CHAPTER 5

Switching networks

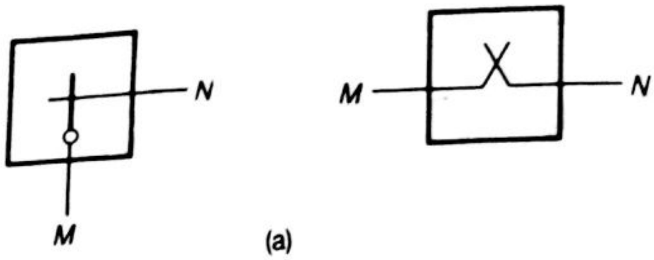
5.1 Introduction

A basic requirement for constructing switching systems, such as telephone exchanges. is to be able to design switching networks having a greater number of outlets than the switches from which they are built. This can be done by connecting a number of switching stages in tandem. For example, the Strowger exchange shown in Figure 3.8 gives access to up to 10 000 line terminations by using three ranks of 100-outlet selectors (two ranks of group selectors and a rank of final selectors). Figure 3.15 shows a two-stage network of ten-outlet crossbar switches giving access to 100 outgoing trunks. Figures 3.18 and 3.19 illustrate supervisory trunks connected to called customers' lines via four stages of crossbar switches. Figure 3.24 shows supervisory trunks connected to calling lines via three stages and to called lines via four stages of reed-relay switches.

In his monumental book on traffic theory[1] Syski wrote: 'At the present stage of development. the theoretical analysis of the telephone exchange as a whole has not yet been attempted.' In his book on switching networks[2] Benes said: 'The general theory of switching systems now consists of some apparently unrelated theorems, hundreds of models and formulas for simple parts of systems, and much practical lore associated with specific systems.' This chapter will attempt to develop such models and formulas for some of the basic networks used in switching systems. It will concentrate on space-division networks and the results will be extended to time-division networks in Chapter 6.

5.2 Single-stage networks

Figure 3.11 shows a single-stage network having M inlets and N outlets, consisting of a matrix of crosspoints. These may, for example, be separate relays or electronic devices or the contacts of a crossbar switch. The network could also be constructed by multipling the banks of M uniselectors or one level of a group of M two-motion



(a)

Figure 5.1 Switch symbols.

selectors, each having N outlets. Future systems may employ *photonic* switches,[3] in which opto-electronic devices are used as crosspoints to make connections between optical-fiber trunks. The network of Figure 3.11 may be represented, in simplified form, by the symbols shown in Figure 5.1. When Figure 5.1(a) is used to represent electromechanical switches, the circle indicates the side of the switch associated with the control mechanism (e.g. the wipers of a Strowger switch or the bridge magnet of a crossbar switch).

The switch shown in Figure 3.11 gives full availability; no calls are lost unless all outgoing trunks are congested. The number of simultaneous connections that can be made is either M (if $M < N$) or N (if $N < M$). The switch contains MN crosspoints. If $M = N$, the number of crosspoints is:

$$C_1 = N^2 \tag{5.1}$$

Thus, cost (as indicated by the number of crosspoints) increases as the square of the size of the switch. However, efficiency (as indicated by the proportion of the crosspoints which can be used at any time, i.e. $N/N^2 = 1/N$) decreases inversely with N. It is therefore uneconomic to use a single-stage network for large numbers of inlets and outlets. For example, a switch with 100 inlets and outlets requires 10 000 crosspoints and only 1% of these can be in use at any time. Switches for making connections between large numbers of trunks are therefore constructed as networks containing several stages of switches.

If the switch shown if Figure 3.11 is used to make connections between N similar circuits, then each circuit is connected to both an inlet and an outlet. Operation of the crosspoint at coordinates (j,k) to connect inlet j to outlet k thus performs the same function as operating crosspoint (k,j) to connect inlet k to outlet j. Consequently, half

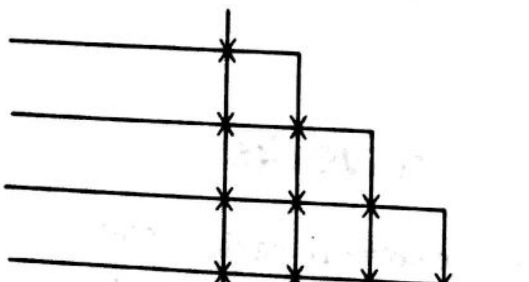


Figure 5.2 Triangular crosspoint matrix for connecting both-way trunks.

the crosspoints are redundant and can be eliminated. This results in the triangular crosspoint matrix shown in Figure 5.2. The number of crosspoints required is

$$C_1 = \tfrac{1}{2}N(N-1) \tag{5.2}$$

Triangular switches are not usually found in telephone switching systems because both-way trunks are not used. The trunks are operated on a one-way basis to facilitate supervision. For example, ringing tone and ringing current are sent over separate one-way trunks depending on whether a customer's line is calling or being called.

5.3 Gradings

5.3.1 Principle

For a route switch or a concentrator it is not necessary for each incoming trunk to have access to every outgoing trunk. It is adequate if each incoming trunk has access to a sufficient number of trunks on each route to give the required grade of service. This is known as *limited availability*. The number of outgoing trunks to which an incoming trunk can obtain connection is called the *availability* and corresponds to the outlet capacity of the switches used.

Figure 5.3(a) shows 20 trunks on an outgoing route to which incoming trunks have access by means of switches giving an availability of only ten (e.g. 20 circuits on an outgoing junction route from a selector level in a step-by-step exchange having 100-outlet two-motion group selectors). In Figure 5.3(a), the outlets of the switches are multipled together in two separate groups and ten outgoing trunks are allocated to

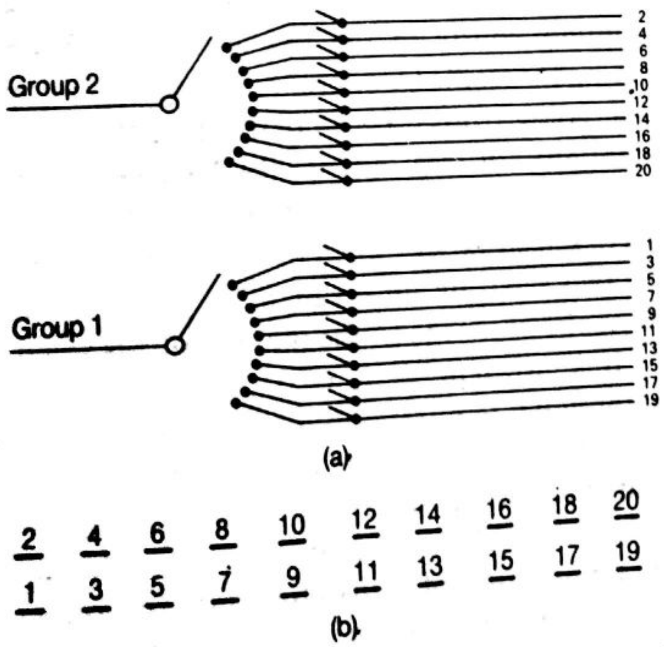


Figure 5.3 **Twenty trunks connected in two separate groups to switches of availability 10. (a) Full diagram. (b) Grading diagram.**

Scanned by CamScanner

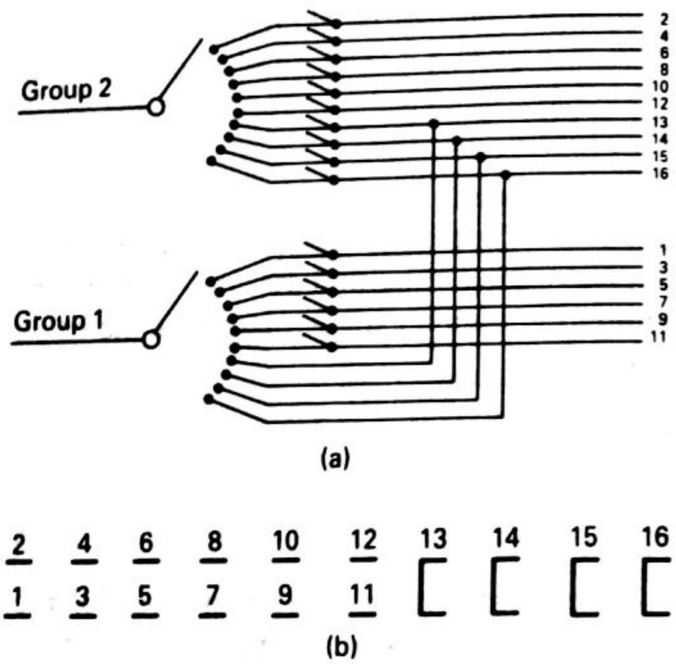


Figure 5.4 Sixteen trunks interconnected to two groups of switches of availability 10. (a) Full diagram. (b) Grading diagram.

each group. If the total traffic offered by the incoming trunks is, say, 8 E, each group of outgoing trunks is offered 4 E and will provide a grade of service (GOS) of better than 0.01. (Table 4.1 shows that a full-availability group of 10 trunks will cater for 4.5 E with a GOS of 0.01). The arrangement shown in Figure 5.3(a) is clearly less efficient than a single full availability group. (Table 4.1 shows that only 15 trunks are required to provide the same GOS for 8 E of traffic.)

If the traffic offered to the two groups of incoming trunks is random, peak loads will seldom occur simultaneously in the two groups. Efficiency can therefore be improved through mixing the traffic by interconnecting the multiples of the two groups so that some of the outgoing trunks are available to both groups of switches. If the switches search sequentially for free outlets, the later-choice outlets carry the least traffic (as shown in Figure 4.8). It is therefore desirable to connect the later-choice trunks to both groups of selectors, as shown in Figure 5.4(a). In this arrangement, the first six outlets are in two separate full-availability groups; the last four outlets are common to both groups and carry the traffic that overflows when the first six outlets of either group are busy. It is shown in Section 5.3.4 that this arrangement will still give a GOS of about 0.01, although it requires only 16 trunks instead of 20. The arrangement shown in Figure 5.4(a) requires only one more outgoing trunk than a full-availability group having a similar grade of service.

The technique described above of interconnecting the multiples of switches is called *grading*. An interconnection of trunks based on this principle is called grading. The conventional diagramatic representation of a grading is shown in Figure 5.4(b). A grading enables a single switching stage to provide access to a number of trunks greater than the availability (i.e. the outlet capacity) of the switches, but not exceeding it by an order of magnitude. A grading provides a poorer grade of service than a full-availability group with the same number of trunks. Some lost calls occur, even when there are free outgoing trunks, when these trunks are in part of the grading which is not accessible to the group of selectors containing the incoming trunk requiring connection.

Gradings of the form shown in Figure 5.4 were extensively studied by G. F. O'Dell [4] in the 1920s and are therefore called *O'Dell gradings*. They are the most widely used form of grading in a class of gradings known as *progressive gradings* because the switches hunt over the outlets sequentially from a fixed home position.

5.3.2 Design of progressive gradings

In order to form a grading, the switches having access to the outgoing route are multipled into a number of separate groups, known as *graded groups*. On early choices, each group has access to individual trunks and on late choices trunks are common, as shown in Figure 5.4. This diagram shows a small grading for only two groups of switches. For larger numbers of outgoing trunks, gradings may contain four or more groups. For example, Figure 5.5 shows four-group gradings. Since the traffic decreases with later choices of outlet, the number of groups connected together increases from individual connections on the early choices through partial commons (doubles) to full commons on the late choices.

In designing a grading to provide access to N outgoing trunks from switches having availability k, the first step is to decide on the number of graded groups g. If all the choices were individual trunks, we would have $N = gk$. If all the choices were full commons, $N = k$. Since the grading contains a mixture of individuals, partial commons and full commons, then $k < N < gk$. A reasonable choice for N is $N = \frac{1}{2}gk$ and traffic simulations have shown that the efficiency of such gradings is near the optimum. The number of groups is thus chosen to be:

$$g = \frac{2N}{k} \tag{5.3}$$

Since the grading should be symmetrical, g must be an even number, so the value of g given by equation (5.3) is rounded up to the next even integer.

It is now necessary to decide how the gk trunks entering the grading are to be interconnected to N outgoing trunks. For a two-group grading there is only one solution. If the number of columns of 'singles' is s and the number of commons is c, then:

$$\begin{aligned} \text{Availability} &= k = s + c \\ \text{No. of trunks} &= N = 2s + c \\ \therefore \quad & s = N - k \text{ and } c = 2k - N \end{aligned}$$

If the grading has more than two groups, there is no unique solution. It is necessary to choose from the possible solutions the best one, i.e. the grading with the greatest traffic

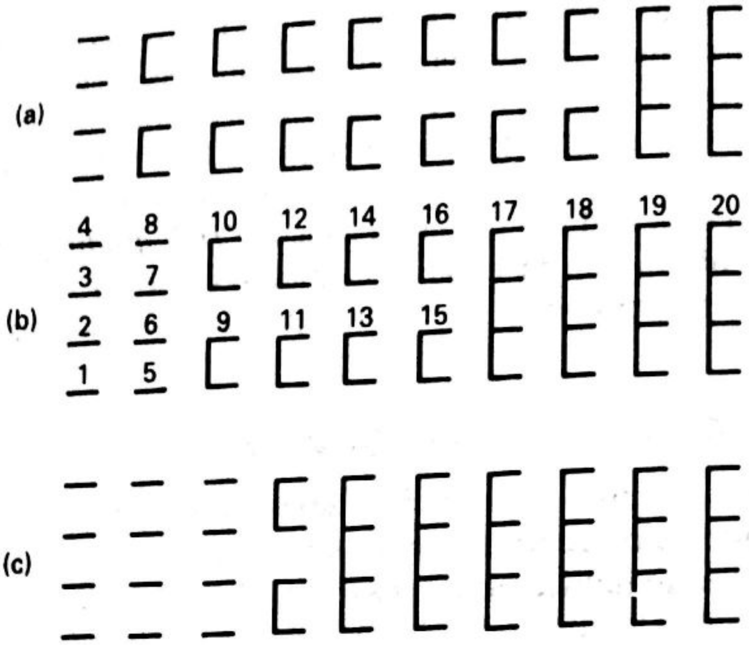


Figure 5.5 Four-group gradings for 20 trunks (availability 10).

capacity. The traffic offered to adjacent outlets will not differ greatly, so they should not be connected to very different sizes of common. There should thus be a smooth progression on the choices from individuals to partial commons, from smaller partial commons to larger ones, and from partial commons to full commons. The numbers of choices of each type in a group should therefore be as nearly equal as possible. This is achieved by minimizing the sum of the successive differences between the number of choices of one type and those of the type following it.

Let g have q factors: $f_1 < f_2 < \dots < f_i \dots < f_q$, where

$$f_1 = 1 \text{ and } f_q = g$$

Let r_i be the number of choices having their incoming trunks connected as f_i tuples.

$$\therefore \sum_{i=1}^{q} r_i = k \tag{5.4}$$

Now each f_i tuple contains g/f_i outgoing trunks.

$$\therefore \sum_{i=1}^{q} r_i g/f_i = N \tag{5.5}$$

Since there are only two equations and more than two unknowns (if $q > 2$), there are a number of different solutions for $(r_1, \dots, r_q)$. These are found and, for each, the sum of the successive differences, D, is given by:

$$D = |r_{-1} - r_2| + |r_2 - r_3| + \dots + |r_{q-1} - r_{1q}| \tag{5.6}$$

The best grading is that having the smallest value of D.

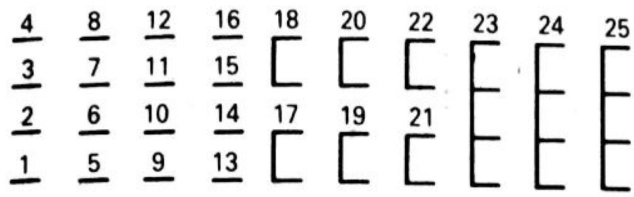


Figure 5.6 Grading of Figure 5.5(b) modified to accommodate 25 trunks.

Example 5.1

Design a grading for connecting 20 trunks to switches having ten outlets.

The number of graded groups, given by equation (5.3) is $g = 40/10 = 4$, and the factors of g are 1, 2 and 4. Let

the number of choices having singles $= s$
the number of choices having doubles $= d$
the number of choices having quadruples $= q$

Substituting in equations (5.4) and (5.5):

$$s + d + q = 10$$
$$4s + 2d + q = 20$$
$$\therefore \quad 3s + d = 10$$

If
$$s = 1 : d = 7 \text{ and } q = 10 - 8 = 2$$
$$s = 2 : d = 4 \text{ and } q = 10 - 6 = 4$$
$$s = 3 : d = 1 \text{ and } q = 10 - 4 = 6$$
$$s \geqslant 4 : d < 0, \text{ so this is not possible.}$$

There are thus three possible gradings, which are shown in Figure 5.5. The sums of the successive differences for these gradings are respectively given by:

$$D_1 = 6 + 5 = 11$$
$$D_2 = 2 + 0 = 2$$
$$D_3 = 2 + 5 = 7$$

The second grading (shown in Figure 5.5(b)) is therefore the best.

If a growth in traffic makes it necessary to increase the number of trunks connected to a grading, this can be done by reducing the number of commons and partial commons and increasing the number of individuals. Figure 5.6 shows the grading of Figure 5.5(b) rearranged to provide access to 25 trunks.

5.3.3 Other forms of grading

In an O'Dell grading, the partial commons are arranged as separate groups, so each is available to only some of the incoming trunks. For example, in Figure 5.5(b) the upper row of pairs serves only the first two groups. However, the principle of grading is based on the sharing of outgoing trunks between different sets of incoming trunks. Efficiency can be improved if this principle can be applied to the whole of a grading instead of only

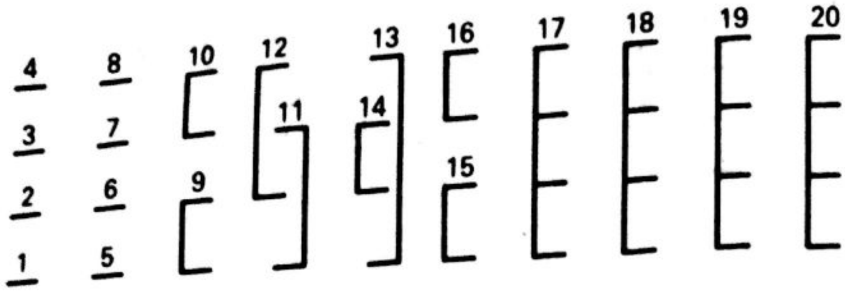


Figure 5.7 Skipped grading.

to parts of it. This can be done by connecting non-adjacent groups, in addition to adjacent groups, as shown in Figure 5.7. This is known as *skipping*.[5]

As an example of the improvement provided, consider the O'Dell grading of Figure 5.5(b) at a time when the upper two groups are carrying heavy traffic but the lower two groups are lightly loaded. Outgoing trunks 11 and 13 may be free but cannot be used. The skipped grading of Figure 5.7 enables these trunks to be used at such a time.

Progressive gradings are intended to be used with switches that hunt sequentially from a fixed home position. However, if switches do not hunt from a fixed home position or they select outlets at random, there is no advantage in connecting some outlets to singles and others to partial or full commons. The grading should then be designed to share each trunk between an equal number of groups, as shown in Figure 5.8. Such gradings are known as *homogeneous gradings*. Sequential hunting is more efficient than random selection because late-choice trunks are left free as long as possible to cater for traffic peaks. However, in practice, the difference between the traffic capacities of sequential and homogeneous gradings is often quite small.[6]

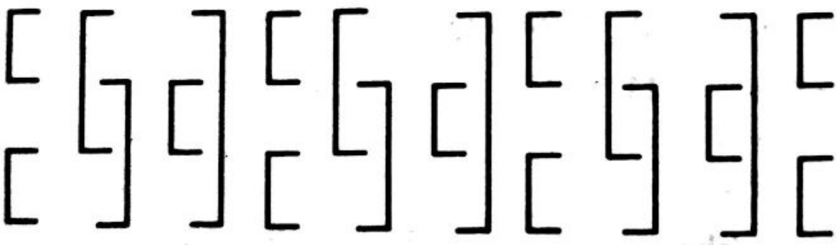


Figure 5.8 Homogeneous grading.

