75	590	16.40	627	17.40	686	19.05	741	20.60
28	520	14.45	618	17.15	656	18.20	717	19.90	773	21.45
29	545	15.15	647	17.95	685	19.05	747	20.75	805	22.35
30	571	15.85	675	18.75	715	19.85	778	21.60	836	23.20
31	597	16.60	703	19.55	744	20.65	809	22.45	868	24.10
32	624	17.35	732	20.35	773	21.45	840	23.35	900	25.00
33	650	18.05	760	21.10	803	22.30	871	24.20	932	25.90
34	676	18.80	789	21.90	832	23.10	902	25.05	964	26.80
35	703	19.55	818	22.70	862	23.95	933	25.90	996	27.65
36	729	20.25	847	23.55	892	24.80	964	26.80	1028	28.55
37	756	21.00	876	24.35	922	25.60	995	27.65	1060	29.45
38	783	21.75	905	25.15	951	26.40	1026	28.50	1092	30.35
39	810	22.50	935	25.95	982	27.30	1057	29.35	1125	31.25
40	837	23.25	964	26.80	1012	28.10	1088	30.20	1157	32.14
41	865	24.05	993	27.60	1042	28.95	1120	31.10	1190	33.05
42	892	24.80	1023	28.40	1072	29.80	1151	31.95	1222	33.95
43	919	25.55	1052	29.20	1103	30.65	1183	32.85	1255	34.85
44	947	26.30	1082	30.05	1133	31.45	1214	33.70	1287	35.75
45	975	27.10	1112	30.90	1164	32.35	1246	34.60	1320	36.65
46	1003	27.85	1142	31.70	1194	33.15	1277	35.45	1352	37.55
47	1030	28.60	1171	32.55	1225	34.05	1309	36.35	1385	38.45
48	1058	29.40	1201	33.35	1255	34.85	1340	37.20	1417	39.35
49	1086	30.15	1231	34.20	1286	35.70	1372	38.10	1450	40.30
50	1115	30.95	1261	35.05	1317	36.60	1403	38.95	1482	41.15

there were 47 trunks on the interconnecting telephone route. The tariffs, from which the telephone company derives revenue, are a function of the erlangs of carried traffic. Suppose we allow \$1.00 per erlang-hour. The very upper limit of service on the route is 47 erlangs, and the telephone company would earn \$47 for the busy hour (much less for all other hours) for that trunk route and the portion of the switches and local plant involved with these calls. As we well know, many of the telephone company's subscribers would be unhappy because they would have to wait excessively to get calls through from X to Y. How, then, do we optimize a trunk route (or serving circuits) and keep the customers as satisfied with service as possible?

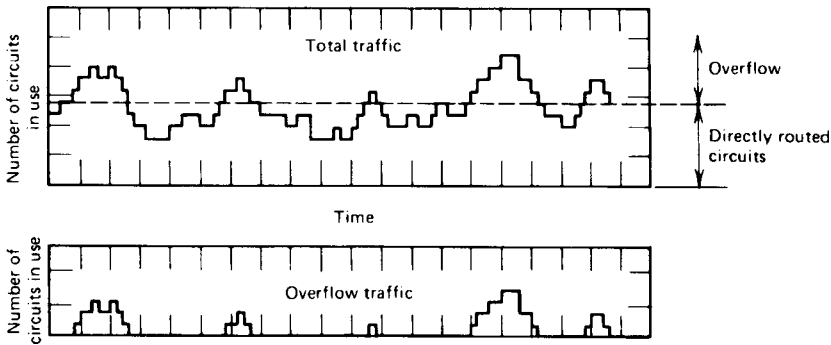
Remember from Table 4.1, with an excellent grade of service of 0.001, that we relate grade of service to subscriber satisfaction (one element of quality of service) and that 47 trunks could carry 30.07 erlangs during the busy hour. Assuming the route did carry 30.07 erlangs, let's say at \$1.00 per erlang, it would earn \$30.07 for that hour. From a revenue viewpoint, that would be the best hour of the day. If the grade of service were reduced to 0.01, 47 trunks would bring in \$35.21 (i.e., 35.21 erlangs) for the busy hour. Note the improvement in revenue at the cost of reducing grade of service.

Here we are relating efficiency on trunk utilization. Trunks not carrying traffic do not bring in revenue. If we are only using some trunks during the busy hour only minutes a day to cover BH traffic peaks, the remainder of the day they are not used. That is highly inefficient. As we reduce the grade of service, the trunk utilization factor improves. For instance, 47 trunks will only carry 30.07 erlangs with a grade of service of 1 in 1000 (0.001), whereas if we reduce the grade of service to 1 in 20 (0.05), we carry 41.54 erlangs (see Table 4.1). Efficiency has improved notably. Quality of service, as a result, has decreased markedly.

**4.2.4.1 Alternative Routing.** One method to improve efficiency is to use alternative routing (called *alternate routing* in North America). Suppose we have three serving areas, X, Y, and Z, served by three switches (exchanges), X, Y, and Z, as illustrated in Figure 4.4. Let the grade of service be 0.005 (1 in 200 in Table 4.1). We find that it would require 48 trunks to carry 34.25 erlangs of traffic during the BH to meet that grade of service between X and Y. Suppose we reduce the number of trunks between X and Y, still keeping the BH traffic intensity at 34.25 erlangs. We would thereby increase efficiency on the X–Y route at the cost of reducing grade of service. With a modification of the switch at X, we could route traffic bound for Y that met congestion on the X–Y route via switch Z. Then Z would route that traffic on the Z–Y link. Essentially this is alternative routing in its simplest form. Congestion would probably only occur during very short peaking periods in the BH, and chances are that these peaks would not occur simultaneously with peaks



**Figure 4.4** Simplified diagram of the alternative (alternate) routing concept. (Solid line represents the direct route, dashed lines represent the alternative route carrying the overflow traffic from X to Y).



**Figure 4.5** Traffic peakedness, the peaks are carried on alternative routes.

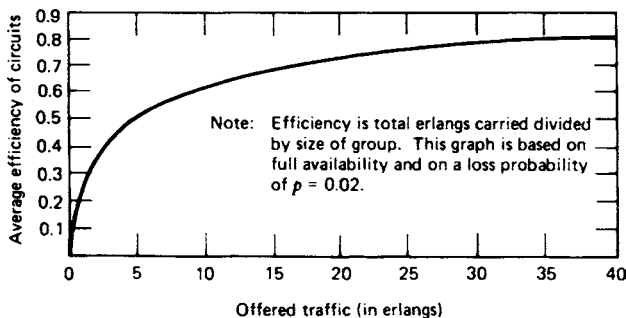
of traffic intensity on the Z–Y route. Furthermore, the incremental load on the X–Z–Y route would be very small. The concept of traffic peakedness that would overflow onto the secondary (X–Z–Y) is shown in Figure 4.5.

**4.2.4.2 Efficiency Versus Circuit Group Size.** In the present context a *circuit group* refers to a group of circuits performing a specific function. For instance, all the trunks (circuits) routed from X to Y in Figure 4.4 make up a circuit group irrespective of size. This circuit group should not be confused with the “group” used in transmission engineering of carrier systems.<sup>6</sup>

If we assume full loading, we find that efficiency improves with circuit group size. From Table 4.1, given a grade of service of 1 in 100, 5 erlangs of traffic require a group with 11 trunks, more than 2:1 ratio of trunks to erlangs, and 20 erlangs requires 30 trunks, a 3:2 ratio. If we extend this to 100 erlangs, 120 trunks are required, a 6:5 ratio. Figure 4.6 shows how efficiency improves with group size.

**4.2.5 Quantifying Data Traffic**

Data traffic usually consists of short, bursty transactions from a few milliseconds duration to several seconds, depending on the data transmission rate (i.e., the number of bits per second). This is particularly true on local area networks (LANs). As the data rate slows down, such as we might find on a wide area network (WAN), transaction time increases,



**Figure 4.6** Group efficiency increases with size.

<sup>6</sup>Carrier systems are frequency-division multiplex systems introduced in Section 4.5.

possibly to a minute or so. For these reasons, it is dangerous to apply speech telephony traffic theory and practice to the data environment.

There is an exception here—that is, when a data protocol specifies a permanent virtual circuit (PVC). This is a circuit that is set up in advance for one or several data transactions. One group of traffic engineers has proposed the milli-erlang for LAN and PVC applications. We think this idea bears merit.

