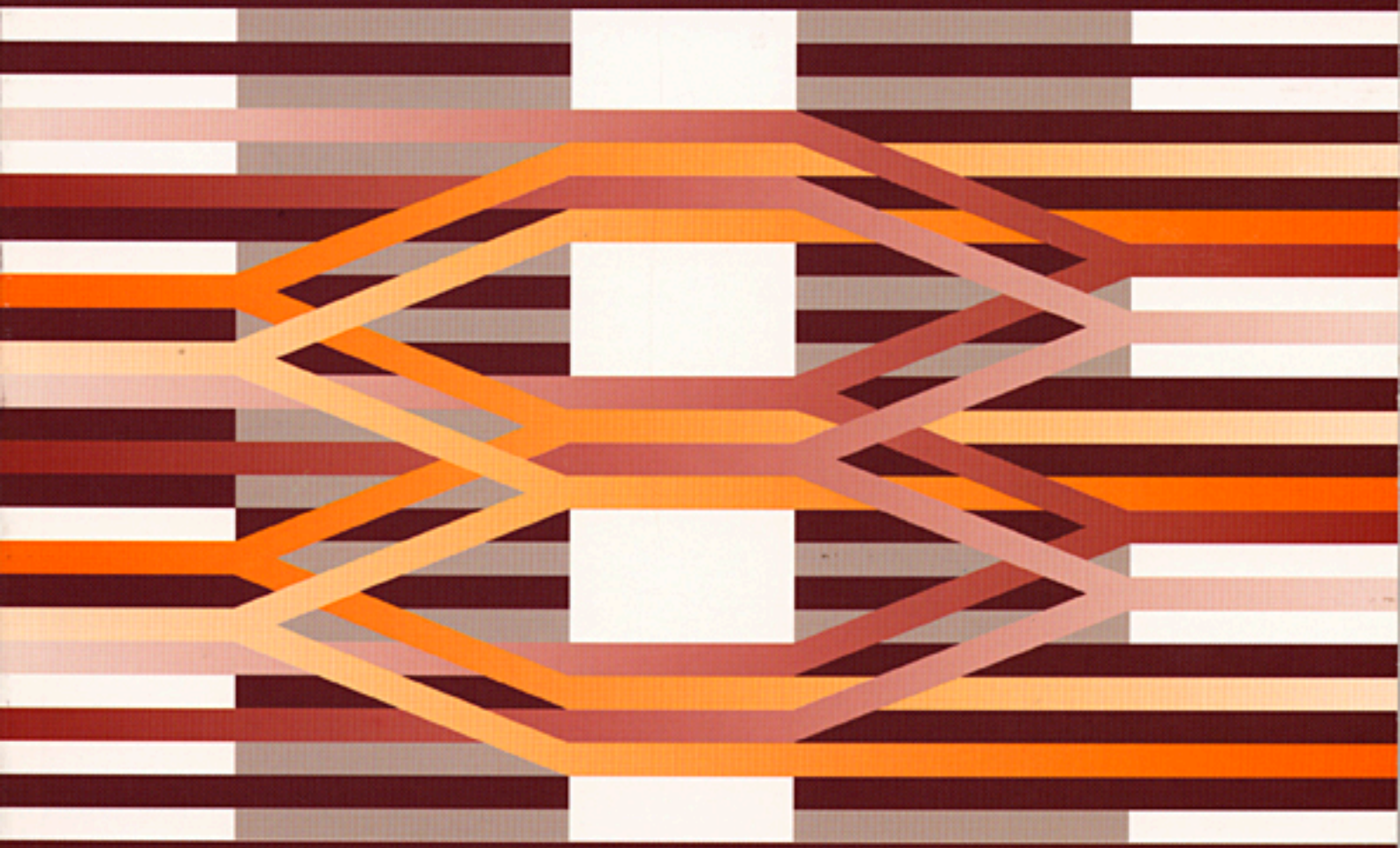
Telektronikk 2.96

Switching



Contents

Feature

Telenor

Telektronikk

Volume 92 No. 2 - 1996
ISSN 0085-7130

Editor:
Ola Espvik
Tel. + 47 63 84 88 83

Status section editor:
Endre Skolt
Tel. + 47 63 84 87 11

Editorial assistant:
Gunhild Luke
Tel. + 47 63 84 86 52

Editorial office:
Telektronikk
Telenor AS, Telenor Research & Development
P.O. Box 83
N-2007 Kjeller, Norway
e-mail: telektronikk@fou.telenor.no