5.3.4 Traffic capacity of gradings

In an ideal grading, the interconnections would ensure that each outgoing trunk carried an identical traffic load.[1] Thus, if total traffic A is carried by N trunks, the occupancy of each trunk is A/N. It is assumed that each trunk being busy is an independent random event. Each call has access k to trunks (where k is the availability), and the probability of all k trunks being busy is thus:

$$B = (A/N)^k$$

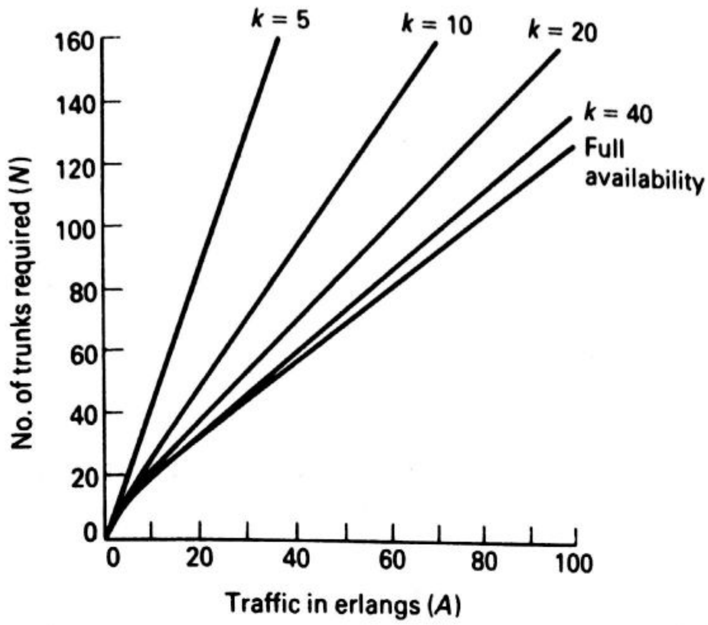


Figure 5.9 Traffic capacity of gradings (from modified Erlang formula, $B = 0.002$).

The number of trunks required to carry A erlangs with a GOS of B is therefore given by:

$$N = AB^{-1/k} \tag{5.7}$$

This is *Erlang's ideal grading formula* and gives a linear relationship between the traffic and the number of trunks required.

Practical gradings do not satisfy the conditions for Erlang's ideal grading. However, it has been found[6] that they do have a linear relationship between traffic capacity A and number of trunks for a given grade of service B. An approximate curve of A against N can therefore be derived from Erlang's full-availability theory for $N \leqslant k$ and extended as a straight line for $N \geqslant k$. From equation (5.7) this line is given by:

$$A = A_k + (N - k)B^{1/k} \tag{5.8}$$

where A_k is the traffic carried by a full-availability group of k trunks (with GOS = B).

Figure 5.9 shows a family of curves plotted from the above modified Erlang formula. This method is, of course, an approximation. More accurate methods have been developed. These include the Palm–Jacobeus formula[6] and the *equivalent random method* of Wilkinson, which is described in Appendix 2. Traffic tables for gradings have also been published.[7]

Example 5.2

Find the traffic capacity of the two-group grading shown in Figure 5.4 if the required grade of service is 0.01.

$k = 10$ and, from Table 4.1, $A_k = 4.5$ E.

From equation (5.8):

$$A = A_k + (N - k)\, B^{1/k}$$
$$= 4.5 + (16 - 10) \times 0.01^{0.1}$$
$$= 4.5 + 6 \times 0.631$$
$$= 8.3 \text{ E}$$

(Note: Table 4.1 shows that a full-availability group of 16 trunks can handle 8.9 E with 0.01 grade of service.)

5.3.5 Applications of gradings

Gradings have been widely employed in step-by-step systems. In Figure 3.10, the trunk distribution frames (TDF) between the ranks of selectors provide cross-connections in the form of gradings.

The use of gradings is not confined to the Strowger system. In the Plessey 5005 crossbar system[8] (the TXK1 system of British Telecom), which is shown in Figure 3.18, there is a grading on each side of the router switch.

Another example of the use of a grading in a link system is in the Bell No.1 ESS system.[9] The subscribers' concentrator of this system is shown in Figure 5.10(a). The number of crosspoints required in the primary switches is reduced by omitting them in a systematic manner. Each primary switch is equivalent to four groups of four-outlet selectors having access to eight trunks through the homogeneous grading shown in Figure 5.10(b). Full availability between the primary-switch inlets and the links would have required eight-outlet selectors, thus doubling the number of crosspoints needed.

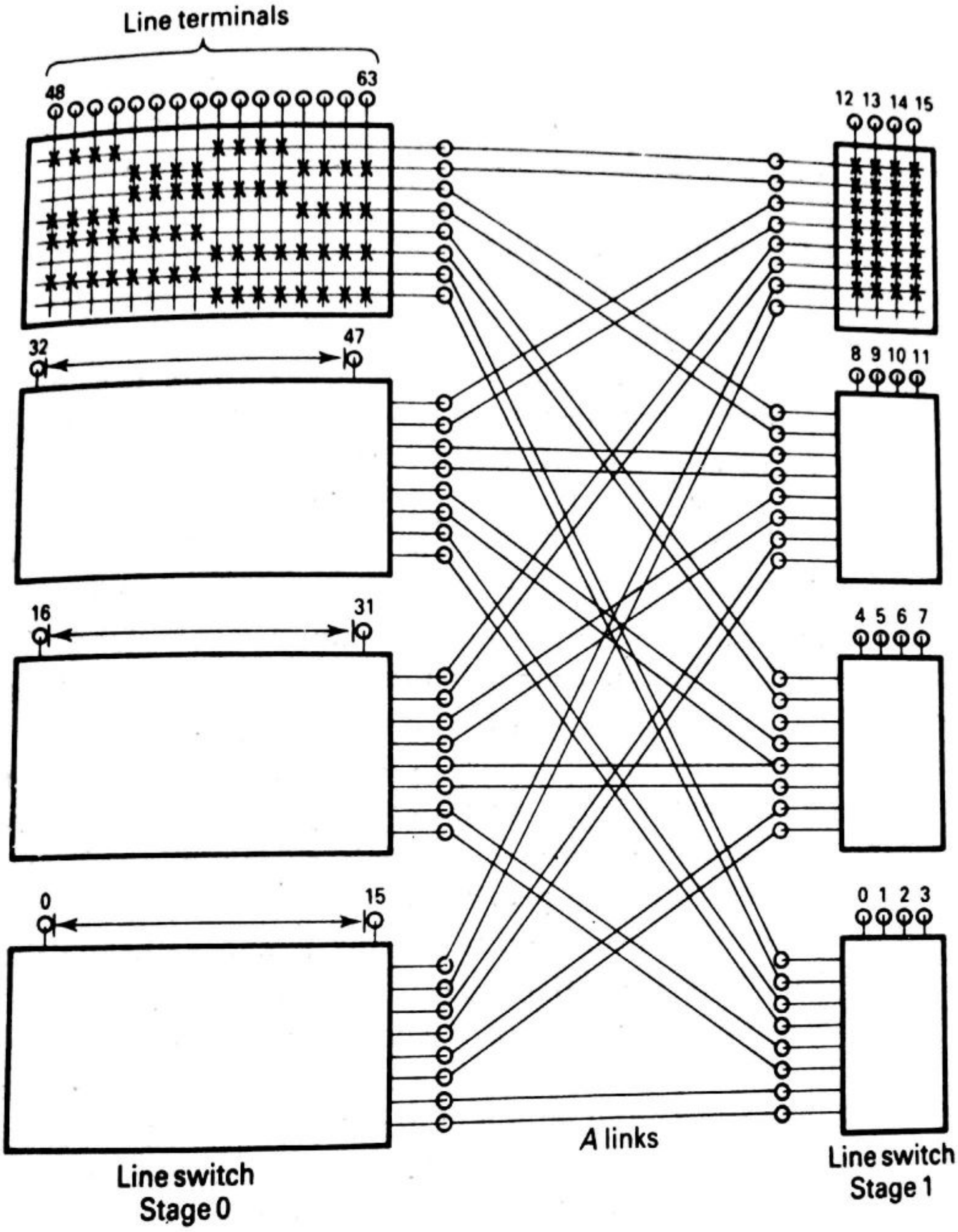
In the Ericsson AXE digital switching system[10] a subscribers' concentrator serving 2048 lines contains 16 modules, each with 128 lines. Each module has access to two PCM highways. One is individual to it and the other is common to all the modules of the concentrator. Traffic overflows to the latter when all PCM channels on the former are busy.

Automatic alternative routing (AAR) presents a situation similar to the use of grading. Traffic to a given destination is offered first to a group of direct circuits. If these are all busy, the traffic overflows to a common indirect route, where it is mixed with traffic to other destinations. In a grading, when all the 'singles' belonging to a group are busy, traffic overflows to common trunks where it mixes with traffic from other groups. Thus, the same theory is applicable to grading and AAR and similar dimensioning methods can be used. The equivalent random method is described in Appendix 2. It was developed by Wilkinson for AAR, but it is equally applicable to gradings.

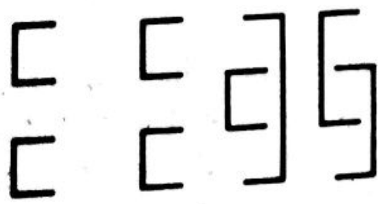
5.4 Link systems

5.4.1 General

Examples of two-stage link systems are shown in Figures 3.2, 3.15 and 3.16 and a four-stage link system is shown in Figure 3.17. In general, a link system may have any



(a)



(b)

Figure 5.10 Two-stage concentrator used in Bell No.1 ESS system. (a) Arrangement of trunks. (b) Four-group homogeneous grading incorporated at *A* stage. (Copyright © 1964 AT&T. All rights reserved. Reprinted with permission.)

number of stages and the number of possible patterns of links between stages is very large. Only a few commonly used configurations will be considered here.

In the two-stage network of Figure 3.15 there is only one link between each primary switch and each secondary switch. Thus, it may be impossible to make a connection from a given incoming trunk to a selected outgoing trunk because the link is already being used for another connection between that primary switch and that secondary switch. This situation is called *blocking*. It is also known as a *mismatch*, because free links exist but none of them can be used for the required connection. If connection must be made to one particular outgoing trunk (e.g. an individual customer's line), the probability of blocking is unacceptably high. For this application, it is therefore necessary to use a network with more stages (e.g. the four-stage network of Figure 3.17), in order to have a choice of paths through the network.

The two-stage network of Figure 3.15 can be used as a route switch. If it serves ten outgoing routes with ten trunks on each route, then trunk no. 1 of each route is connected to secondary switch no. 1, trunk no. 2 is connected to switch no.2, and so on. Thus, an incoming trunk can obtain connection to the selected outgoing route via any of the links outgoing from its primary switch. The call is only lost if all the paths to free outgoing trunks are blocked. The probability of this occurring simultaneously for all links is obviously much smaller than the probability of a single link being busy. Similarly, if the incoming trunks are from several different routes, one trunk from each route is normally terminated on each primary switch.

Step-by-step selection is unsuitable. If the link is chosen before the outgoing trunk, then a free link could be seized that leads to a secondary switch whose trunk on the required outgoing route is already busy. Instead, *conditional selection* is used. The marker does not set up a connection until it has interrogated the busy/free conditions of all the relevant outgoing trunks and internal links. Only when it has found a match between a free outgoing trunk and a free internal link does it operate the switches. With this method of selection, if a free path through the network exists, it can be used. In a multistage step-by-step network, it is possible to encounter congestion at a late stage which would have been avoided if a different choice had been made at an earlier stage of switching, i.e. a free path exists but it has been omitted from the limited search made by the selectors.

A further advantage of conditional selection is that the marker has access at the same time to both ends of the connection through the network. Having set up the connection, it can test it for continuity. If the connection is found to be faulty, the marker can produce a fault record and make a second attempt to set up the connection by choosing a different path through the network.

If any free trunk may be used, as when the network is acting as a concentrator, then an incoming trunk can use any free link from its primary switch. If there are as many links as outgoing trunks, a connection can always be made if there is a free outgoing trunk. Except in this case, a link system will normally give a poorer grade of service than a single full-availability switching stage. This is because calls are lost by internal blocking in addition to congestion of the external trunks.

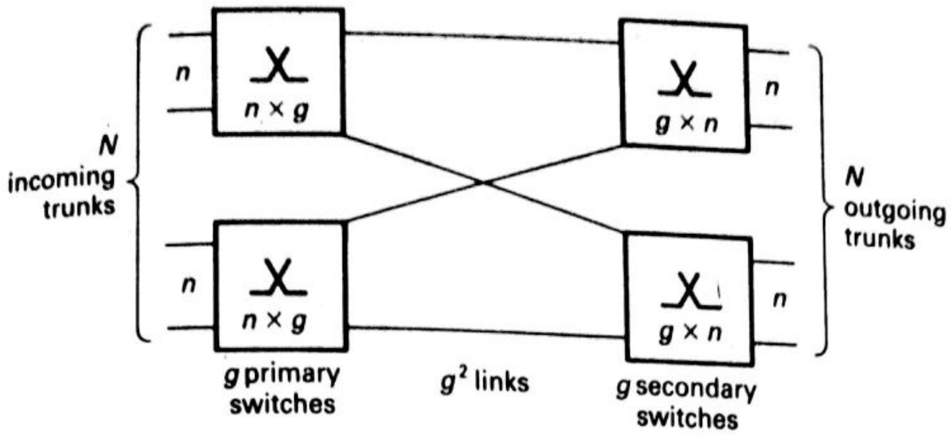


Figure 5.11 Two-stage switching network.

It has been seen that the grade of service of a link system depends on the way it is used. We may classify these uses as follows:

> *Mode* 1: Connection is required to one particular free outgoing trunk. (Since conditional selection is used, an attempt will not be made to set up this connection unless the trunk *is* free.)
> *Mode* 2: Connection is required to a particular outgoing route, but any free trunk on that route may be used.
> *Mode* 3: Connection may be made to any free outgoing trunk.

It will be seen from Figure 3.20 that a concentrator operates in mode 3, a route switch operates in mode 2 and an expander operates in mode 1. The grades of service obtained are considered in Section 5.5.

5.4.2 Two-stage networks

If the two-stage network shown in Figure 5.11 has N incoming and N outgoing trunks and contains primary switches having n inlets and secondary switches having n outlets, then no. of primary switches (g) = no. of secondary switches = no. of outlets per primary switch = no. of inlets per secondary switch, where

$$g = N/n$$

The no. of crosspoints per primary switch = no. of crosspoints per secondary switch $= gn = N$. The total no. of crosspoints (C_2) in the network = (no. of switches) × (crosspoints per switch) i.e.

$$C_2 = 2g\,N = 2N^2/n \tag{5.9}$$

Since there is one link from each primary switch to each secondary switch, the number of links is equal to no. of primary switches × no. of secondary switches, i.e.

$$\text{No. of links} = g^2 = (N/n)^2 \tag{5.10}$$

The number of crosspoints thus varies as $1/n$, but the numbers of link varies as $1/n^2$. If n is made very large to reduce the number of crosspoints, there will be too few links to carry the traffic. Let the number of links be equal to the number of incoming and outgoing trunks, a reasonable choice, since each set of trunks carries the same total traffic.

Then $\quad g^2 = N$

Substituting in equation (5.10) gives

$$n = \sqrt{N} \tag{5.11}$$

Then the total number of crosspoints (from equation (5.9)) is

$$C_2 = 2N^{3/2} \tag{5.12}$$

Equation (5.11) can be only a guide; one should select the nearest integer to n that is a factor of N. Also, in practice, designers are often constrained to use switch units of fixed sizes. For example, crossbar switches may be of sizes 10×10 or 10×20. The Bell No.1 ESS system[9] uses switches constructed from modules of size 8×8 and the British Telecom TXE2 system[11] uses modules of size 5×5.

The number of crosspoints per incoming trunk (from equation (5.12)) is $2N^{1/2}$. The cost per trunk therefore increases fairly slowly with the number of trunks. For large networks, however, it becomes more economic to use networks with more than two stages.

Example 5.3

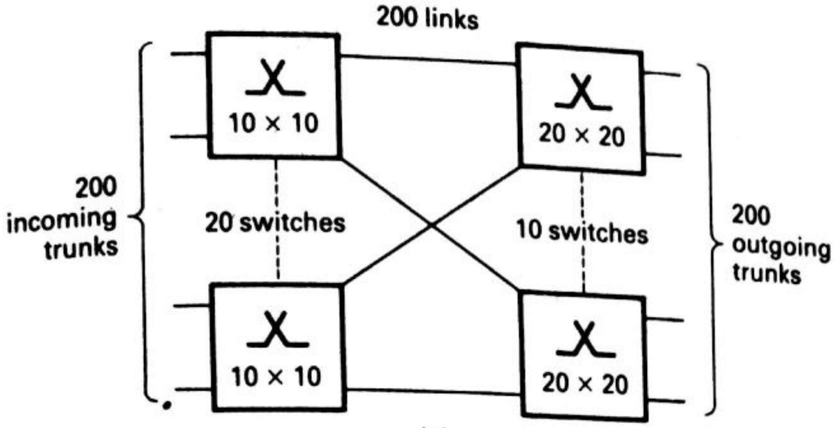
Design a two-stage switching network for connecting 200 incoming trunks to 200 outgoing trunks.

Now, $\sqrt{200} = 14.14$. However, n must be a factor of 200, so the nearest practicable values are $n = 10$ and $n = 20$. Two possible networks are shown in Figure 5.12. Each contains 6000 crosspoints. The network of Figure 5.12(a) is suitable for 20 outgoing routes, each having 10 trunks, and that of Figure 5.12(b) is suitable for 10 outgoing routes, each having 20 trunks.

The network in Figure 5.11 has the same number of outgoing trunks as incoming trunks. However, a concentrator has more incoming than outgoing trunks and an expander has more outgoing than incoming trunks.

Consider a concentrator with M incoming trunks and N outgoing trunks ($M > N$). Let each primary switch have m inlets and each secondary switch have n outlets. Then

No. of primary switches $= M/m$
No. of secondary switches $= N/n$
No. of crosspoints per primary switch $= mN/n$
No. of crosspoints per secondary switch $= nM/m$.



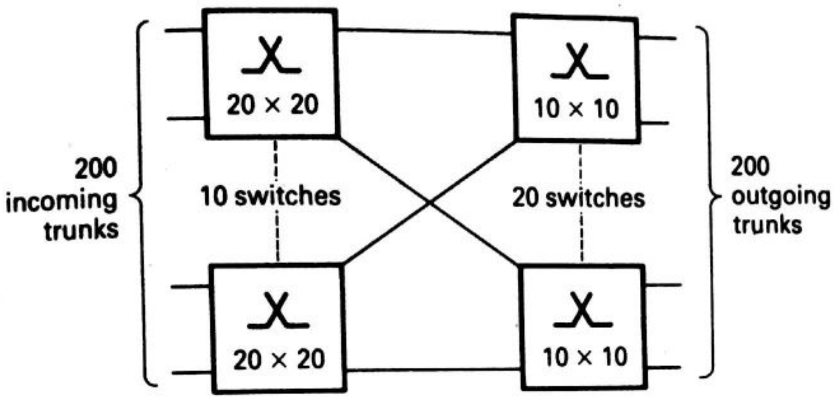


Figure 5.12 Examples of two-stage networks. (a) For 20 outgoing routes (10 trunks on each). (b) For 10 outgoing routes (20 trunks on each).

The total number of crosspoints is:

$$C_2 = \frac{M}{m}\frac{mN}{n} + \frac{N}{n}\frac{nM}{m} = MN\left[\frac{1}{n} + \frac{1}{m}\right] \tag{5.13}$$

The number of links = no. of primary switches × no. of secondary switches

$$= \frac{MN}{mn}$$

Since the traffic capacity is limited by the number of outgoing trunks, there is little point in providing more than this number of links, so let the number of links be N.

$$\therefore \qquad \frac{MN}{mn} = N$$

and

$$n = M/m \tag{5.14}$$

Substituting in equation (5.13) from equation (5.14):

$$C_2 = MN\left[\frac{m}{M} + \frac{1}{m}\right]$$

In order to minimize C_2, treat m as if it were a continuous variable and differentiate with respect to it:

$$\frac{dC_2}{dm} = MN\left[\frac{1}{M} - \frac{1}{m^2}\right]$$

$$= 0 \text{ when } m = \sqrt{M}$$

Hence, from equation (5.14):

$$m = n = \sqrt{M} \tag{5.15}$$

Thus, the number of crosspoints is a minimum when the number of inlets per primary switch equals the number of outlets per secondary switch.