### 4.3 INTRODUCTION TO SWITCHING

In this section our concern is telephone switching, the switching of voice channels. We will deal with some switching concepts and with several specifics. Switching was defined in Section 4.1 in contraposition with transmission.

Actual connectivity is carried out by the switching function. A connectivity may involve more than one switch. As we pointed out in Chapter 1, there are local switches, tandem switches, and *transit switches*. A transit switch is just a tandem switch that operates in the long distance or “toll” service.

A local switch has an area of responsibility. We call this its *serving area*. All subscriber loops in a serving area connect to that switch responsible for the area. Many calls in a local area traverse no more than one switch. These are calls to neighbors. Other calls, destined for subscribers outside of that serving area, may traverse a tandem switch from there to another local serving switch if there is no direct route available. If there is a direct route, the tandem is eliminated for that *traffic relation*. It is unnecessary.

Let us define a *traffic relation* as a connectivity between exchange A and B. The routing on calls for that traffic relation is undetermined. Another connotation for the term *traffic relation* implies that there would be not only a connectivity capability, but also the BH traffic expected on that connectivity.

To carry out these functions, a switch had to have some sort of intelligence. In a manually operated exchange, the intelligence was human, namely, the telephone operator. The operator was replaced by an automatic switch. Prior to the computer age, a switch’s intelligence was “hard-wired” and its capabilities were somewhat limited. Today, all modern switches are computer-based and have a wide selection of capabilities and services. Our interest here is in the routing of a call. A switch knows how to route a call through the dialed telephone number as we described in Section 1.3.2. There we showed that a basic telephone number consists of seven digits. The last four digits identify the subscriber; the first three digits identify the local serving exchange responsible for that subscriber. The three-digit exchange code is unique inside of an area code. In North America, an area code is a three-digit number identifying a specific geographical area. In many countries, if one wishes to dial a number that is in another area code, an access code is required. In the United States that access code is a 1.

#### 4.3.1 Basic Switching Requirements

Conceptually, consider that a switch has inlets and outlets. Inlets serve incoming calls; outlets serve outgoing calls. A call from a calling subscriber enters an exchange through an inlet. It connects to a called subscriber through an outlet. There are three basic switching requirements:

1. An exchange (a switch) must be able to connect any incoming call to one of a multitude of outgoing circuits.

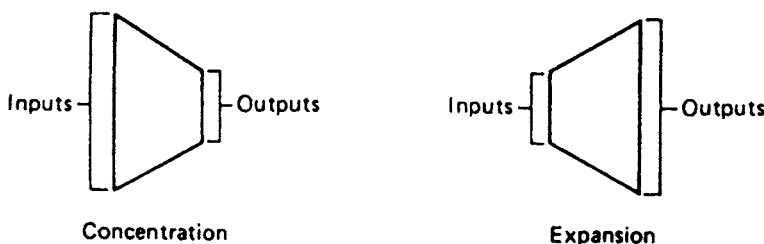
2. It has the ability not only to establish and maintain (or *hold*) a physical connection between a caller and the called party for the duration of the call, but also to be able to disconnect (i.e., “clear”) it after call termination.
3. It also has the ability to prevent new calls from intruding into circuits that are already in use. To avoid this, a new call must be diverted to another circuit that is free or it must be temporarily denied access where the caller will hear a “busy back” (i.e., a tone cadence indicating that the line is busy) or an “all trunks busy” tone cadence signal or voice announcement (i.e., indicating congestion or blockage).

Let’s differentiate local and tandem/transit exchanges. A local exchange connects lines (subscriber loops) to other lines or to trunks. A tandem/transit exchange switches trunks. Local exchanges *concentrate* and *expand*. Tandem and transit exchanges do not.

### 4.3.2 Concentration and Expansion

Trunks are expensive assets. Ideally, there should be one trunk available for every subscriber line (loop). Then there never would be a chance of blockage. Thus, whenever a subscriber wished to connect to a distant subscriber, there would be a trunk facility available for that call. Our knowledge of telephone calling habits of subscribers tells us that during the busy hour, on the order of 30% of subscriber lines will be required to connect to trunks for business customers and some 10% for residential customers. Of course, these values are rough estimates. We’d have to apply the appropriate traffic formula based on a grade of service, as described in Section 4.2.1, for refined estimates.

Based on these arguments, a local exchange serving residential customers might have 10,000 lines, and only 1000 trunks would be required. This is concentration. Consider that those 1000 incoming trunks to that exchange must expand out to 10,000 subscribers. This is expansion. It provides all subscribers served by the switch with access to incoming trunks and local switching paths. The concentration/expansion concept of a local serving exchange is illustrated in the following diagram:



### 4.3.3 The Essential Functions of a Local Switch

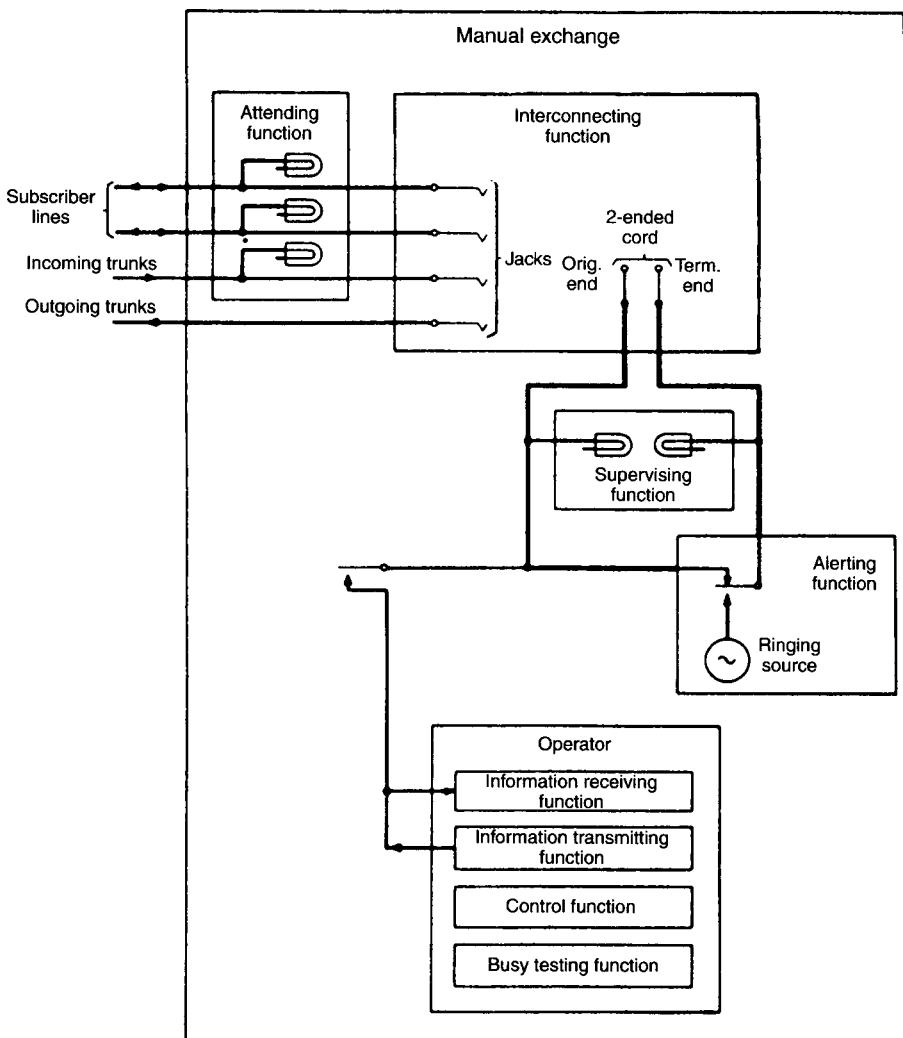
As we mentioned above, means are provided in a local switch to connect each subscriber line to any other in the same exchange. In addition, any incoming trunk must be able to connect to any subscriber line and any subscriber to any outgoing trunk.<sup>7</sup> These switching functions are remotely controlled by the calling subscriber, whether she/he is a local subscriber or long-distance subscriber. These remote instructions are transmitted to the switch (exchange) by “off-hook,” “on-hook,”<sup>8</sup> and dial information. There are eight basic functions that must be carried out by a conventional switch or exchange

<sup>7</sup>The statement assumes *full availability*.

<sup>8</sup>*Off-hook* and *on-hook* are defined in Section 1.3.1.

1. Interconnection
2. Control
3. Alerting
4. Attending
5. Information receiving
6. Information transmitting
7. Busy testing
8. Supervisory

Consider a typical manual switching center illustrated in Figure 4.7. Here the eight basic functions are carried on for each call. The important interconnection function is illustrated by the jacks appearing in front of the operator. There are subscriber-line jacks<sup>9</sup> and jacks



**Figure 4.7** A manual exchange illustrating switching functions.

<sup>9</sup>A jack is an electric receptacle. It is a connecting device, ordinarily employed in a fixed location, to which a wire or wires may be attached, and it is arranged for the insertion of a plug.

for incoming and outgoing trunks. The connection is made by double-ended connecting cords, which can connect subscriber to subscriber or subscriber to trunk. The cords available are always less than half the number of jacks appearing on the board, because one interconnecting cord occupies two jacks, one on either end. Concentration takes place at this point on a manual exchange. Distribution is also carried out because any cord may be used to complete a connection to any of the terminating jacks. The operator is *alerted* by a lamp becoming lit when there is an incoming call requiring connection. This is the *attending–alerting* function. The operator then assumes the *control* function, determining an idle connecting cord and plugging it into the incoming jack. She/he then determines call destination, continuing her/his control function by plugging the cord into the terminating jack of the called subscriber or proper trunk to terminate her/his portion of control of the incoming call. Of course, before plugging into the terminating jack, she/he carries out a *busy test* function to determine that the called line or trunk is not busy. To alert the called subscriber that there is an incoming call, she/he uses the manual ring-down<sup>10</sup> by connecting the called line to a ringing current source as illustrated in Figure 4.7.

Other signaling means are used for trunk signaling if the incoming call is destined for another exchange. On such a call the operator performs the information function orally or by dialing the call information to the next exchange in the routing.

The *supervision function* is performed by lamps to show when a call is completed and the call is taken down (i.e., the patch cord can be removed). The operator conducts numerous control functions to set up a call, such as selecting a cord, plugging it into the originating jack of the calling line, connecting her/his headset to determine calling information, selecting (and busy testing) the called subscriber jack, and then plugging the other end of the cord into the proper terminating jack and alerting the called subscriber by ring-down. Concentration is the ratio of the field of incoming jacks to cord positions. Expansion is the number of cord positions to outgoing (terminating) jacks. The terminating and originating jacks can be interchangeable. The called subscriber at one moment in time can become the calling subscriber at another moment in time. On the other hand, incoming and outgoing trunks may be separated. In this case they would be one-way circuits. If not separated, they would be both-way circuits, accepting both incoming and outgoing traffic.

#### 4.3.4 Introductory Switching Concepts

All local telephone switches have, as a minimum, three functional elements: concentration, distribution, and expansion. Concentration and expansion were discussed in Section 4.3.3. Viewing a switch another way, we can say that it has *originating line appearances* and *terminating line appearances*. These are illustrated in a simplified conceptual drawing in Figure 4.8, which shows three different call possibilities of a typical local exchange:

1. A call originated by a subscriber who is served by the exchange and bound for a subscriber who is served by the same exchange (route A-B-C-D-E).
2. A call originated by a subscriber who is served by the exchange and bound for a subscriber who is served by another exchange (route A-B-F).
3. A call originated by a subscriber who is served by another exchange and bound for a subscriber served by the exchange in question (route G-D-E).

<sup>10</sup>Ring-down is a method of signaling to alert an operator or a distant subscriber. In old-time telephone systems, a magneto was manually turned, thereby generating an alternating current that would ring a bell at the other end. Today, special ringing generators are used.

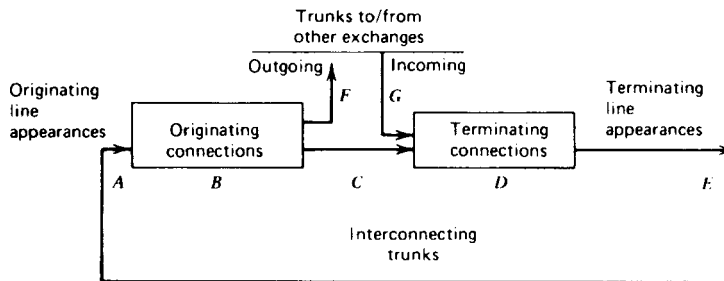


Figure 4.8 Originating and terminating line appearances.

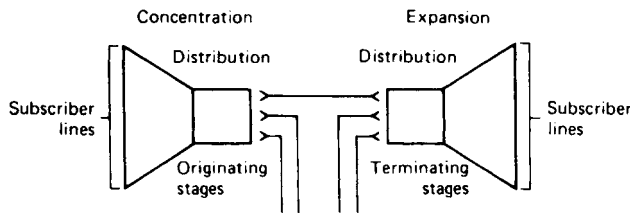


Figure 4.9 The concept of distribution.

Call concentration takes place in B and call expansion at D. Figure 4.9 is simply a redrawing of Figure 4.8 to show the concept of *distribution*. The distribution stage in switching serves to connect by switching the concentration stage to the expansion stage.

### 4.3.5 Early Automatic Switching Systems

**4.3.5.1 Objective.** We summarize several earlier, space division switching systems because of the concepts involved. Once the reader grasps these concepts, the ideas and notions of digital switching will be much easier to understand. First, the operation of the original step-by-step switch is described. This is followed by a discussion of the crossbar switch.

**4.3.5.2 The Step-by-Step Switch.** The step-by-step (SXS) switch use was widespread in the United States prior to 1950, when the crossbar switch tended to replace it. Its application was nearly universal in the United Kingdom, where it was called the *Strowger switch*.

The step-by-step switch has a curious history. Its inventor was Almon B. Strowger, an undertaker in Kansas City. Strowger suspected that he was losing business because the town's telephone operator was directing all requests for funeral services to a competitor, which some say was a boyfriend, others say was a relative. We do not know how talented Strowger was as a mortician, but he certainly goes down in history for his electromechanical talents for the invention of the automatic telephone switch. The first "step" switch was installed in Indiana in 1892. They were popular with independent telephone companies, but installation in AT&T's Bell System did not start until 1911. The step-by-step switch is conveniently based on a stepping relay of 10 levels. In its simplest form, which uses direct progressive control, dial pulses from a subscriber's telephone activate the switch. For example, if a subscriber dials a 3, three pulses from the subscriber subset are transmitted to the switch. The switch then steps to level 3 in the first relay bank. The second relay bank is now connected waiting for the second dialed digit. It accepts the second

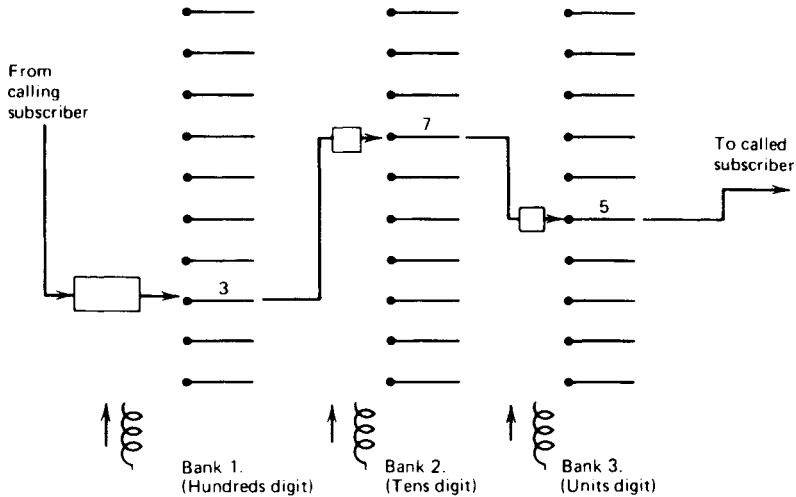


Figure 4.10 Conceptual operation of a step-by-step switch (exchange).

digit from the subscriber and steps to its equivalent position and connects to the third relay bank and so on for four or seven dialed digits. Assume that a certain exchange only serves three-digit numbers. A dialed number happens to be 375 and will be stepped through three sets of banks of 10 steps each. This is conceptually illustrated in Figure 4.10.