Substituting in equation (5.13):

$$C_2 = MN\left[\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{M}}\right]$$

$$= 2M^{\frac{1}{2}} N \tag{5.16}$$

Again, equation (5.15) is no more than a guide; m and n must be integers and factors of M and N, respectively. Moreover, the designer may also be constrained to use switch units of standard sizes. Since $M > N$, equation (5.15) gives larger, and thus fewer, secondary switches than if $n = \sqrt{N}$ were chosen. Consequently, a poorer grade of service is obtained when the network is operated in mode 2. Thus, practical networks sometimes use $n = \sqrt{N}$, $m = M/\sqrt{N}$. For example, this is the case for the crossbar concentrator shown in Figure 3.16. To obtain an expander, M is exchanged with N and m with n.

5.4.3 Three-stage-networks

Figure 5.13 shows a three-stage switching network. There is one link from each primary switch to each secondary switch and one link from each secondary switch to each tertiary switch. A connection from a given inlet on a primary switch to a selected outlet on a tertiary switch may thus be made via any secondary switch, unles its link to the primary switch or its link to the secondary switch is busy. The call can be set up unless this condition applies simultaneously to every secondary switch. The probability of being unable to set up a connection because of blocking is thus much less than for a two-stage network. This network is therefore suitable for operation in mode 1.

If the three-stage network has N incoming trunks and N outgoing trunks and has primary switches with n inlets and tertiary switches with n outlets, then:

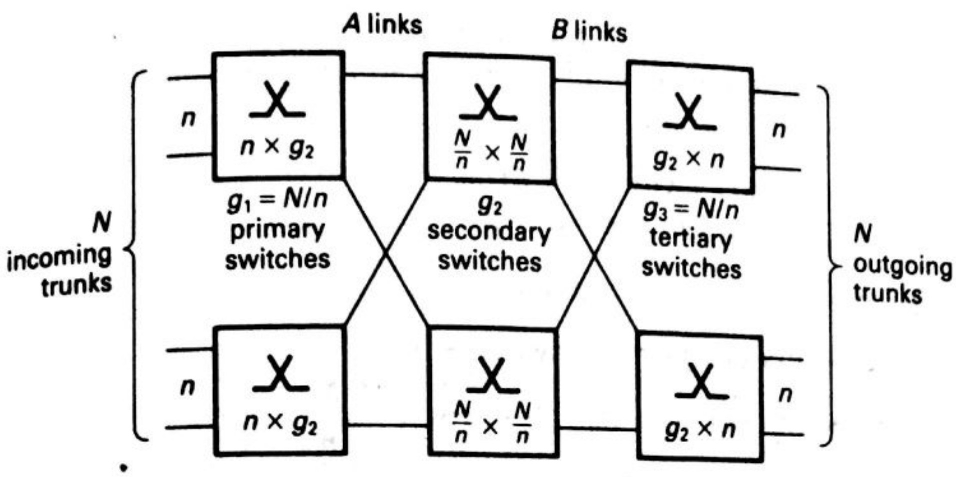


Figure 5.13 Fully interconnected three-stage switching network.

No. of primary switches (g_1) = no. of tertiary switches $(g_3) = N/n$.

$\therefore$ The secondary switches have N/n inlets and outlets.

If the number of primary–secondary links (A links) and secondary–tertiary links (B links) are each N, then the number of secondary switches is

$$g_2 = N \div (N/n) = n$$

= no. of outlets per primary switch = no. of inlets per tertiary switch.

No. of crosspoints in primary stage $= n^2(N/n) = nN$
No. of crosspoints in secondary stage $= n(N/n)^2 = N^2/n$
No. of crosspoints in tertiary stage $= n^2(N/n) = nN$

and the total number of crosspoints is

$$C_3 = N(2n + N/n) \tag{5.17}$$

By differentiating equation (5.17) with respect to n and equating to zero, it can be shown that the number of crosspoints is a minimum when

$$n = \sqrt{N/2} \tag{5.18}$$

and then

$$\begin{aligned} C_3 &= 2\sqrt{2}\, N^{3/2} \\ &= \sqrt{2}\, C_2 \\ &= 2^{3/2} N^{-1/2} C_1 \end{aligned} \tag{5.19}$$

If a three-stage concentrator has M incoming trunks and N outgoing trunks ($M > N$), its primary switches each have m inlets and its tertiary switches each have n outlets, then:

$$\text{No. of primary switches} = M/m$$
$$\text{No. of tertiary switches} = N/n$$

If there are g_2 secondary switches, then

$$\text{Crosspoints per primary switch} = m\, g_2$$
$$\text{Crosspoints per secondary switch} = \frac{M}{m}\frac{N}{n}$$
$$\text{Crosspoints per tertiary switch} = g_2\, n$$

The total number of crosspoints is:

$$\begin{aligned} C_3 &= \frac{M}{m} \times m g_2 + g_2 \times \frac{M}{m}\frac{N}{n} + \frac{N}{n} \times g_2\, n \\ &= g_2\left[M + N + \frac{M}{m}\frac{N}{n}\right] \end{aligned} \tag{5.20}$$

Since $M > N$, let no. of A links = no. of B links = N.

$\therefore$
$$N = g_2\frac{M}{m} = g_2\frac{N}{n}$$

Hence, $g_2 = n$ and $m = n\,M/N$.

Substituting in equation (5.20):

$$C_3 = (M + N)\,n + N^2/n$$

Differentiating with respect to n to find a minimum gives:

$$m = \frac{M}{\sqrt{M+N}}, \quad n = \frac{N}{\sqrt{M+N}} \tag{5.21}$$

$$C_3 = 2N\sqrt{N+M} \tag{5.22}$$

To obtain an expander, M is exchanged with N and m with n.

Example 5.4

Design a three-stage network for connecting 100 incoming trunks to 100 outgoing trunks:

$\sqrt{100/2} = 7.07$ $\therefore$ use $n = 5$ or $n = 10$

1. If $n = 5$, there are:
   20 primary switches of size 5×5

5 secondary switches of size 20 × 20
20 tertiary switches of size 5 × 5.

2. If $n = 10$, there are 10 primary switches, 10 secondary switches and 10 tertiary switches, each of size 10 × 10.

Both networks contain 3000 crosspoints. However, the second has more secondary switches. It therefore provides a greater number of paths between an incoming and an outgoing trunk and will exhibit less blocking.

Example 5.5

Design a three-stage network for 100 incoming trunks and 400 outgoing trunks.

$$100/\sqrt{100 + 400} = 4.47;\ 400/\sqrt{100+400} = 17.89$$

$\therefore$ $m = 4$ or 5; $n = 16$ or 20

1. If $m = 5$, $n = 20$, there are:
20 primary switches of size 5 × 5
5 secondary switches of size 20 × 20
20 tertiary switches of size 5 × 20.

2 If $m = 4$, $n = 16$, there are:
25 primary switches of size 4 × 4
4 secondary switches of size 25 × 25
25 tertiary switches of size 4 × 16.

Both networks contain 4500 crosspoints. However, the first contains more secondary switches and will therefore cause less blocking.

In a three-stage network, the number of the selected outgoing trunk is given by the outlet numbers used on the secondary and tertiary switches. It is not related to the outlet used in the primary switch, since any secondary switch may be used for connection to a given outgoing trunk. For each connection, two sets of links must be interrogated for the busy/free condition and matched to choose a pair connected to the same secondary switch. The control of a three-stage network is thus more complex than that of a two-stage one. For this reason, electromechanical systems usually use trunkings containing a number of separate two-stage networks in tandem. However, systems having electronic central control often employ three-stage switching networks.

A fully interconnected three-stage network (as shown in Figure 5.13) requires a large number of crosspoints when N is large. A reduction can be made in the number of crosspoints (at the expense of an increase in blocking) if the secondary switches have links to only some of the primary and tertiary switches, as shown in Figure 5.14. The secondary and tertiary switches are arranged in separate groups (frames) and are fully interconnected only within their groups. Each primary switch has one link to each of these secondary–tertiary groups. (Alternatively, the primary and secondary switches may be arranged in separate groups, to produce the mirror image of Figure 5.14.)

The number of switches is $3n^2$ and each has n^2 crosspoints, so the total number of crosspoints is:

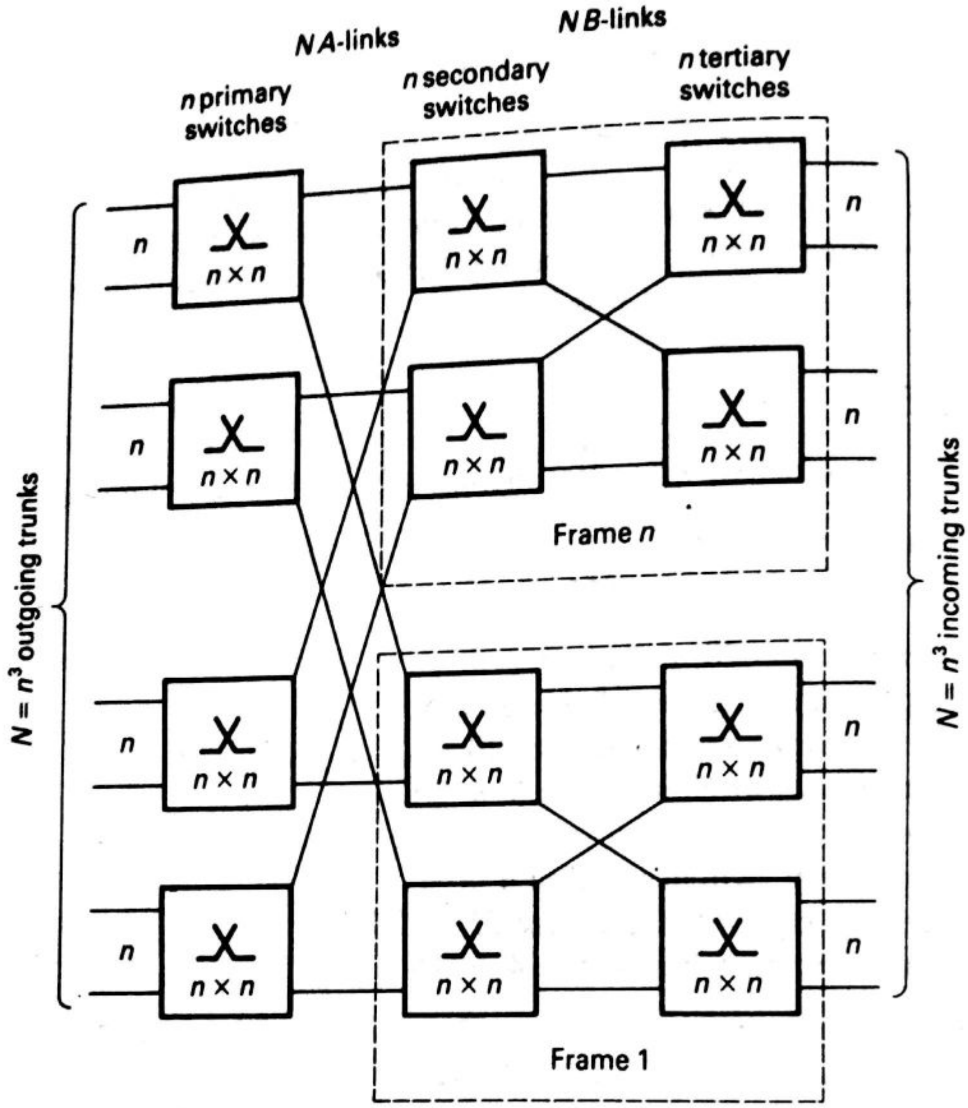


Figure 5.14 Partially interconnected three-stage network.

$$C_3 = 3n^4 = 3N^{4/3} \tag{5.23}$$

and the number of crosspoints per incoming trunk is $3N^{1/3}$.

Since there is only one link from each primary switch to each secondary–tertiary frame, there is only one path from each incoming trunk to each outgoing trunk. Thus, like the two-stage network of Figure 5.11, the network is unsuitable for use in mode 1 (i.e. making a connection from a given incoming trunk to a particular selected outgoing trunk). If the outlets of the tertiary switches serve a number of different routes, having several trunks connected to different secondary–tertiary frames, a sufficient choice of paths is available to give an acceptably low blocking probability in mode 2. A three-stage network of this type is used in some reed-electronic systems. Examples occur in the TXE2 system[11] and TXE4 system[12] of British Telecom.

In Figure 5.13, the third stage added to the two-stage network does not increase the number of outgoing trunks; it increases the mixture of paths available to reach them in order to reduce blocking. The additional stage may therefore be called a *mixing* *stage*. In Figure 5.14 the added third stage does not increase the number of paths to an outgoing trunk; it increases the number of outgoing trunks over which the incoming traffic can be distributed. It is therefore called a *distribution stage*. In Figure 5.13 a primary switch has a link to every secondary switch, so any secondary switch can be used for a connection to a given outgoing trunk. In Figure 5.14, however, there are many more primary switches and each has a link to only one secondary switch of each two-stage frame.

5.4.4 Four-stage networks

A four-stage network can be constructed by considering a complete two-stage network as a single switch and then forming a larger two-stage array from such switches. Figure 3.17 shows a four-stage network for 1000 incoming and 1000 outgoing trunks constructed from two-stage networks (frames) of 100 inlets and 100 outlets using 10×10 switches. It is necessary that one trunk (B link) be connected from each secondary switch of an incoming frame to a primary switch of an outgoing frame. These trunks are connected to switches of corresponding numbers on the two frames, thus facilitating marking of the network. Four-stage networks of this type are used in crossbar systems.

If a four-stage network with N incoming and N outgoing trunks is constructed with switches of size $n \times n$, then $N = n^3$ and the total number of switches is $4n^2$. Thus, the total number of crosspoints is:

$$\begin{aligned} C_4 &= 4n^2 . n^2 \\ &= 4\, N^{4/3} \end{aligned} \tag{5.24}$$

The number of crosspoints per incoming trunk is $4\, N^{1/3}$.

It should be noted that the partially interconnected three-stage network of Figure 5.14 corresponds to the four-stage network of Figure 3.17 truncated at the A links. Adding the fourth stage has not increased the number of trunks, although it has increased the number of crosspoints by one third. However, it has increased the number of paths between each incoming trunk and outgoing trunk from one to ten: i.e. a mixing stage has been added.

5.4.5 Discussion

If a network has N incoming trunks and N outgoing trunks, then the number of crosspoints per incoming trunk for a single stage is proportional to N, for a two-stage network it is proportional to $N^{1/2}$ and for a three-stage partially interconnected network it is proportional to $N^{1/3}$. Thus, for networks having many trunks, it is economic to use more stages than for networks with fewer trunks.

These networks have only distribution stages; there is only one path between an incoming trunk and an outgoing trunk. If more paths are needed to reduce blocking, mixing stages should be added, for example by changing a two-stage network to a fully interconnected three-stage network or changing a partially interconnected three-stage network to a four-stage network. Consequently, the design of a large switching network[13] involves providing distribution stages to obtain crosspoint economy and mixing stages to reduce blocking.

It is rarely possible to use switches of exactly the optimum sizes given by equations (5.11), (5.15), (5.18) and (5.21). Also, the cost of switches is not exactly proportional to the number of crosspoints. For example, doubling the inlet and outlet capacity of a crossbar switch requires four times as many crosspoints but only twice the number of magnets. In addition, the complexity and cost of the associated control equipment increases with the number of stages. Consequently, many switching networks use fewer stages and larger switches than would be indicated by the above equations.

5.5 Grades of service of link systems

5.5.1 General

A simple theory for calculating the probability of loss in link systems, due to C. Y. Lee,[14] will be given here. The method assumes that trunks and links being busy constitute independent random events. If two random events are independent, the probability of both happening at the same time is given by the product of their separate probabilities of occurrence at that time. If two links are to be connected in tandem, and the probability of one being busy is a and of the other being busy is b, then the probabilities of each being free are $1-a$ and $1-b$, respectively, so the probability of both being free is $(1-a)(1-b)$. Therefore, the probability of the path being blocked is $1 - (1-a)(1-b)$.

The occupancy at each stage is the total traffic carried divided by the number of links at that stage. However, if the loss is small (as it should be), little error is introduced by using the traffic offered instead of the traffic carried.

In a practical system the assumption of independence may not be valid, because there is usually some degree of dependence between links. This reduces the probability of blocking, because traffic peaks at different stages coincide more often than would happen if they were independent random events. This overlapping of peaks tends to reduce the total time during which blocking occurs. Consequently, Lee's method overestimates the loss probability. Nevertheless, the method gives reasonably accurate results in most cases. It also has the merit of simplicity. For these reasons, it is widely used. A more accurate, but still not exact, method was published by Jacobaeus.[6,15] This is described in Appendix 3.

An analytical treatment becomes complex for a network having a large number of stages and handling different kinds of traffic. Under these circumstances, it is necessary to resort to a computer simulation in order to determine the GOS with sufficient accuracy. An approximate theoretical calculation may be adequate to enable the designer to choose between alternative trunking schemes, but it will not be sufficiently accurate for determining the amounts of equipment to be provided in exchanges. A small percentage saving in cost on a large network will more than offset the cost of computation!

5.5.2 Two-stage networks

For a two-stage network, as shown in Figure 5.11, let the occupancy of the links be a and the occupancy of the outgoing trunks be b. (If the numbers of links and trunks are equal, then $a = b$.)

For mode 1 (i.e. connection to a particular outgoing trunk) only one link can be used. The probability of this being busy is a and this is the probability of loss. For example, to provide a GOS of $B_1 = 0.01$, each link and outgoing trunk could only carry 0.01 E. This is useless!

For mode 2 (i.e. connection to an outgoing route with one trunk on each secondary switch) any free link can be used. The probability of loss using a particular link is

$$1 - \text{probability that both link and trunk are free}$$
$$= 1 - (1 - a)(1 - b)$$

But there are g paths available. Assuming that each being blocked is an independent random event, the probability of simultaneous blocking for all g paths is:

$$B_2 = [1-(1-a)(1-b)]^g$$
$$= [a+(1-a)b]^g \tag{5.25}$$

where g is the number of secondary switches.

If connection may be made to any outgoing trunk that is free (i.e. mode 3) then it is possible to make the connection unless all the outgoing trunks are busy. Thus, if the numbers of incoming trunks, links and outgoing trunks are equal, no calls can be lost. However, this mode of operation is normally used with a concentrator. The number of incoming trunks is then much larger than the number of outgoing trunks, so the grade of service is given by:

$$B_3 = E_{1,N}(A)$$

where A is the total traffic offered to the network.

Example 5.6

1. Find the grade of service when a total of 30 E is offered to the two-stage switching network of Figure 3.15 and the traffic is evenly distributed over the 10 outgoing routes.

   The link and trunk occupancies are $a = b = 30/100 = 0.3$ E.

   $$B = [1 - (1 - 0.3)(1 - 0.3)]^{10} = 0.51^{10}$$
   $$= 0.0012$$

2. Find the traffic capacity of this network if the grade of service is not to exceed 0.01.

$$B \leq 0.01 = [1 - (1 - a)^2]^{10}$$
$$1 - (1 - a)^2 \leq 0.01^{0.1} = 0.631$$
$$\therefore \quad a \leq 0.39 \quad \text{and} \quad A \leq 39 \text{ E}$$

5.5.3 Three-stage networks

For a fully interconnected three-stage network (as shown in Figure 5.13) let:

Occupancy of A links be a
Occupancy of B links be b
Occupancy of outgoing trunks be c.

For mode 1 (i.e. connection to a particular outgoing trunk), the choice of a secondary switch determines the A and B links.

Probability that both links are free $= (1-a)(1-b)$

$\therefore$ Probability of blocking $= 1 - (1 - a)(1 - b)$

However, there are g_2 secondary switches.