**4.3.5.3 The Crossbar Switch.** Crossbar switching dates back to 1938 and reached a peak of installed lines in 1983. Its life had been extended by using stored program control (SPC)<sup>11</sup> rather than hard-wire control in the more conventional crossbar configuration. The crossbar is actually a matrix switch used to establish the speech path. An electrical contact is made by actuating a horizontal and vertical relay. Consider the switching mechanism illustrated in Figure 4.11. To make contact at point B<sub>4</sub> on the matrix, horizontal relay B and vertical relay 4 must close to establish the connection. Such closing is usually momentary, but sufficient to cause *latching*. Two forms of latching are found in crossbar practice: mechanical and electrical. The latch keeps the speech path connection until an “on-hook” condition occurs. Once the latching occurs, connection B<sub>4</sub> is “busied out,”<sup>12</sup> and the horizontal and vertical relays are freed-up to make other connections for other calls.

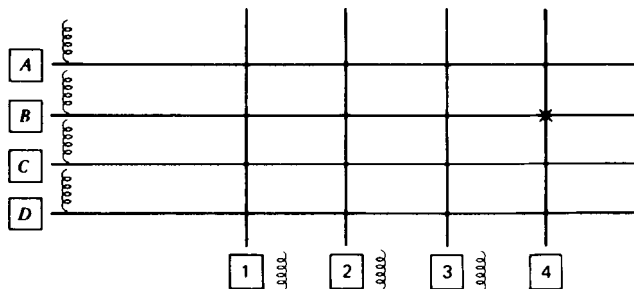


Figure 4.11 The crossbar concept.

<sup>11</sup>SPC, stored program control, simply means that the switch or exchange is computer-controlled. Of course, all modern digital switches are computer-controlled.

<sup>12</sup>“Busy-out” means that a line or connection is taken out of the pool because it is busy, it is being used, and is not available for others to utilize.

### 4.3.6 Common Control (Hard-Wired)

First, we must distinguish *common control* from *direct progressive control* described in Section 4.3.5.2. With direct progressive control a subscriber dialed a digit, and the first relay bank stepped to the dialed digit; the subscriber dialed a second digit, and the second stepping relay bank actuated, stepping to that digit level, and so on, through the entire dialed number. With common control, on the other hand, the dialed number is first stored in a register.<sup>13</sup> These digits are then analyzed and acted upon by a *marker*, which is a hard-wired processor. Once the call setup is complete, the register and marker are free to handle other call setups. The marker was specifically developed for the crossbar switch. Such marker systems are most applicable to specialized crossbar switching matrices of crossbar switches. Stored program control (SPC) is a direct descendent of the crossbar common control system. SPC is described below.

### 4.3.7 Stored Program Control

**4.3.7.1 Introduction.** *Stored program control* (SPC) is a broad term designating switches where common control is carried out entirely by computer. In some exchanges, this involves a large, powerful computer. In others, two or more minicomputers may carry out the SPC function. Still with other switches, the basic switch functions are controlled by distributed microprocessors. Software may be hard-wired on one hand or programmable on the other. There is a natural marriage between a binary digital computer and the switch control functions. In most cases these also work in the binary digital domain. The crossbar markers and registers are typical examples.

The conventional crossbar marker requires about half a second to service a call. Up to 40 expensive markers are required on a large exchange. Strapping points on the marker are available to laboriously reconfigure the exchange for subscriber change, new subscribers, changes in traffic patterns, reconfiguration of existing trunks or their interface, and so on.

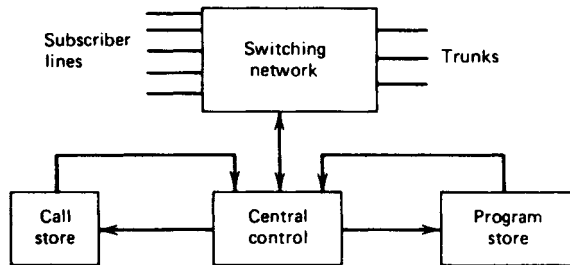
Replacing register markers with programmable logic—a computer, if you will—permits one device to carry out the work of 40. A simple input sequence on the keyboard of the computer workstation replaces strapping procedures. System faults are displayed as they occur, and circuit status may be indicated on the screen periodically. Due to the high speed of the computer, postdial delay is reduced. SPC exchanges permit numerous new service offerings, such as conference calls, abbreviated dialing, “camp-on-busy,” call forwarding, voice mail, and call waiting.

**4.3.7.2 Basic SPC Functions.** There are four basic functional elements of an SPC switching system:

1. Switching matrix
2. Call store (memory)
3. Program store (memory)
4. Central processor (computer)

The earlier switching matrices consisted of electromechanical cross-points, such as a crossbar matrix, reed, correed, or ferreed cross-points. Later switching matrices employed solid-state cross-points.

<sup>13</sup>A *register* is a device that receives and stores signals; in this particular case, it receives and stores dialed digits.



**Figure 4.12** A simplified functional diagram of an SPC exchange.

The *call store* is often referred to as the “scratch-pad” memory. This is temporary storage of incoming call information ready for use, on command from the central processor. It also contains availability and status information of lines, trunks, and service circuits under internal switch-circuit conditions. Circuit status information is brought to the memory by a method of scanning. All speech circuits are scanned for a busy/idle condition.

The *program store* provides basic instructions to the controller (central processor). In many installations, translation information is held in this store (memory), such as DN to EN translation and trunk signaling information.

A simplified functional diagram of a basic SPC system is shown in Figure 4.12.

#### 4.3.8 Concentrators and Remote Switching

In Chapter 5 we discuss the design of a subscriber loop. There we will find that there are very definite length limitations on subscriber loops. As we delve further into subscriber loop design, methods of extending loops still further are described. One way to extend such loops is with a remote concentrator or switch.

The simplest form of extending a switch is to use a concentrator some distance from the switch (exchange). *Concentrators* or *line concentrators* consolidate subscriber loops, are remotely operated, and are a part of the concentration and expansion portion of a switch placed at a remote location. The concentrator may be based on electromechanical facilities or solid-state cross-points for the concentration matrix. For instance, a 10:1 concentrator might serve 100 subscriber loops and deliver 10 trunks to the “mother” exchange. A concentrator does no switching whatsoever. All switching is carried out at the controlling or “mother” exchange. A typical line concentrator is illustrated in Figure 4.13, where 100 subscriber loops are consolidated to 20 trunks plus 2 trunks for control from the nearby “mother” exchange. Of course, the ratio of loops to trunks is a key issue, and it is based on calling habits and whether the subscribers are predominantly business or residential.

A remote switch, sometimes called a satellite, or satellite exchange, originates and terminates calls from the parent exchange. It differs from a concentrator in that local calls (i.e., calls originating and terminating inside the same satellite serving area) are served by the remote switch and do not have to traverse the parent exchange as remote concentrator calls do. A block of telephone numbers is assigned to the satellite serving area and is usually part of the basic number block assigned to the parent exchange. Because of the numbering arrangement, a satellite exchange can discriminate between local calls and calls to be handled by the parent exchange. A satellite exchange can be regarded as a component of the parent exchange that has been dislocated and moved to a distant site. The use of remote switching is very common in rural areas, and the distance a remote switch is from the parent exchange can be as much as 100 miles (160 km). Satellite exchanges range in size from 300 to 2000 lines. Concentrators are cost effective for 300 or less subscribers. However, AT&T’s SLC-96 can serve 1000 subscribers or more.

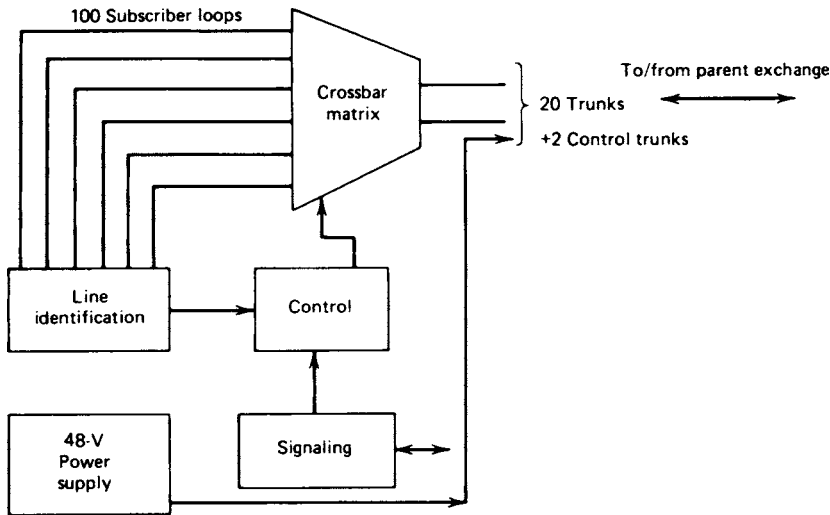


Figure 4.13 A typical concentrator.