$\therefore$ Probability that all g_2 independent paths are simultaneously blocked is

$$B_1 = [1 - (1 - a)(1 - b)]^{g_2} \tag{5.26}$$
$$= [a + (1 - a)b]^{g_2}$$

Thus, for similar occupancies, the three-stage network provides the same GOS for connections to individual trunks as the two-stage network does for connections to a group of trunks (cf. equation (5.25)).

For mode 2 (i.e. a connection to any free trunk in a route having one trunk connected to each tertiary switch):

Probability of blocking for a particular trunk

$$= 1 - (1 - B_1)(1 - c)$$
$$= B_1 + (1 - B_1)c$$

$\therefore$ Probability of simultaneous blocking for all g_3 independent paths is

$$B_2 = [B_1 + c(1 - B_1)]^{g_3} \tag{5.27}$$

where g_3 is the number of tertiary switches.

Example 5.7

1. Compare the grades of service provided by the two networks of Example 5.4 when each operates in mode 1 and is offered 30 E of traffic:

$$a = b = 30/100 = 0.3 \text{ E}$$

For network (a):

$$B = [1 - (1-0.3)(1-0.3)]^5 = 0.51^5 = 0.035$$

For network (b): $B = 0.51^{10} = 0.0012$

2. What is the traffic capacity of each network if the required grade of service is 0.01?

For network (a):

$$[1 - (1 - a)^2]^5 = 0.01$$
$$1 - (1 - a)^2 = 0.01^{1/5} = 0.398$$
$$\therefore \quad a = 0.224$$

Total traffic capacity $= 100 \times 0.224 = 22.4$ E

For network (b):

$$[1 - (1 - a)^2]^{10} = 0.01$$
$$1 - (1 - a)^2 = 0.01^{1/10} = 0.631$$
$$\therefore \quad a = 0.393$$

Total traffic capacity $= 100 \times 0.393 = 39.3$ E

For a partially interconnected three-stage network, as shown in Figure 5.14, there is only one path between an incoming trunk and an outgoing trunk. The probability that this is free is $(1 - a)(1 - b)$ and the probability of blocking is $1 - (1 - a)(1 - b)$.

For a connection to a trunk on an outgoing route with n trunks, each connected to a different frame, the probability of loss using a particular trunk is

$$1 - (1 - a)(1 - b)(1 - c)$$

But there are n such trunks available. Assuming that each being busy is an independent random event, the probability of simultaneous blocking for all paths is

$$B_2 = [1 - (1 - a)(1 - b)(1 - c)]^n \tag{5.28}$$

5.5.4 Four-stage networks

For a four-stage network, as shown in Figure 3.17 let:

Occupancy of A links be a
Occupancy of B links be b
Occupancy of C links be c
Occupancy of outgoing trunks be d

For a connection from a given inlet on an input frame to a particular outlet on an output frame (i.e. mode 1), the call may use any primary switch in the output frame. This switch is connected by a B link to only one secondary switch in the particular input frame. From this switch there is only one A link to the primary switch of the given incoming trunk.

Probability of this path being free is

$$(1 - a)(1 - b)(1 - c)$$

$\therefore$ Probability of this path being blocked is

$$1 - (1 - a)(1 - b)(1 - c)$$

Probability that all g_2 independent paths are simultaneously blocked is

$$B_1 = [1 - (1 - a)(1 - b)(1 - c)]^{g_2} \tag{5.29}$$

where g_2 is number of secondary switches in input frame = number of primary switches in output frame.

For a route of n outgoing trunks:

Probability of loss for a particular trunk

$$= 1 - (1 - B_1)(1 - d)$$
$$= B_1 + (1 - B_1)d$$

∴ Probability of simultaneous blocking for all n independent paths is

$$B_2 = [B_1 + d(1 - B_1)]^n \tag{5.30}$$

5.6 Application of graph theory to link systems

Switching networks, like other forms of network may be studied by means of the branch of mathematics known as graph theory.[16] Lee[14] used graphs to represent switching networks in 1955 and the method has subsequently been developed by Takaki[17] and others.[13]

A graph is a collection of points, known as *nodes* or *vertices*, connected by lines, known as *edges* or *arcs*. For example, the simple two-stage network shown in the trunking diagram of Figure 5.15(a) (whose links are connected as shown in Figure 3.15) may be represented by the graph shown in Figure 5.15(b). The latter shows the links without details of the switches, whereas Figure 5.15(a) focuses attention on the switches without showing the arrangement of the links.

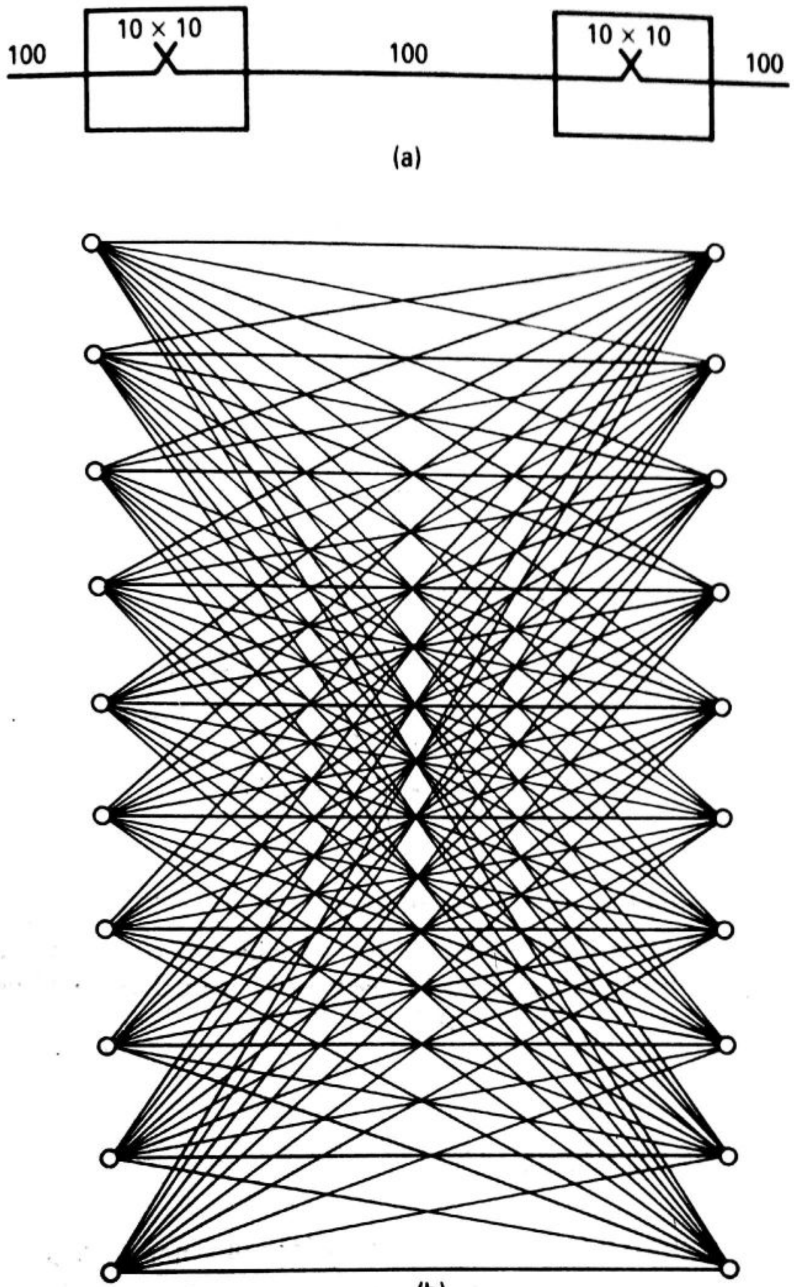
Most switching networks are symmetrical, so the arrangement of links connecting any particular inlet to any particular outlet is topologically the same as for any other inlet–outlet pair. The representation of a switching network by means of a graph can therefore be simplified by drawing only the paths which can be used for making connections between one particular inlet–outlet pair. This graph is called the *channel graph* of the network. Some examples are shown in Figure 5.16. Expressions for blocking probability (e.g. equations (5.26) and (5.29)) can be obtained directly by inspecting the channel graphs.

An important property of the network which is displayed by the channel graph is its *connectivity*. This may be defined as the minimum number of disjoint paths joining the non-adjacent vertices. Thus, in Figure 5.16(a) and (c) the connectivity is unity, whereas in Figure 5.16(b) and (d) its value is 10. Clearly, the larger the value of connectivity, the lower is the probability of blocking. This is shown by equations (5.26) and (5.29) which are of the form

$$B = (1 - x)^k$$

where $0 < x < 1$ and k is the connectivity of the channel graph.

Takagi[17] has developed a method of using channel graphs to design switching



(a)

(b)

Figure 5.15 Two-stage switching network. (a) Conventional representation. (b) Network graph.

networks. He provides a sufficient number of distribution stages to give access to the required number of outgoing trunks and then adds mixing stages to give sufficient connectivity for the required grade of service. The well-known networks shown in Figures 5.13 and 3.17 can be obtained by Takagi's method. The channel graphs in Figure 5.16(b) and (d) show how adequate connectivity is provided by adding a third stage to the two-stage network and a fourth stage to the partially interconnected three-stage network respectively. Takagi has applied his method to networks containing as many as eight stages.[17]

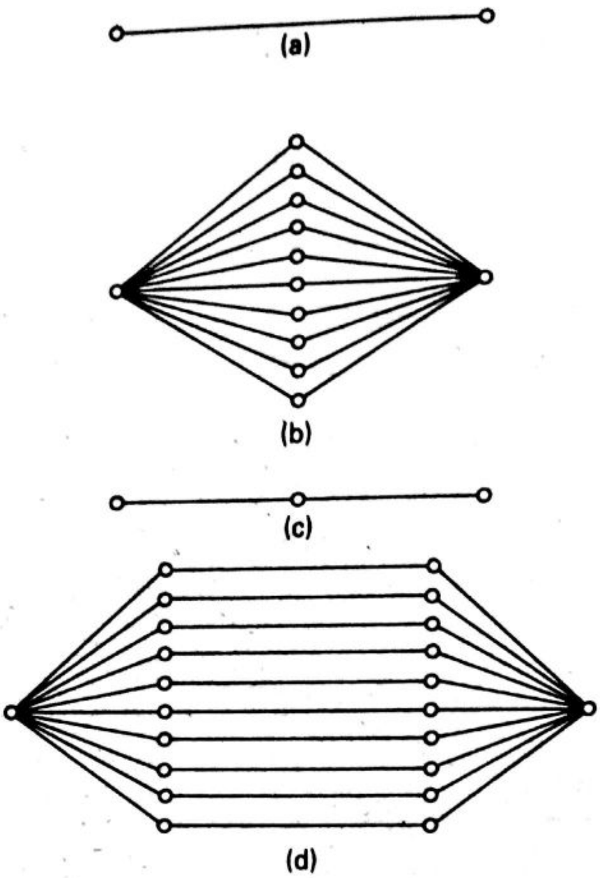


Figure 5.16 Channel graphs of switching networks. (a) Two-stage network of Figure 5.16. (b) Fully interconnected three-stage network of Figure 5.13. (c) Partially interconnected three-stage network of Figure 5.14. (d) Four-stage network of Figure 3.17.

5.7 Use of expansion

It has been shown in Section 5.6 that the blocking loss of a network depends strongly on its connectivity. The loss probability can therefore be reduced by increasing the connectivity by providing more links between switching stages than there are outgoing trunks. This is known as *expansion*.

Expansion is not used in local exchanges. Indeed, because customers' line are lightly loaded, concentration is used, as shown in Figure 5.17(a). Tandem exchanges, as shown in Figure 5.17(b), do not use concentration because junctions are usually heavily loaded. Neither is expansion used, because an adequately low grade of service is obtained without it. Long-distance circuits, particularly international ones, are expensive; however, trunks within an exchange (and their associated switches) are relatively cheap. It is undesirable for expensive circuits to be idle (and losing revenue) because of blocking within an exchange. Expansion is therefore used in trunk-transit and international exchanges, as shown in Figure 5.17(c), to ensure that loss due to blocking is much less than loss due to congestion of outgoing routes.

By using a moderate amount of expansion, the blocking loss of a switching

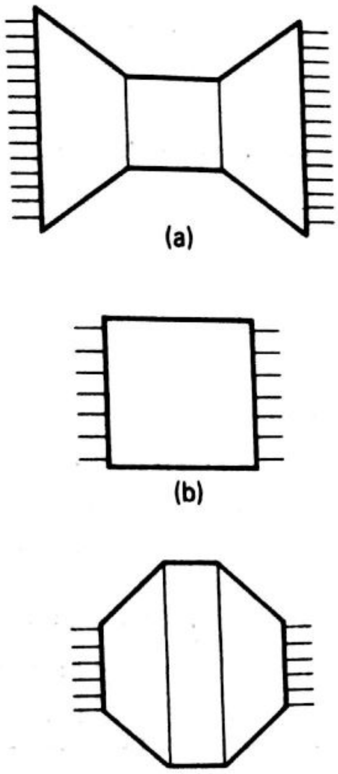


Figure 5.17 Use of concentration and expansion. (a) Local exchange. (b) Tandem exchange. (c) Trunk transit or international exchange.

network can be made negligible compared with the loss due to congestion of the external trunks. Such a network is called a *quasi-nonblocking network*. It will be shown in Section 5.10 that it is actually possible by means of expansion to reduce the blocking to zero.

Example 5.8

A fully interconnected three-stage network has 100 incoming trunks, 100 *A* links, 100 *B* links and 100 outgoing trunks. At each stage, it uses ten switches of size 10×10. As shown in Example 5.7, the grade of service in mode 1 is 0.01 when the link occupancy is 0.39 E.

For the same total offered traffic, what grade of service is obtained if the numbers of links and secondary switches are increased by (1) 20%, (2) 50%?

1. The link occupancy of the modified network is

$$b = c = 0.39/1.2 = 0.325 \text{ E}$$

From equation (5.26), the grade of service is:

$$B = [1 - (1 - 0.325)^2]^{12} = 0.544^{12} = 6.7 \times 10^{-4}$$

2. The link occupancy of the modified network is

$$b = c = 0.39/1.5 = 0.26 \text{ E}$$

From equation (5.26), the grade of service is:

$$B = [1 - (1 - 0.26)^2]^{15} = 0.452^{15} = 6.7 \times 10^{-6}$$

5.8 Call packing

In practice, less blocking is obtained when links are not allocated at random than when they are. The chance of a call being blocked is minimized if the number of possible paths available to it can be maximized. Thus, if each call to be set up is routed through the most heavily loaded part of the network which can still take it, subsequent calls will have a greater choice of paths than if the call were routed through a less heavily loaded part of the network. This is known as *call packing*. Simple as the principle is, there is no general proof that it gives better results than any other alternative. However, this can be demonstrated in a few elementary examples[2] and has been verified by computer simulation for some more complex networks.[18] Increases in traffic capacity of up to 10% have been obtained.[6]

A simple call-packing rule that can be applied to the three-stage network of Figure 5.13 is to select the lowest-numbered secondary switch which has free links to the required primary and tertiary switches. Thus, secondary switch 2 is only used if the A or B link to switch 1 is busy, and so on.

When call packing is used, calls are offered to switches in a predetermined order. A faulty first-choice switch can therefore cause serious degradation of service during periods of low traffic. (This is a well-known defect of the Strowger system.) To obviate this, common-control systems using crosspoint switches usually provide a second-attempt feature. When an attempt to set up a connection fails, the marker makes a second attempt. In order to avoid selecting the same faulty switch, it is necessary to search for suitable free links from a different starting point from that used in the first attempt.

5.9 Rearrangeable networks

Call packing reduces blocking by selecting the most heavily loaded part of a network for each new connection set up. A further reduction in blocking could be obtained if it were possible to ensure that connections which exist through lightly loaded parts of the network were cleared down before those through heavily loaded parts. Clearly, this is not possible, because call terminations are caused by the actions of customers. However, the equivalent result would be obtained if, every time a call ended, all the remaining connections were cleared down and set up again using a call-packing rule. If some of these connections were already made through a heavily loaded part of the network it would not be necessary to move them; it would probably be sufficient to alter only a few connections. Moreover, this need not be done each time a connection clears down; it can be done as part of the selection process when the next connection is set up. A network that is operated in this mode is called a *rearrangeable network*.

Benes has shown[2,19] that by using rearrangement it is possible to obtain networks which completely eliminate blocking. Such networks are said to be *nonblocking in the wide sense*. This means that zero blocking is not guaranteed by the structure of the network; it is obtained by means of the control algorithm used to set up connections.

If a multistage network with N incoming and N outgoing trunks is to avoid blocking, there must clearly be at least N links at each intermediate stage. Also, if a network having input switches with n inlets and output switches with n outlets is to be made nonblocking, its intermediate links must permit n simultaneous connections between each input switch and each output switch. The fully interconnected three-stage network of Figure 5.13 and the four-stage network of Figure 3.17 satisfy these conditions; the two-stage network of Figure 5.11 and the partially interconnected three-stage network of Figure 5.14 do not. The condition for a rearrangeable fully interconnected three-stage network to be nonblocking is simply $g_2 \geq n$. It can be shown[2] that the maximum number of existing connections to be moved to enable a new connection to be made is $n - 1$.

Rearrangeable networks are not used in space-division telephone exchanges. The clicks caused by the interruptions of current when connections were rearranged would be objectionable to users. Rearrangement could be used in time-division exchanges, because each connection is made and disconnected eight times every millisecond. A reduction in the number of crosspoints could thus be obtained. However, there is no incentive to obtain this reduction because crosspoints are used very economically by being time shared between a large number of connections, as described in Chapter 6. Moreover, time-division switching networks normally have a high connectivity, which provides very low blocking.

5.10 Strict-sense nonblocking networks

It was shown in Section 5.9 that rearrangeable networks can have zero blocking loss. These networks are nonblocking in the wide sense, because zero blocking is obtained by using a prescribed control algorithm. If a network can never have blocking, no matter what existing connections are present and without any need to rearrange connections, the network is said to be *nonblocking in the strict sense*.

The single-stage network of Figure 5.1 is obviously strictly nonblocking. A connection can always be made to a free outgoing trunk.