## 4.4 ESSENTIAL CONCEPTS IN TRANSMISSION

### 4.4.1 Introduction

In this section we discuss two-wire and four-wire transmission and two impairments that are commonly caused by two-wire-to-four-wire conversion equipment. These impairments are *echo* and *singing*. The second part of this section is an introduction to multiplexing. *Multiplexing* allows two or more communication channels to share the same transmission bearer facility.

### 4.4.2 Two-Wire and Four-Wire Transmission

**4.4.2.1 Two-Wire Transmission.** A telephone conversation inherently requires transmission in both directions. When both directions are carried on the same pair of wires, it is called *two-wire transmission*. The telephones in our homes and offices are connected to a local switching center (exchange) by means of two-wire circuits. A more proper definition for transmitting and switching purposes is that when oppositely directed portions of a single telephone conversation occur over the same electrical transmission channel or path, we call this *two-wire operation*.

**4.4.2.2 Four-Wire Transmission.** Carrier and radio systems require that oppositely directed portions of a single conversation occur over separate transmission channels or paths (or use mutually exclusive time periods). Thus we have two wires for the transmit path and two wires for the receive path, or a total of four wires, for a full-duplex (two-way) telephone conversation. For almost all operational telephone systems, the end instrument (i.e., the telephone subset) is connected to its intervening network on a two-wire basis.

Nearly all long-distance (toll) telephone connections traverse four-wire links. From the near-end user the connection to the long-distance network is two-wire or via a two-wire link. Likewise, the far-end user is also connected to the long-distance (toll) network via a two-wire link. Such a long-distance connection is shown in Figure 4.14. Schematically, the four-wire interconnection is shown as if it were a single-channel wire-line

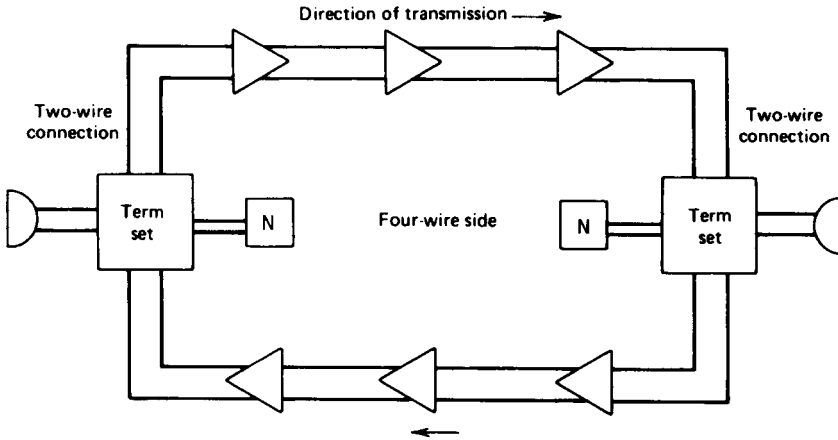


Figure 4.14 A typical long distance (toll) connection model.

with amplifiers. However, it would more likely be a multichannel multiplexed configuration on wire/fiber-optic cable or over radio. Nevertheless, the amplifiers in Figure 4.14 serve to convey the ideas this section considers. As illustrated in Figure 4.14, conversion from two-wire to four-wire operation is carried out by a *hybrid*, which is a four-port, four-winding transformer.

**4.4.2.3 Operation of a Hybrid.** A hybrid, in terms of telephony (at voice frequency), is a transformer with four separate windings. Based on a simplified description, a hybrid may be viewed as a power splitter with four sets of wire-pair connections. A functional block diagram of a hybrid device is shown in Figure 4.15. Two of the wire-pair connections belong to the four-wire path, which consists of a transmit pair and a receive pair. The third pair is a connection to the two-wire link, which is eventually connected to the subscriber subset via one or more switches. The last pair of the four connects the hybrid to a resistance–capacitance balancing network, which electrically balances the hybrid with the two-wire connection to the subscriber subset over the frequency range of the balancing network. *Balancing*, in this context, means matching impedances—that is, the impedance of the two-wire side to the hybrid two-wire port.

Signal energy entering from the two-wire subset connection divides equally. Half of it dissipates (as heat) in the impedance of the four-wire side receive path and the other half goes to the four-wire side transmit path, as illustrated in Figure 4.15. Here the *ideal* situation is that no energy is to be dissipated by the balancing network (i.e., there is

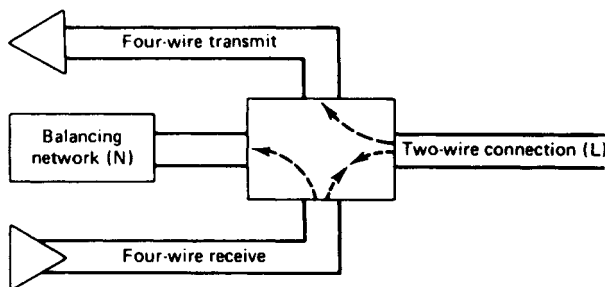


Figure 4.15 Operation of a hybrid transformer.

a perfect balance or impedance match). The balancing network is supposed to display the characteristic impedance of the two-wire line (subscriber connection) to the hybrid.<sup>14</sup> Signal energy entering from the four-wire side receive path is also split in half in the ideal situation where there is a perfect balance (i.e., a perfect match). Half of the energy is dissipated by the balancing network (N) and half at the two-wire port (L) (see Figure 4.15.)

The reader should note that in the description of a hybrid, in every case, ideally half of the signal energy entering the hybrid is used to advantage and half is dissipated or wasted. Also keep in mind that any passive device inserted in a circuit, such as a hybrid, has an insertion loss. As a rule of thumb, we say that the insertion loss of a hybrid is 0.5 dB. Thus there are two losses here that the reader must not lose sight of:

Hybrid insertion loss	0.5 dB	
Hybrid dissipation loss	<u>3.0 dB</u>	(half of the power)
	3.5 dB	(total)

As far as this section is concerned, any signal passing through a hybrid suffers a 3.5-dB loss. This is a good design number for gross engineering practice. However, some hybrids used on short subscriber connections purposely have higher losses, as do special resistance-type hybrids.

In Figure 4.15, consider the balancing network (N) and the two-wire side of the hybrid (L). In all probability (L), the two-wire side will connect to a subscriber through at least one switch. Thus the two-wire port on the hybrid could look into at least 10,000 possible subscriber connections: some short loops, some long loops, and some loops in poor condition. Because of the fixed conditions on the four-wire side, we can generally depend on holding a good impedance match. Our concern under these conditions is the impedance match on the two-wire side—that is, the impedance match between the compromise network (N) and the two-wire side (L). Here the impedance can have high variability from one subscriber loop to another.

We measure the capability of impedance match by *return loss*. In this particular case we call it *balance return loss*:

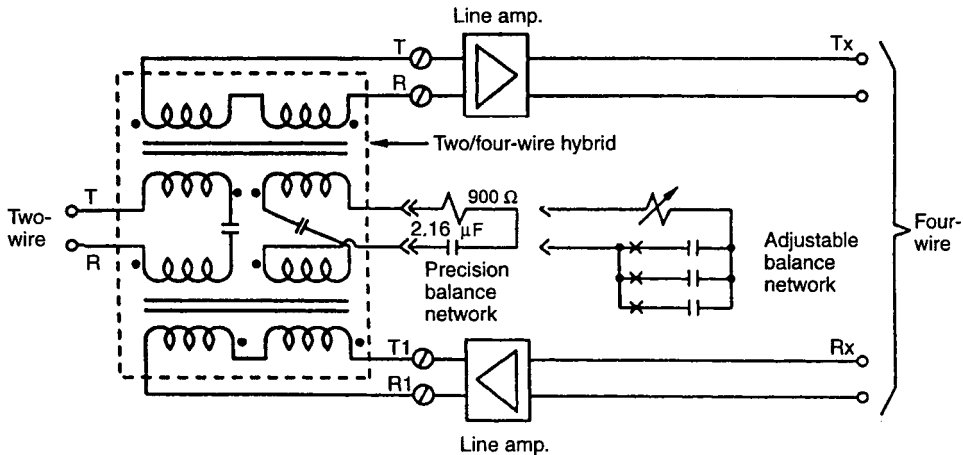
$$\text{Balance return loss}_{\text{dB}} = 20 \log_{10} \frac{Z_L + Z_N}{Z_L - Z_N}.$$