In order to make the two-stage network of Figure 5.11 strictly nonblocking, it is necessary to make the number of links from a primary switch to a secondary switch equal to the number of outlets of the secondary switch. If the network has N incoming and outgoing trunks, the primary switches have n inlets and the secondary switches

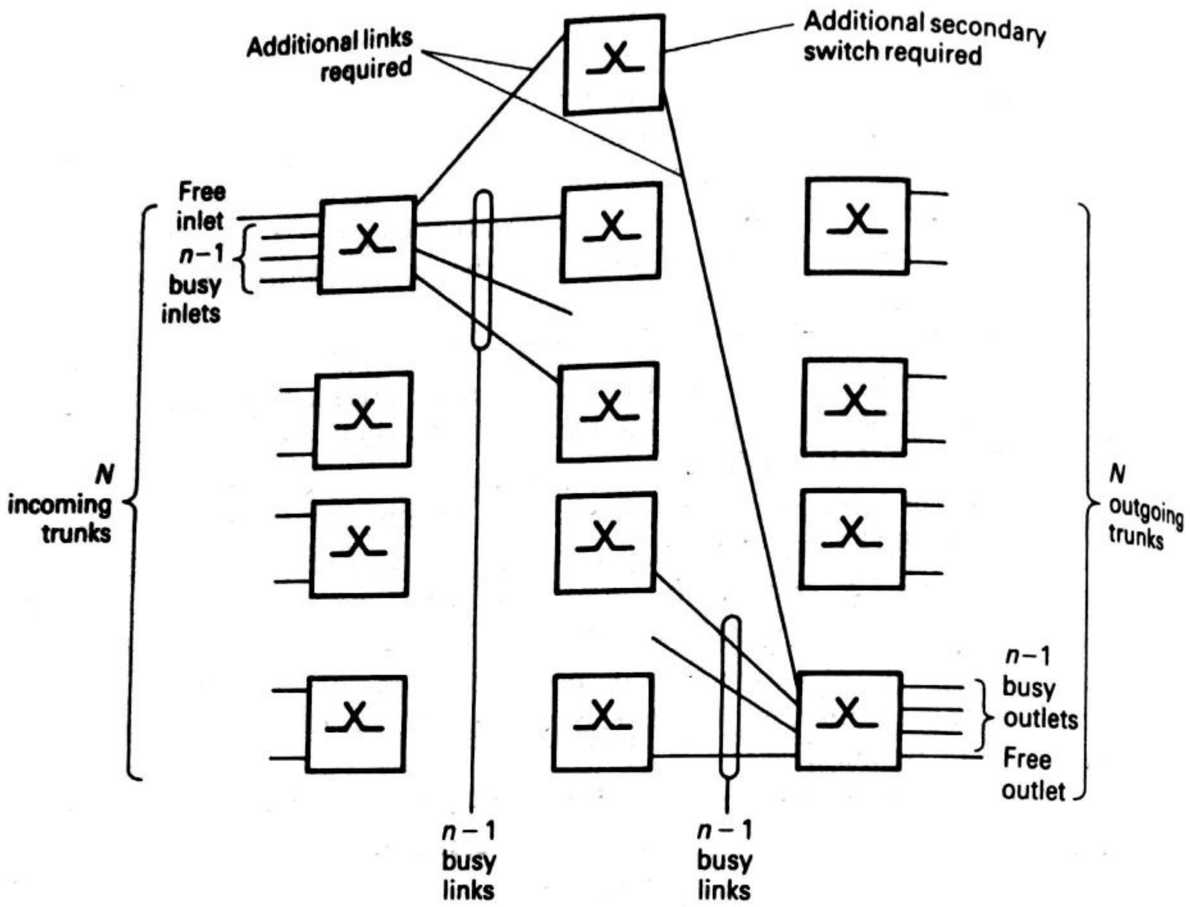


Figure 5.18 Principle of three-stage strictly nonblocking network.

have n outlets (where $n = \sqrt{N}$), then the primary switches need N outlets and the secondary switches need N inlets. Thus, each needs nN crosspoints. The total number of crosspoints is: $C_2 = 2n \times nN = 2N^2$. This is twice as many as needed by a single stage!

If an attempt is made to modify the four-stage network of Figure 3.17 to make it strictly nonblocking, a similar change must be made to each of its switches. Thus, the total number of crosspoints is again greater than that required for a single stage.

Networks having odd numbers of stages can be made strictly nonblocking while requiring fewer crosspoints than a single stage. A method of designing such networks was discovered by Clos[20] . These networks are therefore often called *Clos networks*.

Consider a three-stage network, as shown in Figure 5.18, in which each primary switch has n inlets and each tertiary switch has n outlets. The greatest number of calls that a primary switch can be carrying and still be capable of accepting another is $(n - 1)$. These calls will occupy $(n - 1)$ links to different secondary switches. Similarly, the selected tertiary switch can be carrying $(n - 1)$ calls and occupying $(n - 1)$ links from secondary switches. The worst case for blocking will occur when the busy links from the primary switch and the busy links to the tertiary switch terminate on different secondary switches. In order to be able to make the new connection, there must still be one secondary switch to which there are free links. The minimum number of secondary switches, g_2, required is therefore

$$g_2 = (n-1) + (n-1) + 1 = 2n - 1 \quad (5.31)$$

Then:

No. of primary switches = N/n
No. of secondary switches = $2n-1$
No. of tertiary switches = N/n
No. of crosspoints per primary switch = $n(2n-1)$
No. of crosspoints per secondary switch = $(N/n)^2$
No. of crosspoints per tertiary switch = $(2n-1)n$
and the total no. of crosspoints is:

$$C_3 = (2n-1)\,[2N + (N/n)^2] \quad (5.32)$$

Differentiating equation (5.32) w.r.t. n and equating to zero gives:

$$2n^3 - Nn + N = 0 \quad (5.33)$$

This has integer roots: $n = 2$, $N = 16$ and $n = 3$, $N = 27$. However, if $n \gg 1$, equation (5.33) approximates to $2n^2 - N = 0$ and $n = \sqrt{N/2}$. Substituting in equation (5.32) gives:

$$C_3 = 2^{5/2}\,N^{3/2} - 4N \quad (5.34)$$

(Clos[20] used the nonoptimum value $n = \sqrt{N}$ and obtained $C_3 = 6N^{3/2} - 3N$.)

It follows that a three-stage nonblocking network has fewer crosspoints than a single stage if $N \geq 28$. However, comparing equation (5.34) with equation (5.19) shows that a three-stage nonblocking network contains nearly twice as many crosspoints as a conventional fully interconnected three-stage network.

For networks having large values on N, fewer crosspoints are required if more than three stages are used. A five-stage network[20] can be considered as a three-stage network in which each secondary switch is replaced by a 'level' which itself consists of a three-stage nonblocking network, as shown in Figure 5.19. The number of such levels required is, of course, $2n - 1$. It can be shown that the number of crosspoints (C_5) in this network is a minimum when $n = (2N)^{1/3}$. Then:

$$C_5 = 3 \times 2^{7/3}\,N^{4/3} - 14N + 2^{5/3}\,N^{2/3} \quad (5.35)$$

Similarly, a seven-stage network[20] can be designed as a three-stage network in which each secondary stage is replaced by a level containing five stages. The method can be extended in this way to design nonblocking networks having any odd number of stages.

Nonblocking networks are not used in commercial space-division switching systems because they are uneconomic. However, they have been used in certain military applications where economic considerations are secondary to operational

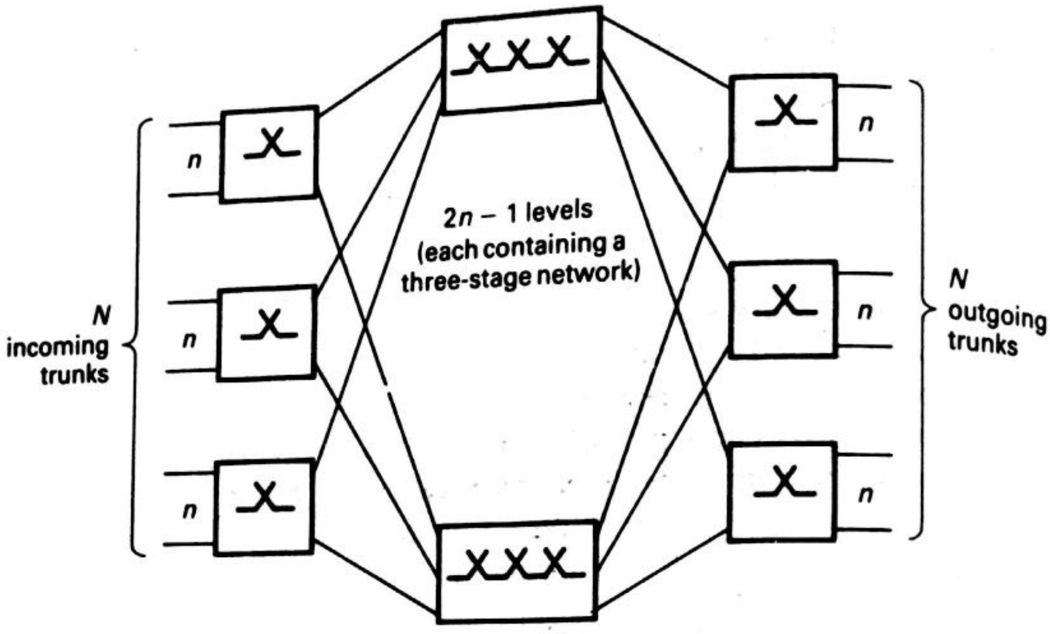


Figure 5.19 Five-stage strictly nonblocking network.

requirements. It can be economic to use nonblocking networks in time-division switching systems, since crosspoints are time-shared between a large number of connections.

Example 5.9

Design a strictly nonblocking network for 100 incoming and 100 outgoing trunks.

Choose a three-stage network. The minimum number of crosspoints is obtained when

$$n = \sqrt{N/2} = \sqrt{50} = 7.07 \therefore \text{ use } n = 5$$

Thus, the number of secondary switches is $2n-1 = 9$ and the network is as shown in Figure 5.20(a).

The total number of crosspoints is:

$$C_3 = 20 \times 5 \times 9 + 9 \times 20 \times 20 + 20 \times 9 \times 5 = 5400$$

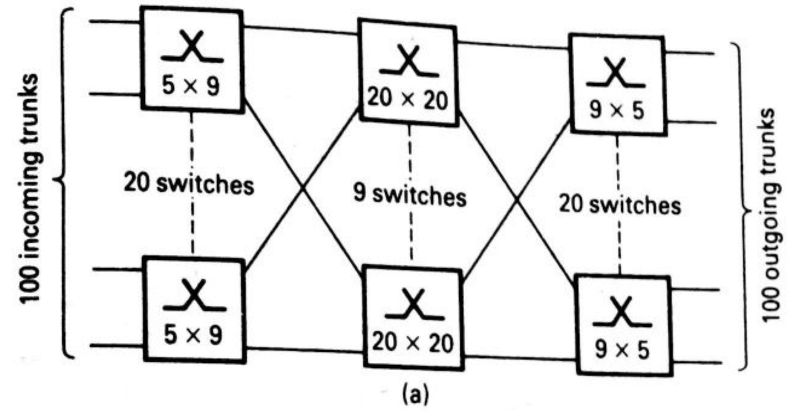
Example 5.10

Design a strictly nonblocking network for 1000 incoming and 1000 outgoing trunks.

Choose a five-stage network. The minimum number of crosspoints is obtained when

$$n = (2N)^{1/3} = 2000^{1/3} = 12.599 \therefore \text{ use } n = 10$$

The number of levels needed is $2n-1 = 19$.



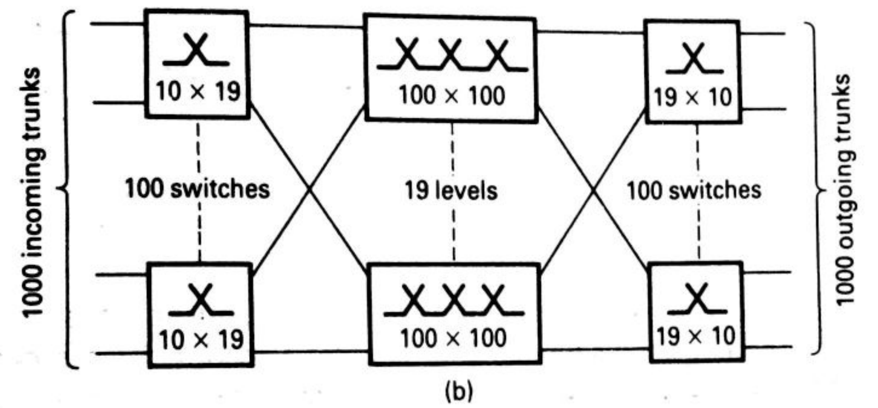


Figure 5.20 Examples of strictly nonblocking networks. (a) Three-stage network with 100 incoming and outgoing trunks. (b) Five-stage network with 1000 incoming and outgoing trunks. (Each level consists of a three-stage network as shown in Figure 5.20(a).)

No. of switches in 1st stage = no. of switches in 5th stage = 1000/10 = 100.

Thus, the network is as shown in Figure 5.20(b). Each level has 100 incoming and 100 outgoing links. Therefore, it can consist of a three-stage network as shown in Figure 5.20(a).

The number of crosspoints is:

$$\begin{aligned} C_5 &= 100 \times 10 \times 19 + 19 \times 5400 + 100 \times 19 \times 10 \\ &= 140\,600 \end{aligned}$$

5.11 Sectionalized switching networks

Since trunks in small groups normally have a lower traffic capacity than trunks in large ones, the division of a large switching network into smaller sections that are not interconnected would appear to be inefficient. However, it is possible to provide an electronic central control to determine which of the networks is best able to deal with

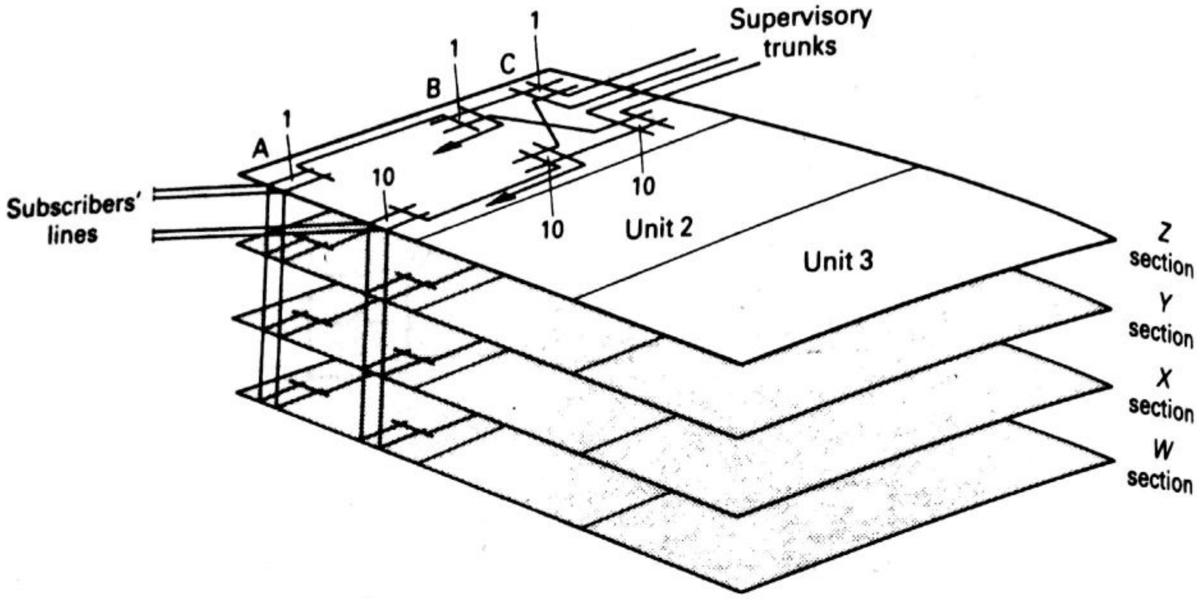


Figure 5.21 Sectionalized switching network.

each call. By doing so, it can coordinate the combined activities of the separate sections as if they were a homogeneous whole. This enables a sectionalized exchange to use fewer crosspoints than the equivalent unsectionalized switching network.[21]

Figure 5.21 shows a sectionalized switching network containing three separate units, each serving up to 1000 subscribers' lines. Each unit consists of a partially interconnected three-stage switching network of the form shown in Figure 5.14, but having concentration. The network contains four sections (shown as planes W, X, Y, Z in the figure). Each section provides only one path between each inlet (subscriber's line) and each outlet (supervisory trunk). It also contains its own marker. The supervisory trunks provide connections between different units in the same plane and between units in different planes.

Call packing is used. In order to set up a connection, the central control instructs the marker in each section which trunks and links to interrogate. The resulting pattern of busy/free conditions is signalled to the central control for choice of the route to be used. This choice is made in accordance with a program of priorities to select the route least likely to cause blocking of subsequent calls. The marker in the selected section then sets up the connection.

The trunking shown in Figure 5.21 can be considered as a number of minimal exchanges connected in parallel. It enables exchanges to be constructed from standard modules, each consisting of one section of one unit. The number of units provided depends on the number of lines and the number of sections (planes) depends on the traffic. Failure of one section results in a degradation of service to all customers, but it does not cause loss of service to any. Thus, the only equipment that needs to be replicated to ensure security is the central control unit.

The sectionalized three-stage network shown in Figure 5.21 is employed in the

British Telecom TXE4 reed–electronic system.[12] It uses from six to eight planes. A fourth switching stage (the D switch), which is shown in Figure 3.24, acts as a mixing stage to improve the occupancy of the supervisory trunks.

Switching networks consisting of separate planes in parallel are now used in some digital time-division switching systems. Usually, there are two planes. If one fails, the other can carry the traffic.

Note

1. This would occur if originating calls were allotted to outgoing trunks by a perfectly random selection process.

Problems

1. Explain briefly the meanings of the following terms applied to gradings: graded group, availability, progressive grading, skipped grading, homogeneous grading.

2. A grading is required to connect 30 outgoing trunks to switches of availability 10.
(a) Design a progressive grading.
(b) Design a homogeneous grading.

3. (a) Estimate the traffic capacity of the above gradings if the grade of service is to be:
(i) 0.001, (ii) 0.01.
(b) If the gradings are each offered 12 E of traffic evenly distributed over the graded groups, how much traffic is carried by the first-choice trunks and the second-choice trunks?

4. Define the following terms: Full availability, link system, blocking, conditional selection, distribution stage, mixing stage.

5. (a) Sketch a two-stage network using switches of size 3 × 4 to connect any of 9 incoming trunks to any one of 16 outgoing trunks. How many crosspoints are needed?
(b) Sketch a fully interconnected three-stage network for the same number of incoming and outgoing trunks, using 3 × 4 switches in the primary stage and 4 × 4 switches in the tertiary stage. How many crosspoints are needed? Comment on this result.
(c) Demonstrate by means of a simple example that the probability of blocking is less with network (b) than with network (a).

6. (a) Design a two-stage network, using switch modules of size 10 × 10, suitable for connecting 200 incoming trunks to 200 outgoing trunks. The outgoing trunks serve ten routes with 20 trunks on each route. How many switch modules are needed?
(b) The occupancies of the incoming and outgoing trunks are each 0.5 E. Estimate the grade of service obtained.

7. (a) A two-stage space-division network acts as a concentrator. It has M incoming trunks, N outgoing trunks and N links between the two switching stages (where $M > N$). Two methods of designing the network are as follows:

(i) To use the same number of switches in each stage, but to have larger switches in the first stage.
(ii) To use switches in the primary stage which have the same number of inlets as the secondary switches have outlets, but to use a different number of switches in each stage.

Which requires the smaller number of crosspoints?
(b) Verify the answer to part (a) by designing

a network of each type for a two-stage concentrator having 400 incoming trunks and 100 outgoing trunks. How many crosspoints does each contain?

(c) The concentrator serves 10 outgoing routes with ten trunks on each route. Each incoming trunk originates 0.1 E of traffic, of which the same proportion is directed to each outgoing route. Estimate the grade of service given by each network and state any assumptions you make. Comment briefly on the results obtained.

8. A three-stage fully interconnected switching network is to connect 600 incoming trunks to 100 outgoing trunks. It is to use switches assembled from blocks of size 5 × 5. Design a suitable network and determine the number of switch blocks required.

9. A switching network is to have N incoming trunks, N outgoing trunks and not less than N links between stages.

(a) Show that a two-stage network has fewer crosspoints than either a single stage or a partially interconnected three-stage network for $6 \le N \le 19$.

(b) The cost per switch is $K(1 + C_1/100)$, where C_1 is its number of crosspoints (e.g. the total cost of a 20 × 20 switch is 25% more than the cost of its crosspoints). Show that a two-stage network costs less than either a single-stage or a partially interconnected three-stage network for $42 \le N \le 400$.

10. Define the following terms: Connectivity, call packing, expansion, quasi-nonblocking network, wide-sense nonblocking network, strict-sense nonblocking network.

11. (a) Design a fully interconnected three-stage network with 200 incoming and 200 outgoing trunks with the minimum number of crosspoints. How many crosspoints does the network contain?

(b) If the occupancy of the trunks is 0.4 E and connections are to be made to particular outgoing trunks, estimate the grade of service. State any assumptions you make.

(c) Redesign the above network to provide a grade of service better than 1 in 1000. How many crosspoints does this network require?

(d) Redesign the network to be strictly nonblocking. How many crosspoints does this network require?

12. A partially interconnected three-stage network as shown in Figure 5.14 consists of switches of size 10 × 10 and connects 1000 incoming trunks to 1000 outgoing trunks. The outgoing trunks serve ten routes, each having one trunk connected to each tertiary switch.

(a) (i) Compare the number of crosspoints in the network with the number required by a three-stage fully interconnected network with the same number of incoming and outgoing trunks.
(ii) What advantage would the fully interconnected network have over that shown in Figure 5.14?