Let us say, for argument's sake, that we have a perfect match. In other words, the impedance of the two-wire subscriber loop side (L) on this particular call was exactly 900  $\Omega$  and the balancing network (N) was 900  $\Omega$ . Substitute these numbers in the preceding formula above and we get

$$\text{Balance return loss}_{\text{dB}} = 20 \log \frac{900 + 900}{900 - 900}.$$

Examine the denominator. It is zero. Any number divided by zero is infinity. Thus we have an infinitely high return loss. And this happens when we have a perfect match, an ideal condition. Of course it is seldom realized in real life. In real life we find that the balance return loss for a large population of hybrids connected in service and serving a large population of two-wire users has a median more on the order of 11 dB with a

<sup>14</sup>*Characteristic impedance* is the impedance that the line or port on a device is supposed to display. For most subscriber loops it is 900  $\Omega$  with a 2.16- $\mu\text{F}$  capacitor in series at 1000 Hz for 26-gauge wire pair or 600- $\Omega$  resistive. The notation for characteristic impedance is  $Z_0$ .



**Figure 4.16** Schematic diagram of two-wire to four-wire conversion using a hybrid. (From Figure 5-9, p. 104, Ref. 5. Reprinted with permission of the IEEE Press.)

standard deviation of 3 dB (Ref. 4). This is valid for North America. For some other areas of the world, balance return loss median may be lower with a larger standard deviation.

When the return loss becomes low (i.e., there is a poor impedance match), there is a reflection of the speech signal. That is, speech energy from the talker at her/his distant hybrid leaks across from the four-wire receive to the four-wire transmit side (see Figure 4.15). This signal energy is heard by the talker. It is delayed due to the propagation time. This is echo, which can be a major impairment depending on its intensity and amount of time it is delayed. It can also be very disruptive on a data circuit.

We define the cause of echo as any impedance mismatch in the circuit. It is most commonly caused by this mismatch that occurs at the hybrid. Echo that is excessive becomes singing. *Singing* is caused by high positive feedback on the intervening amplifiers (Figure 4.14). Singing on the analog network could take the network down by overloading multiplex equipment. The possibility of singing on the digital network is very low. The control of echo and singing is discussed in Chapter 8. Figure 4.16 is a schematic diagram of a hybrid circuit.

## 4.5 INTRODUCTION TO MULTIPLEXING

### 4.5.1 Definition

*Multiplexing* is used for the transmission of a plurality of information channels over a single transmission medium. An information channel may be a telephone voice channel, data channel, or a channel carrying image information. Our discussion below will concentrate on a telephone channel. A *telephone channel* is a channel optimized for carrying voice traffic, in this case the voice of a single telephone user. We will define it as an analog channel with occupying the band of frequencies between 300 and 3400 Hz (CCITT/ITU-T definition).

Before launching into our discussion, keep in mind that all multiplex equipment is four-wire equipment. If we look at one side of a circuit, there will be a multiplexer used for transmission and a demultiplexer used for reception.

The number of channels that can be multiplexed on a particular circuit depends on the bandwidth of the transmission medium involved. We might transmit 24 or 48 or

96 channels on a wire pair, depending on the characteristics of that wire pair. Coaxial cable can support many thousands of voice channels; line-of-sight microwave radio is capable of carrying from several hundred to several thousand voice channels. A single fiber-optic thread can support literally tens of thousand of channels. A communication satellite transponder can carry between 700 and 2000 such voice channels, depending on the transponder's bandwidth.

There are essentially two generic methods of multiplexing information channels:

1. In the frequency domain; we call this frequency division multiplex (FDM).
2. In the time domain, which we call time division multiplex (TDM).

The concepts of frequency division multiplexing are discussed in this chapter. Time division multiplexing (pulse code modulation) is covered in Chapter 6.

## 4.5.2 Frequency Division Multiplex

**4.5.2.1 Introduction.** With FDM the available channel bandwidth is divided into a number of nonoverlapping frequency slots. Each frequency slot or bandwidth segment carries a single information-bearing signal such as a voice channel. We can consider an FDM multiplexer as a frequency translator. At the opposite end of the circuit, a demultiplexer filters and translates the frequency slots back into the original information bearing channels. In the case of a telephone channel, a frequency slot is conveniently 4 kHz wide, sufficient to accommodate the standard 300- to 3400-kHz voice channel. Figure 4.17 illustrates the basic concept of frequency division multiplex.

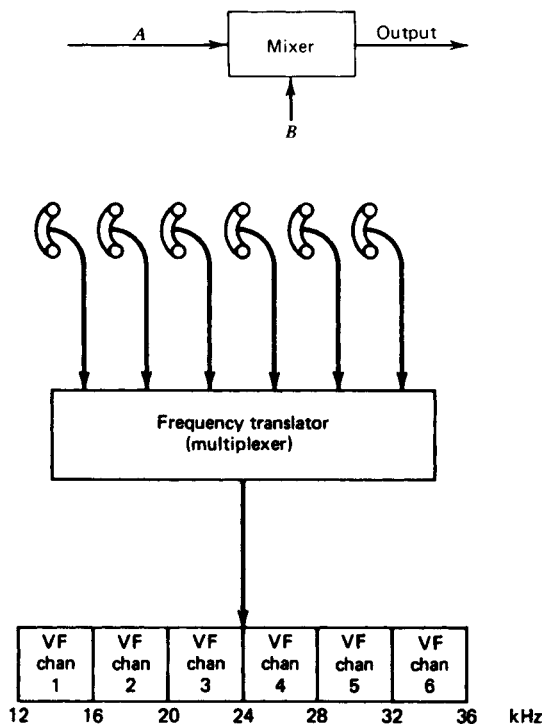


Figure 4.17 The frequency division multiplex concept illustrated.

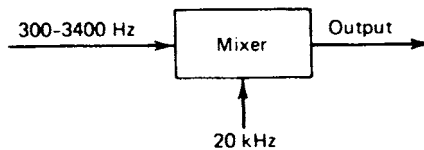
In practice, the *frequency translator* (multiplexer) uses single sideband modulation of radio-frequency (RF) carriers. A different RF carrier is used for each channel to be multiplexed. This technique is based on *mixing* or heterodyning the signal to be multiplexed, typically a voice channel, with an RF carrier.

An RF carrier is an unmodulated RF signal of some specified frequency. In theory, because it is not modulated, it has an indefinitely small bandwidth. In practice, of course, it does have some measurable bandwidth, although very narrow. Such a carrier derives from a simple frequency source such as an oscillator or a more complex source such as a synthesizer, which can generate a stable output in a range of frequencies.

A simplified block diagram of an FDM link is shown in Figure 4.18.

**4.5.2.2 Mixing.** The heterodyning or mixing of signals of frequencies  $A$  and  $B$  is shown as follows. What frequencies may be found at the output of the mixer? Both the original signals will be present, as well as the signals representing their sum and their difference in the frequency domain. Thus at the output of the illustrated mixer we will have present the signals of frequency  $A$ ,  $B$ ,  $A + B$ , and  $A - B$ . Such a mixing process is repeated many times in FDM equipment.

Let us now look at the boundaries of the nominal 4-kHz voice channel. These are 300 Hz and 3400 Hz. Let us further consider these frequencies as simple tones of 300 Hz and 3400 Hz. Now consider the following mixer and examine the possibilities at its output:



First, the output may be the sum or

$$\begin{array}{r}
 20,000 \text{ Hz} \\
 + \quad 300 \text{ Hz} \\
 \hline
 20,300 \text{ Hz}
 \end{array}
 \qquad
 \begin{array}{r}
 20,000 \text{ Hz} \\
 + \quad 3,400 \text{ Hz} \\
 \hline
 23,400 \text{ Hz}
 \end{array}$$

A simple low-pass filter could filter out all frequencies below 20,300 Hz.

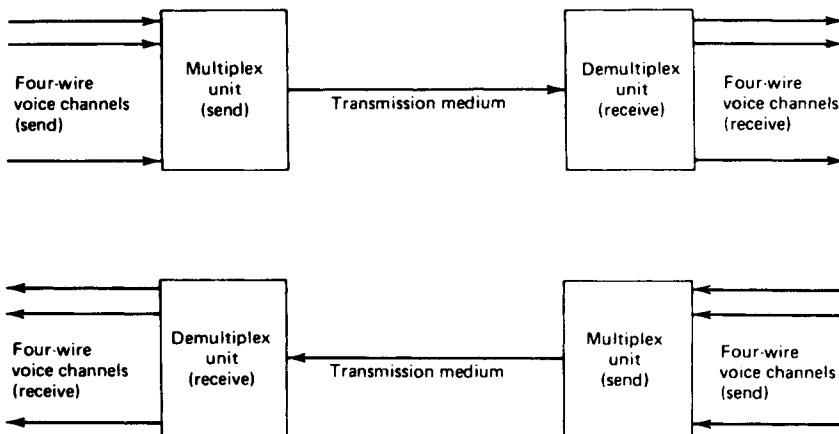
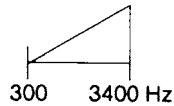
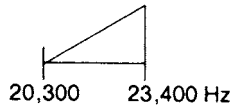


Figure 4.18 Simplified block diagram of an FDM link.

Now imagine that instead of two frequencies, we have a continuous spectrum of frequencies between 300 Hz and 3400 Hz (i.e., we have the voice channel). We represent the spectrum as a triangle:



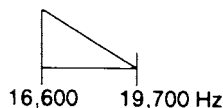
As a result of the mixing process (translation) we have another triangle, as follows:



When we take the sum, as we did previously, and filter out all other frequencies, we say we have selected the upper sideband. Therefore we have a triangle facing to the right, and we call this an upright or erect sideband. We can also take the difference, such that

$$\begin{array}{r}
 20,000 \text{ Hz} \\
 - 300 \text{ Hz} \\
 \hline
 19,700 \text{ Hz}
 \end{array}
 \qquad
 \begin{array}{r}
 20,000 \text{ Hz} \\
 - 3,400 \text{ Hz} \\
 \hline
 16,600 \text{ Hz}
 \end{array}$$

and we see that in the translation (mixing process) we have had an inversion of frequencies. The higher frequencies of the voice channel become the lower frequencies of the translated spectrum, and the lower frequencies of the voice channel become the higher when the difference is taken. We represent this by a right triangle facing the other direction:



This is called an inverted sideband. To review, when we take the sum, we get an erect sideband. When we take the difference, frequencies invert and we have an inverted sideband represented by a triangle facing left.

Again, this modulation technique is called single-sideband suppressed carrier (SSBSC). It is a type of amplitude modulation (AM). With conventional AM, the modulation produces two sidebands, an upper sideband and a lower sideband, symmetrical on either side of the carrier. Each sideband carries the information signal. If we tune to 870 kHz on the AM dial, 870 kHz is the frequency of the RF carrier, and its sidebands fall on either side, where each sideband is about 7.5 kHz wide. Thus a radio station on the AM dial requires about 15 kHz of spectrum bandwidth.

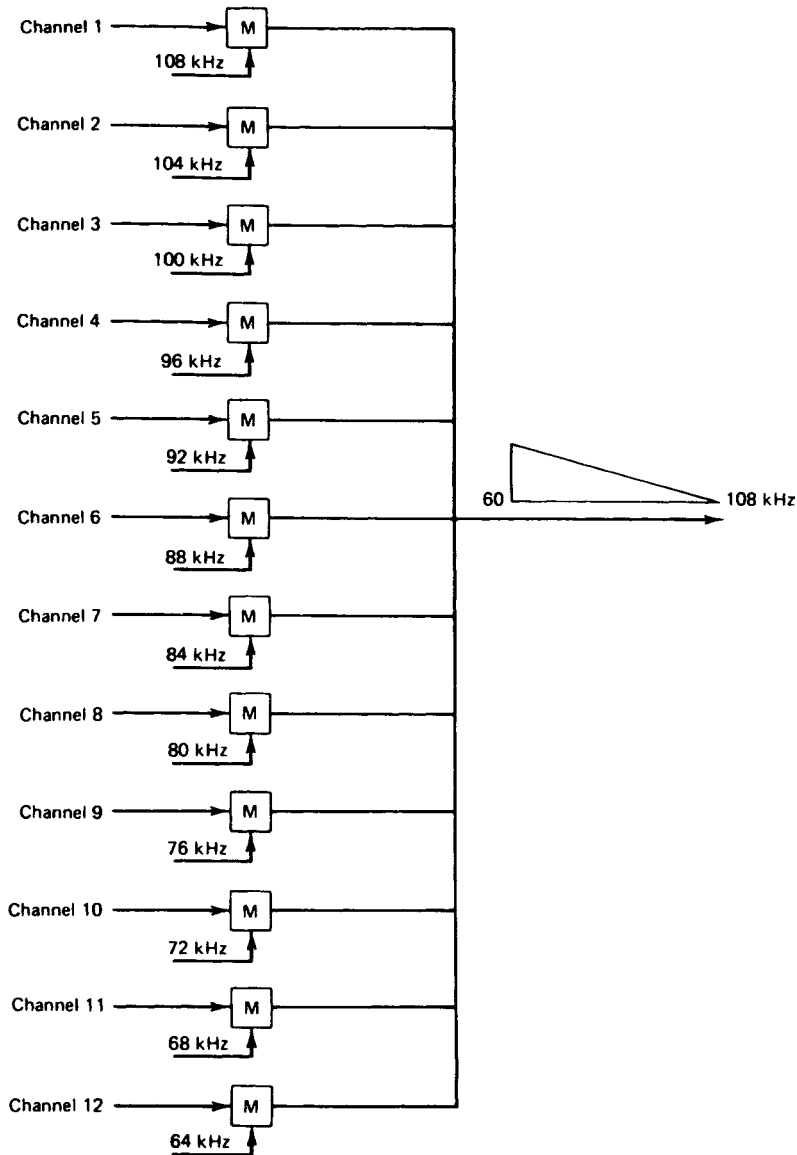
In our case, there are also two sidebands extending about 3.4 kHz either side of the carrier frequency. Of course, the carrier frequency is the local oscillator frequency, which is suppressed at the output, as is the upper sideband. All that remains is the lower sideband, which contains the voice channel information.

### 4.5.2.3 CCITT Modulation Plan

4.5.2.3.1 Introduction. A modulation plan sets forth the development of a band of frequencies called the *line frequency* (i.e., ready for transmission on the line or transmission

medium). The modulation plan usually is a diagram showing the necessary mixing, local oscillator mixing frequencies, and the sidebands selected by means of the triangles described previously in a step-by-step process from voice channel input to line frequency output. The CCITT has recommended a standardized modulation plan with a common terminology. This allows large telephone networks, on both national and multinational systems, to interconnect. In the following paragraphs the reader is advised to be careful with terminology.

**4.5.2.3.2 Formation of the Standard CCITT Group.** The standard *group*, as defined by the CCITT, occupies the frequency band of 60 kHz to 108 kHz and contains 12 voice channels. Each voice channel is the nominal 4-kHz channel occupying the 300-Hz to



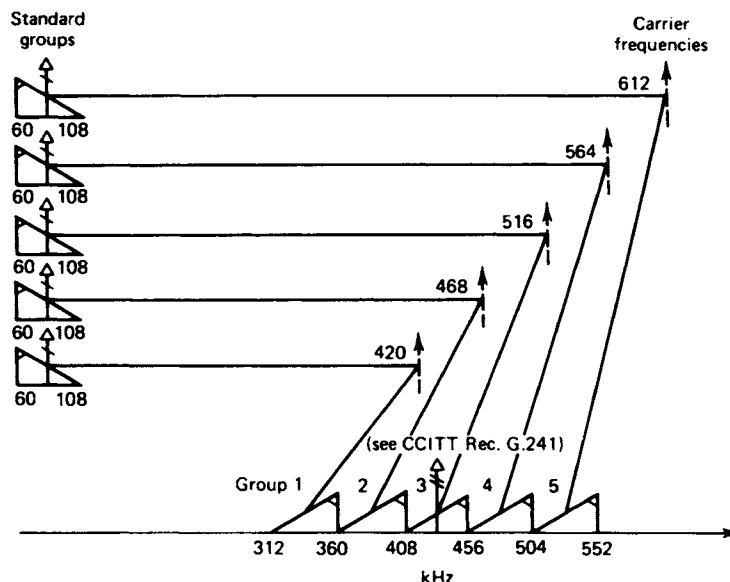
**Figure 4.19** Formation of the standard CCITT group.