(b) (i) Assuming that the traffic is evenly distributed over the incoming and outgoing trunks, determine the grade of service when each incoming trunk originates 0.6 E of traffic.
(ii) If the total traffic offered to the network is unchanged but the traffic offered to one primary switch increases by 20%, find the grade of service for these calls.
(iii) If the total traffic offered to the network is unchanged but the traffic offered to one outgoing route increases by 20%, find the grade of service for calls offered to that route.

References

[1] Syski, R. (1986), *Introduction to Congestion Theory in Telephone Systems*, 2nd edn, Oliver & Boyd, Edinburgh.

[2] Benes, V.E. (1965), *Mathematical Theory of Connecting Networks and Telephone Traffic*,

CHAPTER 6
Time-division switching

6.1 Introduction

As described in Chapter 3, the first application of digital time-division switching was to provide tandem switching of PCM junction and trunk circuits[1]. Examples of such systems[2] are the Bell ESS No.4 system and the French E 12 system. Since a tandem or trunk exchange is similar to the route switch in a local exchange, local-exchange systems were developed by adding reed-relay space-division concentrators. Examples of such systems[2] include the initial versions of System X and the AXE 10 and E 10 systems. However, developments in the technology of solid-state integrated circuits eventually solved the BORSCHT problem and enabled TDM concentrators to replace space-division concentrators, thus extending digital operation to the customer's line circuit.[2–5]

This has led to the evolution of *integrated digital networks* (IDN), in which compatible digital transmission and switching are used throughout a network. Finally, extension of digital transmission over customers' lines has enabled *integrated-services digital networks* (ISDN) to be introduced. An ISDN can provide the customer with a wide variety of services, based on 64 kbit/s transmission, over a single line from a local exchange.

As an example, the architecture of a System X local exchange[6] is shown in Figure 6.1. The digital switching subsystem (DSS) corresponds to the route switch of Figure 3.20. The line-terminating units of PCM junctions are connected to it directly. Voice-frequency junctions are connected via a signalling interworking subsystem (SIS) and an analog line terminating subsystem (ALTS) that provides analog/digital and digital/analog conversion. Analog and digital customers' lines are connected to the DSS via a concentrator, known as the digital subscribers' switching subsystem (DSSS). These concentrators may be in the main exchange or located remotely. The processor utility subsystem uses software built largely from modules corresponding to the hardware subsystems that they control, as shown in Figure 6.1. The system has a capacity for 60000 lines and 10000 E of traffic. A tandem or trunk exchange uses a similar DSS, but it has no concentrators.

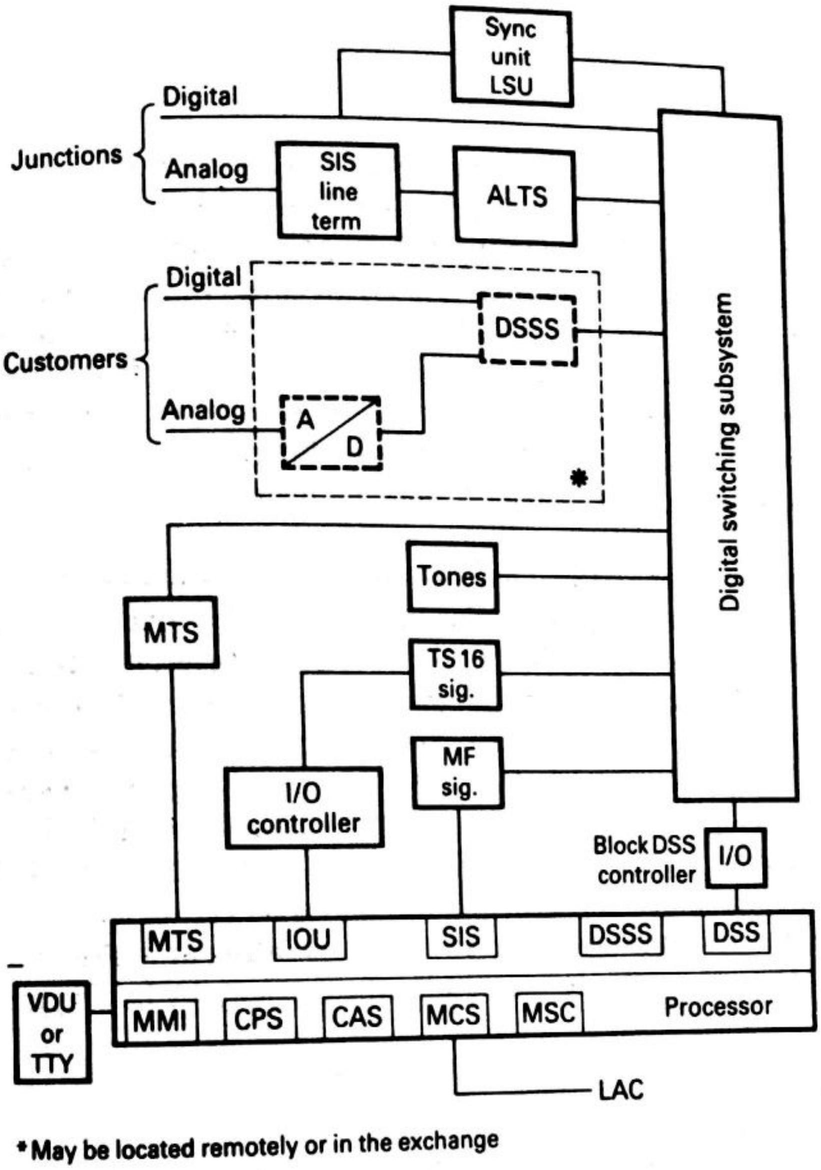


Figure 6.1 System X local exchange. ALTS = analog line terminating subsystem, CAS = call-accounting subsystem, CPS = call-processing subsystem, DSS = digital switching subsystem, DSSS = digital subscribers' switching subsystem, LAC = local administration centre, MCS = maintenance control subsystem, MMI = man–machine interface, MSS = management statistics subsystem, MTS = message transmission subsystem, SIS = signalling interworking subsystem, VDU = visual display unit.

A broadly similar architecture is used by other digital telephone-exchange systems.[2,7] Examples of such systems include the AXE-10 system (developed in Sweden), the DMS-10 system (Canada), the E-10 system (France), the No. 5 ESS system (USA), the EWS-D system (Germany) and the NEAX system (Japan).

6.2 Space and time switching

6.2.1 General

A tandem switching centre, or the route switch of a local exchange, must be able to connect any channel on one of its incoming PCM highways to any channel of an outgoing PCM highway. The incoming and outgoing highways are spatially separate, so the connection obviously requires space switching. In general, a connection will occupy different time-slots on the incoming and outgoing highways. Thus, the switching network must be able to receive PCM samples from one time-slot and retransmit them in a different time-slot. This is known as *time-slot interchange*, or simply as *time switching*. Consequently, the switching network of a tandem exchange, or the route switch of a local exchange, must perform both space switching and time switching.

Simple time-division switching networks make connections between channels on highways carrying a primary multiplex group, i.e they operate at 1.5 Mbit/s or 2 Mbit/s. A 2 Mbit/s line system has 32 time-slots. However, it only carries 30 speech channels; time-slot 0 is used for frame alignment and time-slot 16 for signalling. Within an exchange, time-slot 0 is not needed for frame alignment since all switches are driven synchronously from the clock-pulse generator of the exchange. It is also unnecessary to use time-slot 16 for signalling associated with the channels on the highway when this is handled over a separate path (e.g. when common-channel signalling is used, as described in Chapter 8). In this case, all 32 time-slots can be used to switch speech connections. Some systems have large switches operating at multiples of the primary rate (e.g. at 8 Mbit/s) in order to increase traffic capacity by having more time-slots.

6.2.2 Space switches

Connections can be made between incoming and outgoing PCM highways by means of a crosspoint matrix of the form shown in Figure 3.11. However, different channels of an incoming PCM frame may need to be switched by different crosspoints in order to reach different destinations. The crosspoint is therefore a two-input AND gate. One input is connected to the incoming PCM highway and the other to a *connection store* that produces a pulse at the required instants. A group of crosspoint gates can be implemented as an integrated circuit, for example by using a multiplexer chip.

Figure 6.2 shows a space switch with k incoming and m outgoing PCM highways, each carrying n channels. The connection store for each column of crosspoints is a memory with an address location for each time-slot, which stores the number of the crosspoint to be operated in that time-slot. This number is written into the address by the controlling processor in order to set up the connection. The numbers are read out cyclically, in synchronism with the incoming PCM frame. In each time-slot, the number stored at the corresponding store address is read out and decoding logic converts this into a pulse on a single lead to operate the relevent crosspoint.

Since a crosspoint can make a different connection in each of the n time-slots, it is

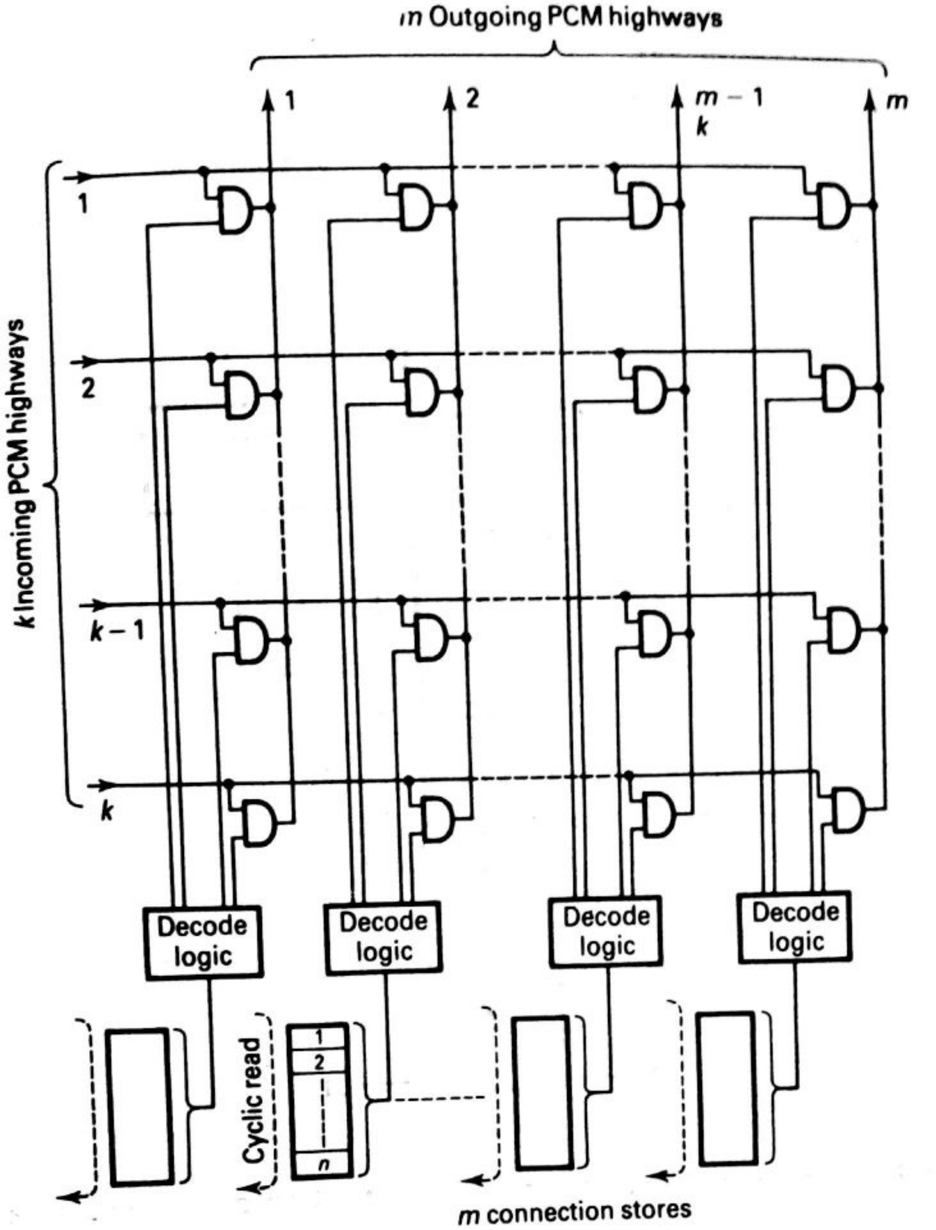


Figure 6.2 Space switch.

equivalent to n crosspoints in a space-division network. The complete space switch is thus equivalent to n separate $k \times m$ switches in a space-division switching network.

6.2.3 Time switches

The principle of a time switch is shown in Figure 6.3(a). It connects an incoming n-channel PCM highway to an outgoing n-channel PCM highway. Since any incoming channel can be connected to any outgoing channel, it is equivalent to a space-division crosspoint matrix with n incoming and n outgoing trunks, as shown in Figure 6.3(b).

Time-slot interchange is carried out by means of two stores, each having a storage address for every channel of the PCM frame. The *speech store* contains the data of each

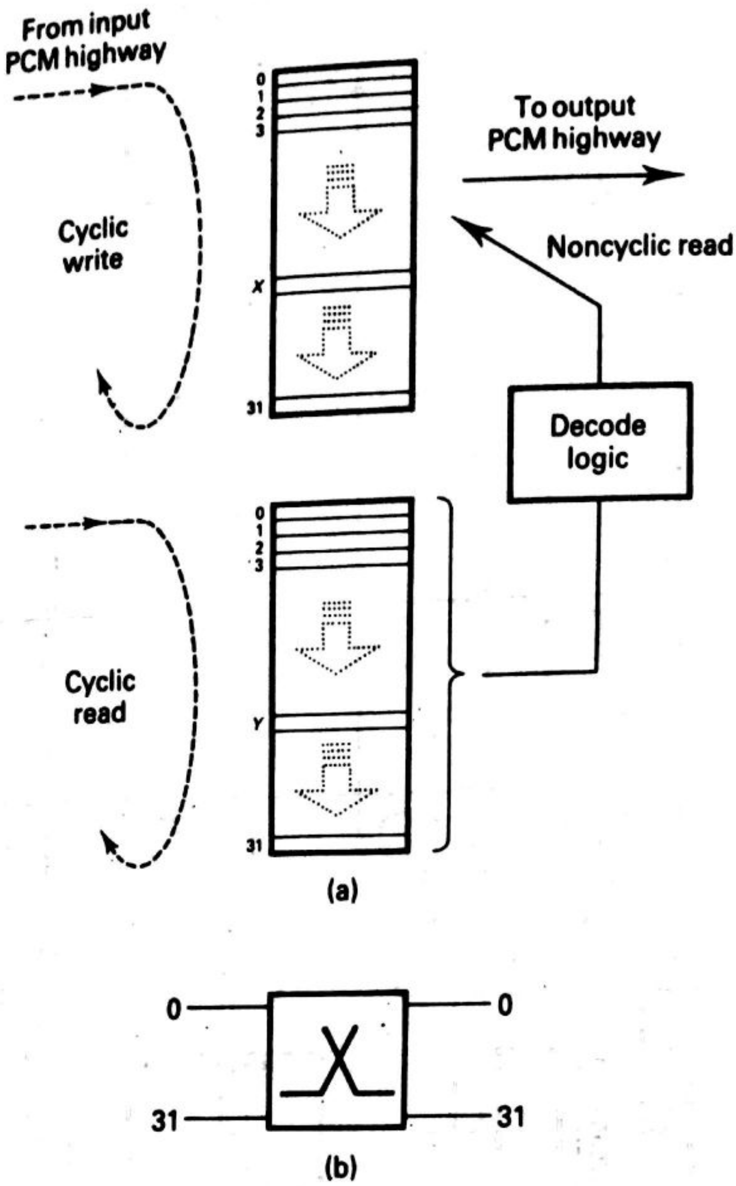


Figure 6.3 Time switching. (a) Time switch. (b) Space-division equivalent.

of the incoming time-slots (i.e. its speech sample) at a corresponding address. Each address of the *connection store* corresponds to a time-slot on the outgoing highway. It contains the number of the time-slot on the incoming highway whose sample is to be retransmitted in that outgoing time-slot. Information is read into the speech store cyclically, in synchronism with the incoming PCM system; however, random-access read-out is used. The connection store has cyclic read-out, but writing in is noncyclic.

To establish a connection, the number (X) of the time-slot of an incoming channel is written into the connection store at the address corresponding to the selected outgoing channel (Y). During each cyclic scan of the speech store, the incoming PCM sample from channel X is written into address X. During each cyclic scan of the connection store, the number X is read out at the beginning of time-slot Y. This is decoded to select address X of the speech store, whose contents are read out and sent over the outgoing highway.

An alternative way of implementing a time switch uses a speech store with

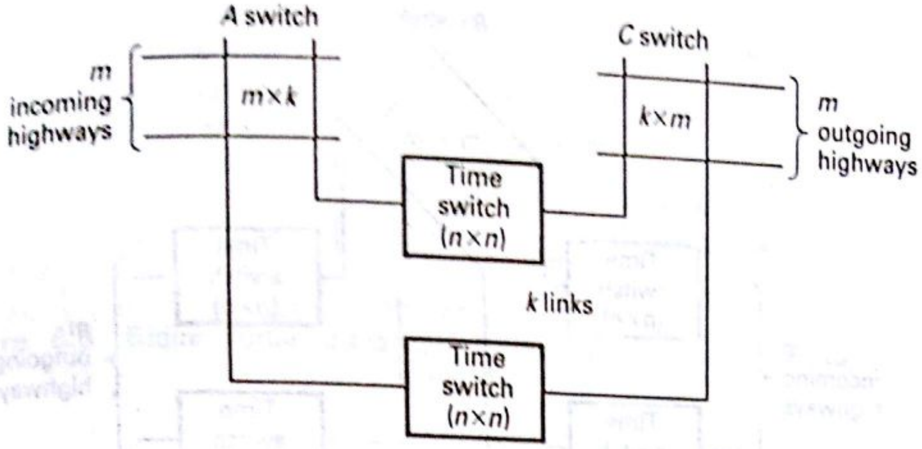


Figure 6.4 Space–time–space (S–T–S) switching network. m = no. of PCM highways, n = no. of time-slots.

random access for writing and cyclic access for reading. In order to transfer data from time-slot X on the incoming highway to time-slot Y on the outgoing highway, the connection store holds Y at address X. This is therefore read out at time X and decoded to write the incoming sample in the speech store at address Y. The cyclic scan of the speech store then reads out the sample at time Y for retransmission on the outgoing highway.

Time switching introduces delay. If $Y > X$, the output sample occurs later in the same frame as the input sample. If $Y < X$, the output sample occurs in the next frame. In a multi-link connection, several such delays occur. Since these are in addition to propagation delay, they adversely affect the echo performance of the connection.

6.3 Time-division switching networks

6.3.1 Basic networks

Figure 6.4 shows a space–time–space (S–T–S) switching network. Each of the m incoming PCM highways can be connected to k links by crosspoints in the A switch, and the other ends of the links are connected to the m outgoing PCM highways by crosspoints in the C switch. Each link contains a time switch. To make a connection between time-slot X of an incoming PCM highway and time-slot Y of an outgoing highway, it is necessary to select a link having address X free in its speech store and address Y free in its connection store. The time switch is then set to produce a shift from X to Y. The connection is completed by operating the appropriate A-switch crosspoint at time X and the appropriate C-switch crosspoint at time Y in each frame.