3400-Hz spectrum. The group is formed by mixing each of the 12 voice channels with a particular carrier frequency associated with each channel. Lower sidebands are then selected, and the carrier frequencies and the upper sidebands are suppressed. Figure 4.19 shows the preferred approach to the formation of the standard CCITT group. It should be noted that in the 60-kHz to 108-kHz band, voice channel 1 occupies the highest frequency segment by convention, between 104 kHz and 108 kHz. The layout of the standard group is illustrated in Figure 4.19. Single sideband suppressed carrier (SSBSC) modulation techniques are utilized universally.

**4.5.2.3.3 Formation of the Standard CCITT Supergroup.** A supergroup contains five standard CCITT groups, equivalent to 60 voice channels. The standard supergroup, before further translation, occupies the frequency band of 312 kHz to 552 kHz. Each of the five groups making up the supergroup is translated in frequency to the supergroup frequency band by mixing with the appropriate carrier frequencies. The carrier frequencies are 420 kHz for group 1, 468 kHz for group 2, 516 kHz for group 3, 564 kHz for group 4, and 612 kHz for group 5. In the mixing process, in each case, the difference is taken (i.e., the lower is selected). This frequency translation process is illustrated in Figure 4.20.

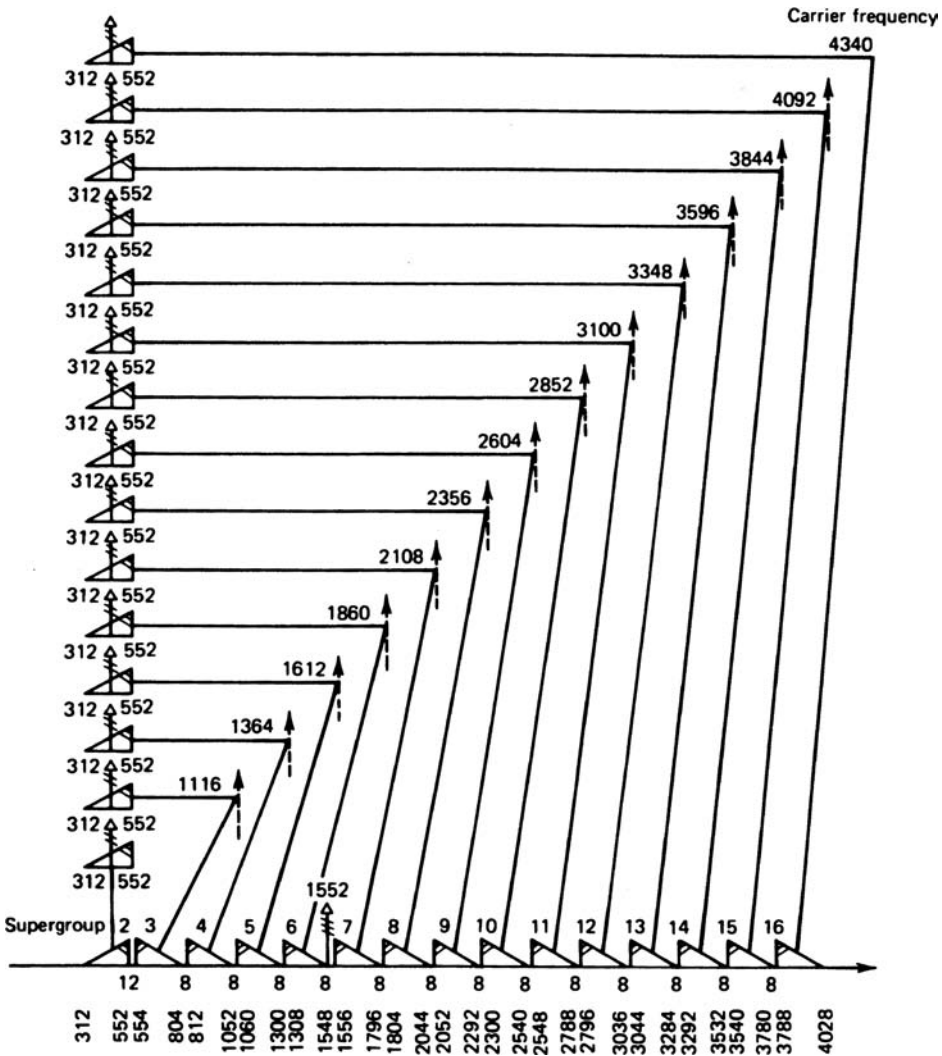
**4.5.2.4 Line Frequency.** The band of frequencies that the multiplexer applies to the line, whether the line is a radiolink, wire pair, or fiber-optic cable, is called the line frequency. Some texts use the term *high frequency* (HF) for the line frequency. This is not to be confused with HF radio, which is a radio system that operates in the band of 3 MHz to 30 MHz.

The line frequency in this case may be the direct application of a group or supergroup to the line. However, more commonly a final frequency translation stage occurs, particularly on high-density systems.<sup>15</sup> An example of line frequency formation is illustrated in



**Figure 4.20** Formation of the standard CCITT supergroup. The vertical arrows show the frequencies of the group level regulating pilot tones. (From CCITT Rec. G.233, courtesy of ITU-T Organization, Ref. 6.)

<sup>15</sup>“High-density” meaning, in this context, a system carrying a very large number of voice channels.



**Figure 4.21** Makeup of the basic CCITT 15-supergroup line frequency assembly. (From CCITT Rec. G.233, courtesy of the ITU-T Organization, Geneva, Ref. 6.)

Figure 4.21. This figure shows the makeup of the basic 15-supergroup assembly. Its capacity is  $(15 \times 60)$  900 voice channels (Ref. 7).

### 4.5.3 Pilot Tones

Pilot tones are used to control level in FDM systems. They may also be used to actuate maintenance alarms. In Figures 4.20 and 4.21, the pilot tones are indicated by the vertical lines with little triangles on top. For example in the 15-supergroup assembly shown in Figure 4.21, the pilot tone is at 1552 kHz.

A pilot tone provides a comparatively constant amplitude reference for an automatic gain control (AGC) circuit. Frequency division multiplex equipment was designed to carry speech telephony. The nature of speech, particularly its varying amplitude, makes it a poor prospect as a reference for level control. Ideally, simple single-sinusoid (a sine

wave signal), constant-amplitude signals with 100% duty cycles provide simple control information for level regulating equipment (i.e., the AGC circuit).<sup>16</sup>

Modern FDM equipments initiate a level-regulating pilot tone on each group on the transmitting side of the circuit. Individual level-regulating pilots are inserted on each supergroup and other frequency configurations. The intent is to maintain the system level within  $\pm 0.5$  dB.

Pilots are assigned frequencies that are part of the transmitted FDM spectrum yet do not interfere with the voice channel operation. They are standardized by CCITT and are usually inserted on a frequency in the guard between voice channels or are residual carriers (i.e., partially suppressed carriers).

#### **4.5.4 Comments on the Employment and Disadvantages of FDM Systems**

FDM systems began to be implemented in the 1950s, reaching a peak employment in the 1970s. All long-haul (long-distance) broadband systems, typically line-of-sight microwave, satellite communication, and coaxial cable systems, almost universally used FDM configurations. One transcontinental system in North America, called the L-5 system, carried 10,800 voice channels on each cable pair. (Remember the system is four-wire, requiring two coaxial cables per system.) There were 10 working cable pairs and one pair as spare. The total system had 108,000 voice channel capacity.

Few, if any, new FDM systems are being installed in North America today. FDM is being completely displaced by TDM systems (i.e., digital PCM systems, see Chapter 6). The principal drawback of FDM systems is noise accumulation. At every modulation–demodulation point along a circuit, noise is inserted. Unless the system designer was very careful, there would be so much noise accumulated at the terminal end of the system that the signal was unacceptable and the signal-to-noise ratio was very poor. Noise does not accumulate on digital systems.

We incorporated this section on frequency division multiplex so that the reader would understand the important concepts of FDM. Many will find that frequency division techniques are employed elsewhere such as on satellite communications, cellular, and PCS systems. Also keep in mind that FDM as described herein is still widely used in emerging nations. However, these networks are also starting to phase out FDM in favor of a digital network based on TDM (PCM). TDM techniques are covered in Chapter 6.