Figure 6.5 shows a time–space–time (T–S–T) switching network. Each of the m incoming and m outgoing PCM highways is connected to a time switch. The incoming and outgoing time switches are connected by the space switch. To make a connection between time-slot X of an incoming highway and time-slot Y of an outgoing highway, it is necessary to choose a time-slot Z which is free in the connection store of the incoming

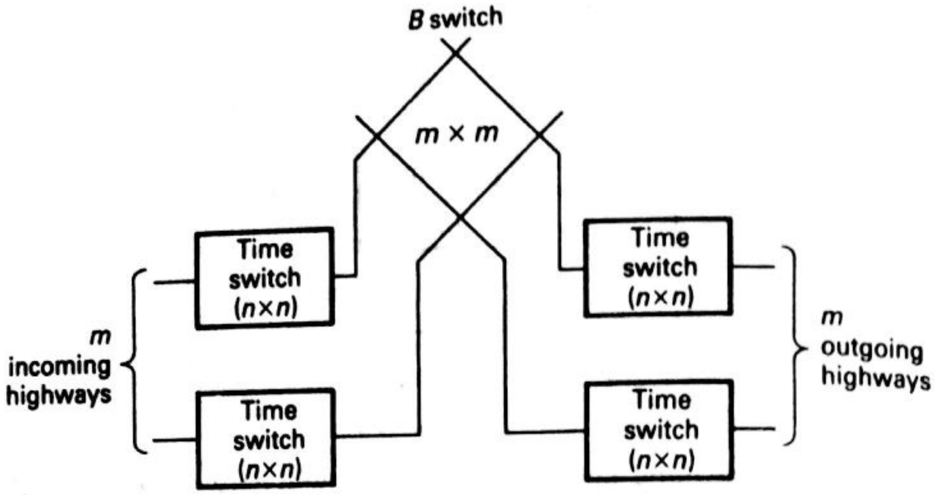


Figure 6.5 Time–space–time (T–S–T) switching network. m = no. of PCM highways, n = no. of time-slots.

highway and the speech store of the outgoing highway. The connection is established by setting the incoming time switch to shift from X to Z, setting the outgoing time switch to shift from Z to Y and operating the appropriate crosspoint at time Z in each frame.

Early designs of digital switching systems used S–T–S networks. This was because storage of speech samples, and hence time switching, was expensive. With the advent of semiconductor memories, time switching is no longer expensive. As a result, most current systems use T–S–T networks. However, a small switching network may need only two stages (e.g. T–S). Large switching networks may have more than three stages. For example, the Bell No.4 ESS toll switching system[8] uses a five-stage (T–S–S–S–T) network.

6.3.2 Bidirectional paths

The switching networks described above provide a connection for only one direction of transmission. Since PCM transmission systems use four-wire circuits, it is necessary to provide separate paths for the 'send' and 'receive' channels. One way of doing this would be to provide a separate switching network for each direction of transmission. However, this may be avoided by connecting the 'send' highways of both incoming and outgoing circuits to one side of the switch and the 'receive' highways to the other side, as shown in Figure 6.6.

In an S–T–S network the same speech-store address in the time switch may be used for each direction of transmission. For a connection between time-slot X on one trunk and channel Y on another, for one direction of transmission, the contents at the address are written at the end of time-slot X and are read at the beginning of time-slot Y. For the opposite direction of transmission, they are written at the end of the same

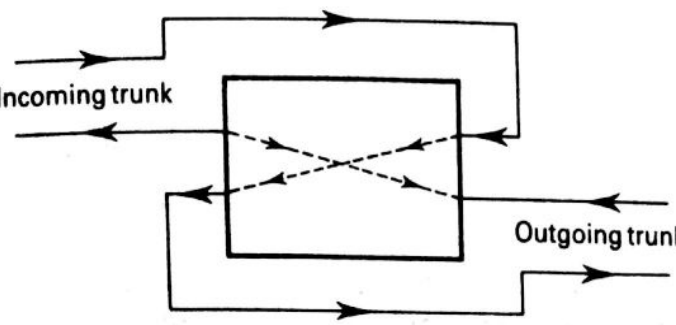


Figure 6.6 Bidirectional transmission through time-division switching network.

time-slot Y and read at the beginning of the next time-slot X. This method cannot be used if both external circuits use the same time-slots; however, this is rarely necessary.

In a T–S–T network, speech in the two directions must be carried through the space switch using different time-slots. In order to simplify control of the switching network, the time-slots for the two directions of transmission have a fixed time difference. Usually, the time-slots have a phase difference of 180°. In a 32-channel system, if time-slot 12 is used for one direction of transmission, then time-slot $(12 + 16) = 28$ is used for the reverse direction. One advantage of this arrangement is that if one time-slot is found to be free, the associated time-slot will also be free. Since the same time-slot is received from the input highway of a junction and sent to its output highway, the same connection store can be used to control the time switches of both. Speech stores associated with input highways have cyclic read-in and noncyclic read-out (as described in Section 6.2.3). However, speech stores associated with output highways have noncyclic read-in and read out cyclically to the output highways.

6.3.3 More complex switching networks

Many variations are possible on the basic three-stage T–S–T network shown in Figure 6.5. These include:

1. Increasing the size of stores in the time switch, so that each can serve more than one PCM highway.
2. Operating the space switch at a higher bit rate than the incoming and outgoing PCM systems. This enables each crosspoint to serve more than 32 channels, thus effectively increasing the size of the switch.
3. Using parallel instead of serial transmission of PCM words through the space switch. This has the same effect as (1) without an increase in speed; however, it increases the numbers of highways and crosspoint gates.
4. Duplicating, or even triplicating, the switching network to increase security of operation in the presence of faults. This is uneconomic for a space-division network. It is also unnecessary, because failure of individual switches has little effect on the overall grade of service. It is practicable in a time-division switching network because of the cost reduction brought about by time-sharing. It is also

desirable, because failure of an individual unit (e.g. a connection store) has much more serious consequences than in a space-division switching network.

Each of these techniques is used in the Mark 1 digital switching sub-system of System X,[9] shown in Figure 6.7. The receive and send time switches each have a speech store containing 1024 locations and so can serve up to 32 PCM systems. The complete network can contain up to 96 time switches, thus serving up to 3072 PCM systems.

The space switch therefore requires a maximum size of 96 × 96 and each crosspoint must be able to switch 1024 channels. This is done by using parallel transmission at a digit rate of 8.194 Mbit/s. To minimize problems of pulse distribution, the 1024-channel highways are each split into two highways of 512 channels. The space switch is therefore divided into two segments (A and B), switching odd and even time-slots respectively.

To provide adequate security, the complete network is duplicated. The duplicated systems operate in synchronism and faults are detected by means of a parity check. The PCM words transmitted across the switching network therefore use 9 bits, instead of the normal 8 bits used for transmission.

This Mark 1 digital switching subsystem has now been superseded by a more complex Mark 2 subsystem.[10] It uses a five-stage switching network.

6.3.4 Concentrators

A concentrator connects to a PCM highway a number of customers' line units greater than the number of time-slots on the highway. In a simple concentrator, the customers' codecs are all connected to the common highway and each may use any time-slot. A codec is operated in the required time-slot by means of a connection store. This method is used, for example, in the AXE system[2,4]. However, each 128-line concentrator module gives access to two PCM highways. One is individual to it and, when all its channels are busy, calls overflow to a second highway that is common to 16 modules.

Alternatively, a group of codecs equal to the number of available time-slots (e.g. 24 or 30) uses fixed channel times on a highway. Several such highways are concentrated onto a second highway by means of a time switch. System X uses this method[10]. A concentrator connects up to 2048 customers' lines to eight 32-channel PCM systems (i.e. a concentration of 8 to 1).

Since a concentrator is connected to the route switch by a PCM highway, it may be located at a distance from the main exchange. The concentrator can be controlled by the central processor in the main exchange by means of signals sent over the PCM link (e.g. in time-slot 16 of a 30-channel system). If the PCM link between a remote concentrator unit and the main exchange fails, customers on the concentrator lose all service. Duplicate PCM links are therefore often provided.

The control functions of the concentrator may be enhanced to enable it to connect calls between its own customers (but not with others) if the PCM link fails. Facilities must be added to receive and analyze address signals, generate tones and make cross-switch connections between customers' lines. The unit is then known as a *remote switching unit*.

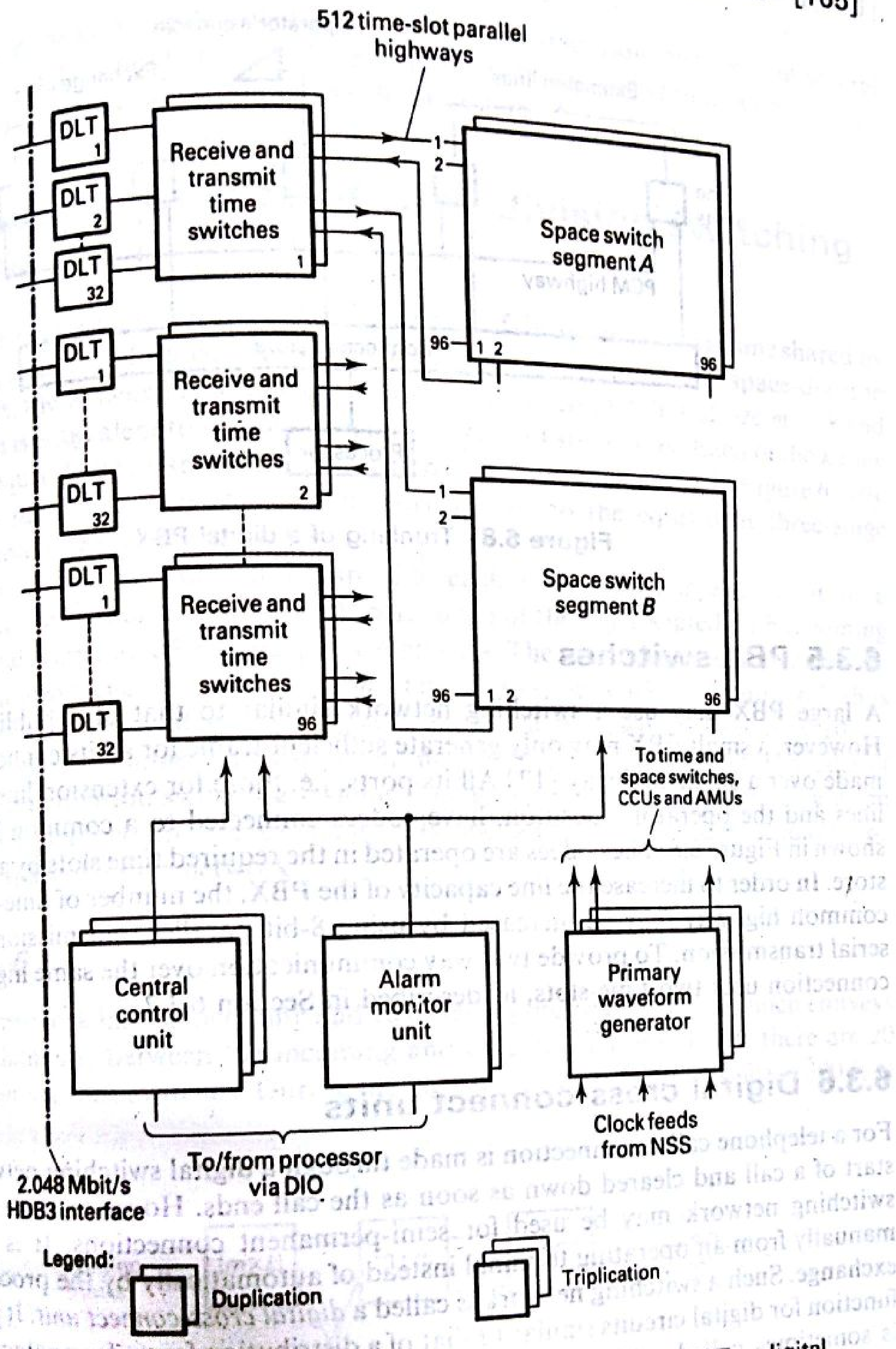


Figure 6.7 Mark 1 digital switching subsystem of System X. DLT = digital line terminating unit.

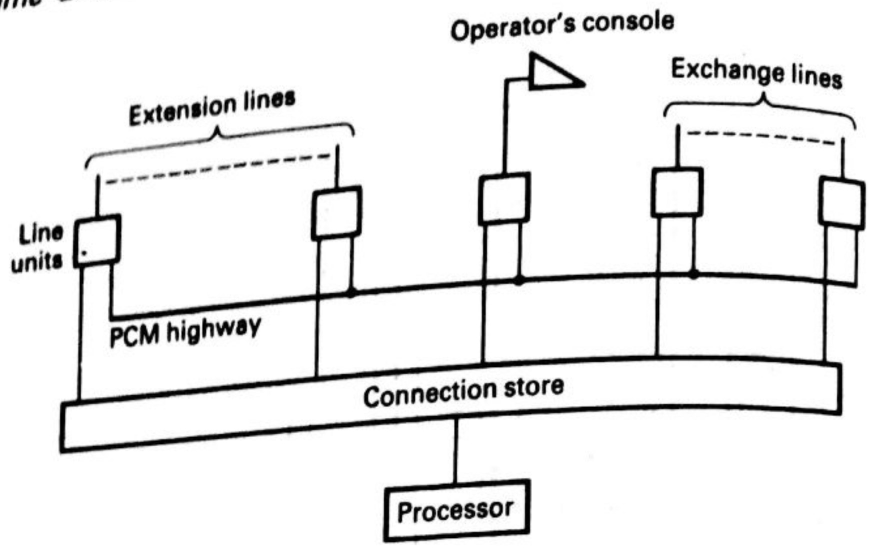


Figure 6.8 Trunking of a digital PBX.

6.3.5 PBX switches

A large PBX may use a switching network similar to that of a public exchange. However, a small PBX may only generate sufficient traffic for all its connections to be made over a single highway.[12] All its ports, i.e. those for extension lines, exchange lines and the operator's position, have codecs connected to a common highway, as shown in Figure 6.8. The codecs are operated in the required time slots by a connection store. In order to increase the line capacity of the PBX, the number of time-slots on the common highway may be increased by using 8-bit parallel transmission instead of serial transmission. To provide two-way communication over the same highway, each connection uses two time-slots, as described in Section 6.3.2.

6.3.6 Digital cross-connect units

For a telephone call, a connection is made through a digital switching network at the start of a call and cleared down as soon as the call ends. However, a similar digital switching network may be used for semi-permanent connections. It is controlled manually from an operating terminal instead of automatically by the processor of an exchange. Such a switching network is called a *digital cross-connect unit*. It performs a function for digital circuits similar to that of a distribution frame for analog circuits. It is sometimes called a 'slow switch', in contrast to a 'fast switch' used to connect telephone calls and the connections made by a digital cross-connect unit are sometimes called 'nailed-up time-slots'.

Two functions that can be performed by digital cross-connect units are *grooming* and *consolidation*. In grooming, 64 kbit/s channels on a common PCM bearer are separated for routing to different destinations. For example, a line from a customer's PBX may carry a mixture of PCM channels, some to the public exchange and some to other PBXs in the customer's private network. In consolidation, channels on several PCM bearers that are not fully loaded are combined onto a smaller number of bearers, thereby improving the utilization of the PCM systems.

6.4 Grades of service of time-division switching networks

In the S–T–S network of Figure 6.4 each crosspoint of the space switch is time shared by n channels. It is therefore equivalent to n separate crosspoints in a space-division switch. Thus, the A switch is equivalent to n space-division switches of size $m \times k$ and the C switch is equivalent to n space-division switches of size $k \times m$. Each of the k time switches is equivalent to a space-division switch of size $n \times n$, as shown in Figure 6.3(b). The S–T–S network of Figure 6.4 thus corresponds to the equivalent three-stage space-division network of Figure 6.9.

In the T–S–T network of Figure 6.5, each time switch is equivalent to a space-division switch of size $n \times n$ and there are m of them associated with incoming highways and m associated with outgoing highways. The space switch is equivalent to n space-division switches of size $m \times m$. The T–S–T network of Figure 6.5 thus corresponds to the equivalent three-stage space-division network of Figure 6.10.

As a result, it is unnecessary to invent new traffic theory to determine grades of service for time-division switching systems. The loss probability obtained for a given traffic offered to a time-division switching network can be determined by studying the equivalent space-division network.

Example 6.1

An S–T–S network has 16 incoming and 16 outgoing highways, each of which conveys 24 PCM channels. Between the incoming and outgoing space switches there are 20 links containing time switches. During the busy hour, the network is offered 300 E of

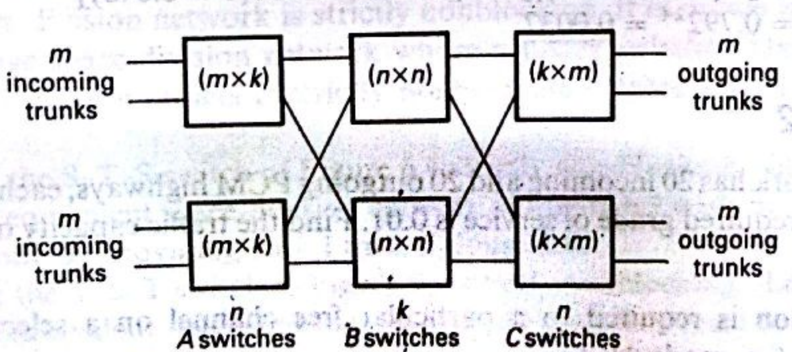


Figure 6.9 Space-division equivalent of S–T–S switch: m = no. of PCM highways, n = no. of time-slots, k = no. of time-switch links.

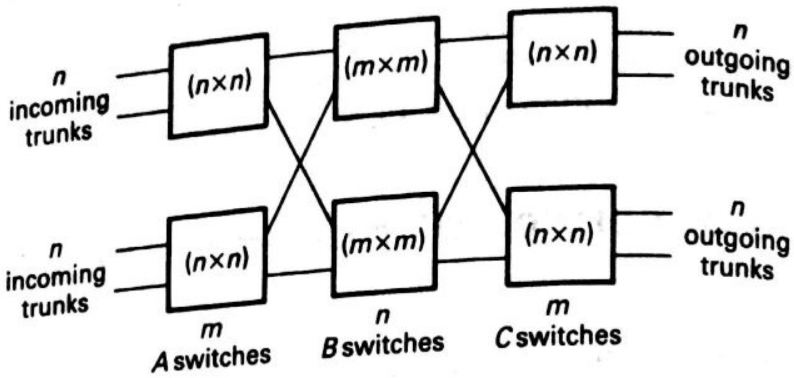


Figure 6.10 Space-division equivalent of T–S–T switch. m = no. of PCM highways, n = no. of time-slots.

traffic and it can be assumed that this is evenly distributed over the outgoing channels. Estimate the grade of service obtained if:

1. Connection is required to a particular free channel on a selected outgoing highway (i.e. mode 1).
2. Connection is required to a particular outgoing highway, but any free channel on it may be used (i.e. mode 2).

1. For the equivalent space-division network shown in Figure 6.9: $m = 16$, $n = 24$, $k = 20$. The occupancy of a link is $b = 300/(24 \times 20) = 0.625$ E.
From equation (5.26):

$$B_1 = [1 - (1 - b)^2]^k = [1 - (1 - 0.625)^2]^{20}$$
$$= 0.859^{20} = \underline{0.048}$$

2. The occupancy of a highway is

$$c = 300/(24 \times 16) = 0.781 \text{ E}$$

From equation (5.27):

$$B_2 = [B_1 + c(1 - B_1)]^n = [0.048 + 0.781(1 - 0.048)]^{24}$$
$$= 0.792^{24} = \underline{0.0037}$$

Example 6.2

A T–S–T network has 20 incoming and 20 outgoing PCM highways, each conveying 30 channels. The required grade of service is 0.01. Find the traffic capacity of the network if:

1. Connection is required to a particular free channel on a selected outgoing highway (i.e. mode 1).
2. Connection is required to a particular outgoing highway, but any free channel on it may be used (i.e. mode 2).

1. For the equivalent space-division network shown in Figure 6.10: $m = 20, n = 30$.
Let the occupancies of the mn links and trunks be b.
From equation (5.26):

$$B^1 = [1 - (1 - b)^2]^n = [1 - (1 - b)^2]^{30} = 0.01$$
$$\therefore \quad 1 - (1 - b)^2 = 0.01^{0.0333} = 0.858$$
$$(1 - b)^2 = 0.142 \text{ and } b = 0.623.$$

$\therefore$ Total traffic capacity of network is
$0.623 \times 20 \times 30 = \underline{374 \text{ E}}$

2. All the channels on a route are provided by the same outgoing time switch; therefore all the trunks on a route are connected to the same C switch in the equivalent space-division network.

The probability of blocking for a connection to this C switch is B_1.
The probability that all trunks outgoing from the C switch are busy is approximately b^n.
$\therefore$ The probability that connection can be made to a free outgoing trunk is $(1-B_1)(1-b^n)$ and the loss probability is $B_2 = 1 - (1-B_1)(1-b^n)$.

However, if n is large, b^n is very small (e.g. if $b = 0.623$, $b^{30} = 6.8 \times 10^{-7}$).

$$\therefore \quad B_2 \doteq B_1$$

Thus, approximately the same loss probability is obtained in either mode and the traffic capacity of the network for $B_2 = 0.01$ is again 374 E.

6.5 Nonblocking networks

Time-division switching networks often have large values of connectivity and are therefore quasi-nonblocking. For example, the network described in Section 6.3.3 has $n = 1024$ and it can carry up to 0.95 E per time-slot.[9] Time-division networks can also be strictly nonblocking.

A time-division switching network will be nonblocking in the strict sense if its equivalent space-division network is strictly nonblocking. It is shown in Section 5.10 that a three-stage space-division network whose primary switches have n inlets and tertiary switches have n outlets is strictly nonblocking if there are $2n-1$ secondary switches.

To make the S–T–S switch of Figure 6.4 strictly non-blocking, the number of B switches in the equivalent space-division network of Figure 6.9 must be at least $2m-1$. This can be done by providing $2m-1$ time-shifting links.

To make the T–S–T switch of Figure 6.5 strictly nonblocking, the number of B switches in the equivalent space-division network of Figure 6.10 must be at least $2n-1$. This can be done by operating the space switch at a higher bit rate than the external highways in order to provide more time-slots (e.g. by doubling the speed to provide 64 time-slots instead of 32).

In either case, since switches are time shared, the cost penalty involved in making a network nonblocking is much less than for a space-division network. Consequently, nonblocking networks are used in some commercial digital switching systems.

6.6 Synchronization

6.6.1 Frame alignment

For correct operation of a time-division switching network the PCM frames on all the incoming highways must be exactly aligned. However, since incoming PCM junctions come from different places, their signals are subjected to different delays. Thus, even if all exchange clock-pulse generators are in perfect synchronism, there will be time differences between the starting instants of different PCM frames entering a digital exchange.

To solve this problem, the line-terminating unit of a PCM junction stores the incoming digits in a *frame-alignment buffer*, as shown in Figure 6.11. Digits are read into this buffer at the rate, f_a, of the incoming line, beginning at the start of each frame. They are then read out at the rate, f_b, of the exchange clock, beginning at the start of the PCM frame of the exchange. To cater for the maximum amount of misalignment between a digital line system and the exchange, the aligner must have a buffer capacity of at least one frame (e.g. 256 bits for a 2 Mbit/s PCM system). This introduces delay additional to that caused by time switching.

A frame-alignment buffer caters perfectly for a constant misalignment. The fill of the buffer is constant and its level depends on the phase difference between the incoming line system and the exchange. It will also cope with a misalignment that changes slowly between limits (e.g. due to temperature changes in cables). However, if the exchanges at the two ends of a line have slightly different clock frequencies, the contents of the buffer will change until it either overflows or empties. If the buffer overflows, its contents are erased so that it can start refilling. If the buffer empties completely, the contents of the previous frame are repeated to refill it. In either case, a complete frame is in error. This is known as a *frame slip*. Of course, slips can also arise

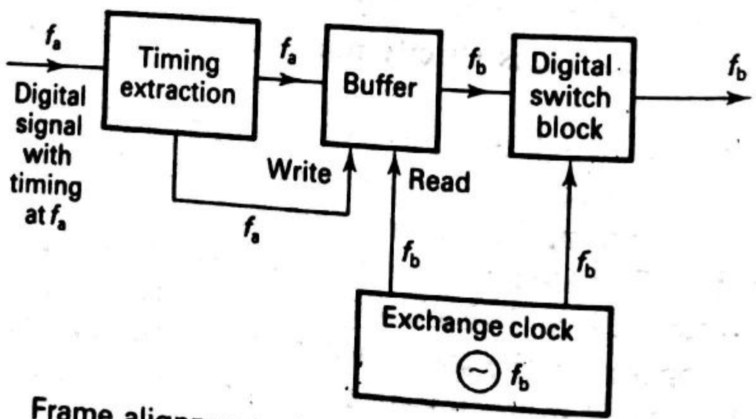


Figure 6.11 Frame alignment of PCM signals entering a digital exchange.

from malfunctions in transmission or switching systems. These are called *uncontrolled slips*, whereas a slip made deliberately to regain frame alignment is called a *controlled slip*.

A digital network may be plesiochronous (i.e. each exchange has an independent clock) or it may be synchronous (i.e. all exchange clocks are controlled by a single master clock). If plesiochronous working is employed, with crystal-controlled clocks having a frequency accuracy of 1 part in 10^7, about 68 frame slips per day will occur in an exchange.[5] For a connection of seven inter-exchange links in tandem, this would cause about 20 slips per hour. For telephony, only about one slip in 25 results in an audible click,[13] so this is tolerable. However, regular slips would be much more serious for data transmission. Either frequent adjustment of clock frequencies would be needed or clocks of atomic standard would be required in all exchanges. Consequently, all the exchanges in an integrated digital network are usually synchronized by a common master clock, as described below.

6.6.2 Synchronization networks

In a synchronous digital network just one or two atomic reference clocks control the frequencies of the clocks of all the exchanges in the network.[14,15] This is sometimes called *despotic control*. For this purpose, a synchronizing network is added to the PSTN in order to link the exchange clocks to the national reference standard. Under normal conditions the network will be free of slip, whereas a plesiochronous network will always experience some slips.

The local clock in each exchange is provided by a crystal oscillator whose frequency can be adjusted by a control voltage. This control voltage is derived from the incoming digit stream on a synchronizing link, which is used to determine whether the exchange clock rate should be increased, decreased or left unchanged. Adjustments are made periodically, as a single quantum increase or decrease. This ensures that exchanges maintain the same long-term average frequency, although short-term deviations may occur. This is known as *mesochronous working*.

Synchronizing links may be unilateral or bilateral. In the first case, there is a 'master–slave' relationship; the clock frequency of the exchange at one end of the link is controlled solely by the exchange at the other end. In the second case, there is a mutual relationship; each exchange influences the frequency of the other. The principles of these methods are shown in Figure 6.12.

A unilateral sync system is shown in Figure 6.12(a). Exchange A is the 'master' and exchange B is the 'slave'. Exchange B determines the phase difference between its own clock and that of exchange A by the fill of the aligner buffer on the incoming link. A change in phase causes a step increase or decrease in clock frequency lasting for a few milliseconds. If there is more than one sync link into exchange B, its correction is based on a majority decision.

In a single-ended bilateral sync link, as shown in Figure 6.12(b), the above decision process is made at each end of the link. As a result, both exchange clocks achieve the same average frequency. In a mesh of such sync nodes, the exchanges would

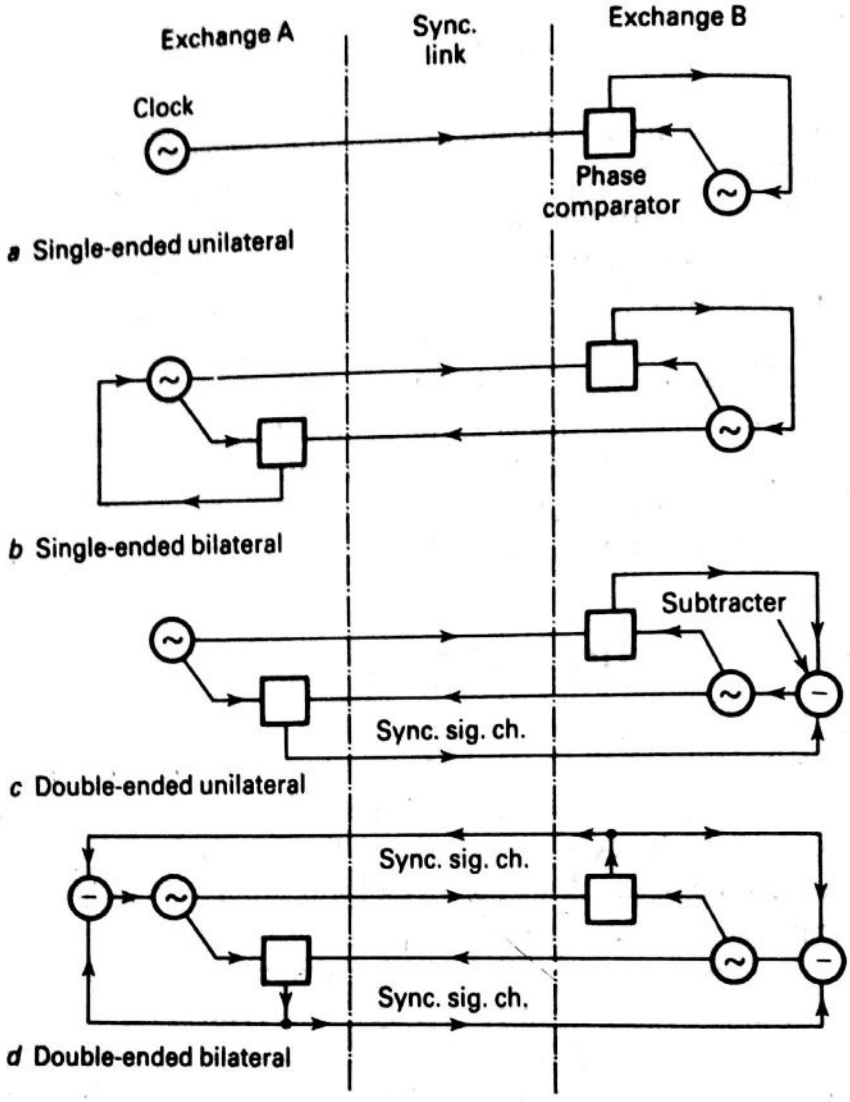


Figure 6.12 Exchange synchronization systems. (a) Single-ended unilateral system. (b) Single-ended bilateral system. (c) Double-ended unilateral system. (d) Double-ended bilateral system.

mutually agree on a common frequency without being controlled by an overall master clock.

A disadvantage of the single-ended unilateral and bilateral sync systems is that the phase comparators are unable to distinguish between phase changes due to frequency drift and those due to changes in propagation time (e.g. caused by cable temperature changes). The former necessitates frequency adjustment, but the latter does not. This disadvantage of the single-ended unilateral and bilateral sync systems is overcome by the double-ended systems shown in Figure 6.12(c) and (d). These eliminate the influence of propagation-delay variations by subtracting the change in phase determined at one end of the link from that determined at the other end.

Let the phase error detected at exchange A be $\delta(\phi_A - \phi_B) + \delta\phi_T$, where $\delta(\phi_A - \phi_B)$ is the phase change due to discrepancy between the clocks and $\delta\phi_T$ is that due to a

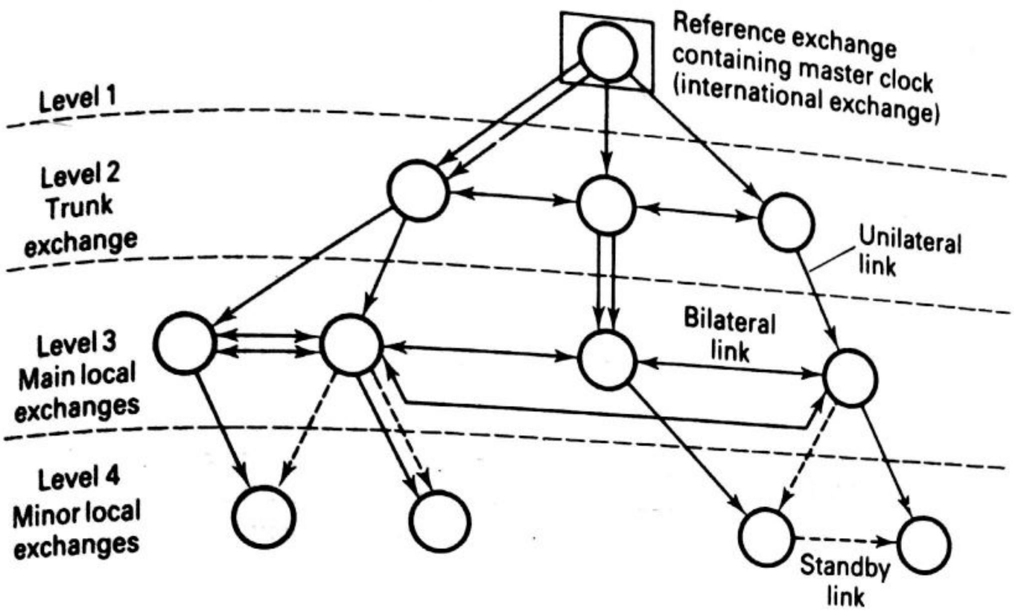


Figure 6.13 Synchronization hierarchy of an integrated digital network.

change in propagation time. Then the phase change detected at exchange B is $\delta(\phi_B - \phi_A) + \delta\phi_T$. Since $\delta(\phi_B - \phi_A) = -\delta(\phi_B - \phi_B)$, the difference between the two measurements is $2\delta(\phi_A - \phi_B)$ and $\delta\phi_T$ is eliminated.

A signalling channel is required to carry the result of the phase comparison to the other end of the link in order to make the subtraction. For a unilateral link, as shown in Figure 6.12(c), this channel is needed in only one direction. For a bilateral link, a signalling channel is needed in each direction, as shown in Figure 6.12(d). If the network uses 2 Mbit/s PCM systems, this signalling information can be carried in the spare capacity of time-slot 0 of the 32 time-slot frame.[5]

A synchronizing network for an IDN is shown in Figure 6.13. Since this auxiliary network must link all exchanges in the IDN, the sync network has the same nodes and the same hierarchical structure as the parent PSTN. The sync links are provided by PCM systems that carry normal traffic between the exchanges.

Frequency control is exerted downwards from the national reference standard by unilateral links from each exchange to those in the next lower level. However, bilateral links are used between exchanges in the same level of the hierarchy. Thus, if there is an incomplete network, or failure of the sync signal from the master source, exchanges at that level mutually determine their own clock frequency and synchronize lower-level exchanges to it.

Problems

1. **(a)** Sketch an S–T–S network to connect *m* incoming highways to *m* outgoing highways, each carrying *n* PCM channels and having *k* time-switch links. Explain briefly how it works.

(b) An S–T–S network has ten incoming highways, ten outgoing highways and ten time-switch links. The highways convey 32 PCM time-slots. The average occupancy of incoming PCM channels is 0.7 E.

(i) Derive an equivalent space-division network.
(ii) Estimate the blocking probability.
(iii) Estimate the grade of service when an incoming call must be connected to a selected outgoing highway but may use any free channel on it.

2. (a) Sketch a T–S–T network to connect m incoming lines to m outgoing lines, each carrying n PCM channels. Explain briefly how it operates.

(b) A T–S–T network has ten incoming highways and ten outgoing highways, each carrying 32 PCM channels. The average occupancy of the incoming channels is 0.6 E.

(i) Derive an equivalent space-division network.
(ii) Estimate the blocking probability.
(iii) Estimate the grade of service when an incoming call must be connected to a selected outgoing highway but may use any free channel on it.

3. Redesign the S–T–S network of question 1(b) and the T–S–T network of question 2(b) to be strictly nonblocking.

4. A T–S–T and S–T–S switch both have 32 incoming and outgoing highways, each having 32 PCM channels. The S–T–S network has 32 time-switch links, so both networks give the same blocking probability.

Compare the numbers of crosspoints and bytes of storage required for these two networks.

5. Since bits of storage are very much cheaper than crosspoints, an approximate measure that has been proposed for comparing the relative costs of networks is the 'complexity', defined as

Complexity $= N_c + N_b/100$, where
N_c = number of space-stage crosspoints
N_b = number of bits of memory

Compare the complexities of the S–T–S network in question 1(b) and the T–S–T network in question 2(b).

6. A two-stage digital switching network is to make connections between m incoming PCM highways and m outgoing PCM highways, each having n channels. Each call from an incoming PCM highway is to be connected to a selected outgoing PCM highway but may use any free channel on it.

(a) The network may use either space-time switching or time-space switching. Draw an equivalent space-division network for each. Hence, state which is preferable and explain why.
(b) For each network, determine the grade of service when the network has ten incoming and ten outgoing 30-channel PCM systems, each channel having an average occupancy of 0.4 E.

7. A digital PBX has all its ports (serving extensions, exchange lines and the operator) connected to a common bus, as shown in Figure 6.8. The bus carries 128 time-slots and can be used for making a connection between any two ports.

In the busy hour, the average both-way traffic per extension is 0.1 E and 10% of this constitutes traffic to and from the public exchange. The occupancy of the operator's position is 0.3 E. Since extension users can dial outside calls directly, traffic between extensions and the operator is very small.

If the required grade of service is 0.01, how many extensions can the PBX accommodate and how many lines are needed to the public exchange? (Table 4.1 may be used.)

8. In the digital switching subsystem of System X, as shown in Figure 6.7, the sending and receiving time switches each have a speech store containing 1024 storage locations and so can serve 32 PCM systems (each having 32 time-slots). The complete switching network contains up to 96 time switches, serving up to 3072 PCM systems. The space switch is thus of size 96 × 96 and handles 1024 channels. However, to minimize problems of pulse distribution, the 1024-channel highways are each split into two highways of 512 channels each. For security, the complete network is duplicated.

If the switch is fully equipped and the average occupancy of incoming channels is less than 0.45 E, show that connections to outgoing routes (which may use any free time-slot) have negligible loss probabilty.

9. (a) What are the principal requirements for the interface between a PCM line system and an incoming time switch?

(b) Show the application of this in designing a time switch capable of handling 32 PCM systems.

10. A plesiochronous digital network uses 2.048 Mbit/s line systems whose frames contain 256 bits. The exchange clocks have a frequency stability of 1 part in 10^7.

(a) Prove that a connection of seven inter-exchange links can have about 20 slips per hour.

(b) On average, what proportion of bits are then received in error?

(c) In what percentage of 3-minute telephone calls on such connections will clicks be audible?

References

[1] Chapuis, R.J. and Joel, A.E. (1990), *Electronics, Computers and Telephone Switching: a book of technological history*, North-Holland, Amsterdam.

[2] Joel, A.E. (ed.) (1982), *Electronic Switching: digital central office systems of the world*, IEEE Press, New York.

[3] McDonald, J.C. (ed.) (1983), *Fundamentals of Digital Switching*, Plenum Press, New York.

[4] Ronayne, J. (1986), *Introduction to Digital Switching*, Pitman, London.

[5] Redmill, F.J. and Valdar, A.R. (1990), *SPC Digital Telephone Exchanges*, Peter Peregrinus, Stevenage.

[6] Tippler, J. (1979), 'Architecture of System X: Part I', *PO Elect. Engrs. Jour.*, **72**, 138–41.

[7] 'The No.5 ESS switching system', (1985), Special Issue of *ATT Bell Labs, Tech. Jour.*, **70**, Part 2.

[8] 'The 4ESS switch', (1977), Special Issue of *Bell Syst. Tech. Jour.*, **56**, Part 7.

[9] Risbridger, J.N.A. (1979), 'System X subsystems, Part 1: the digital switching subsystem', *PO Elec. Engrs. Jour.*, **73**, 19–26.

[10] Maddern, T.S. and Ozdmar, M. (1989), 'Advances in digital circuit switching architecture', *Proc. 2nd IEE National Conf. on Telecommunications*, pp. 303–8.

[11] Foxton, M.C. (1980), 'System X subsystems, Part 5: the subscribers' switching subsystem', *PO Elec. Engrs. Jour.*, **73**, 216–22.

[12] Rayfield, D.A.T. and Gracie, S.J. (1979), 'The Plessey PDX: a new digital PABX', *Int. Conf. on Private Electronic Switching Systems*, London, IEE Conf. Pub. no. 163, pp. 211–14.

[13] McLintock, R.W. (1991), 'Transmission performance', Chapter 4 in Flood, J.E. and Cochrane, P. (eds), *Transmission Systems*, Peter Peregrinus, Stevenage.

[14] Boulter, R.A. and Bunn, W. (1977), 'Network synchronisation', *PO Elect. Engrs. Jour.*, **70**, 21–8.

[15] Abate, J., Cooper, C., Pạn, J. and Shapiro, I. (1979), 'Synchronization considerations for the switched digital network', *Int. Commun. Conf. Rec.*, June.