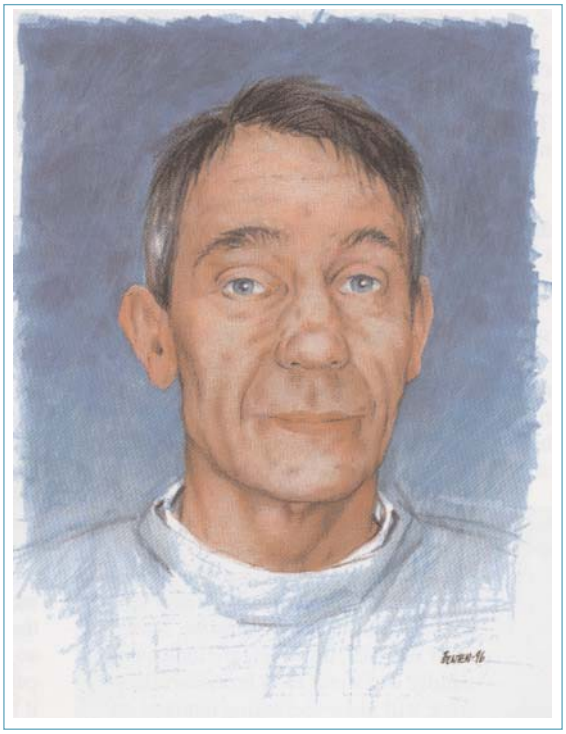
Editorial board:
Ole P Håkonsen, Senior Executive Vice President
Oddvar Hesjedal, Vice President, Research & Development
Bjørn Løken, Director

Graphic design:
Design Consult AS

Layout and illustrations:
Gunhild Luke, Britt Kjus, Åse Aardal
Telenor Research & Development

Guest editorial

BY KARL KILVIK



This issue of *Telektronikk* is intended to provide some basic knowledge about switching theory as well as some special topics concerning the use of switching. It is obvious that one issue cannot give all the information, neither on switching theory nor on the large number of telecommunication systems.

First, there is a logical basic question: What is a switch, or what is switching? The answer is that it can be everything from a light switch that everybody knows to a sophisticated switch in a modern tele/data communication system. The fact is that the switch in some form or other is used in a large number of equipment and is often hidden within a black box or within a piece of software. In fact, it would take up too much space to list all places where the switch is a key component. The fact that the function of the switch can also be represented by a piece of software, indicates that the view of what a switch is should not be restricted to a physical construction of some kind. So, in trying to establish a switching theory it is important not to restrict the definition of a switch to a specific technology or use.

Another place where switching theory ought to be useful is in the planning of the whole communication system. This should intuitively be of great value since the whole system (LAN, MAN, WAN, etc.) can be viewed as a distributed switching network. The technological advance moving from the electromechanical to today's beginning of photonic switching systems is expected to give a cheaper and more versatile system with a great increase in services. It is often stated that this will need a system with more intelligence. However, this need not be true, as the evolution could also be towards a more passive network and terminal equipment with more intelligence. A telecommunication network can be said to consist of two parts: the transmission network and the switching nodes. It is the first part which attracts most attention among people when a new system is announced. Names like "electronic motorways", broadband services etc., are often used when the new technology is explained to people. This is of course correct, but at the same time this may give the wrong impression of the system since the lack of the other main part of the system, i.e. the switching nodes, has to be able to "use" the transmission capacity. This can degrade the whole system even with an "excellent" transmission network.

The reliability of a telecommunication system is often taken for granted. Everybody seems to expect that the telephone operates every time it is used. The reliability problem of telecommunication systems is expected to be dealt with in a later issue and will not be mentioned further here. In data communication the switching aspect is often hidden in a complicated use of protocols, etc. This is perhaps a serious drawback since it is obvious that the switching theory can be used when trying to get many protocols to cooperate.

Switching in telecommunications

BY KARL KILVIK

1 Introduction

Historically, the modern telecommunication networks started with the invention of the telegraph (in the 1830s). This made it possible to have rapid communication between far away places.

The next invention – the telephone – also made it possible to transport voice and this gradually resulted in the modern telecommunication networks giving a very large contribution to the society.

A modern communication network can be divided into two main parts, the transmission part and the switching part. The transmission part transports the information while the switching part gets the right connections between the transmitters and the receivers. This classification can also be used in systems where switching and transmission are more or less distributed. A modern telecommunication network must be able to transmit voice, numbers, characters and pictures. These various types of information have resulted in different networks, each specially adapted to the specific information type. However, the trend now is to transform the various types of information to the same representation, transport this through the network and transform the information back to the users' representation at destination. This common representation is now digital, and the digitalisation process is going on world-wide. To get an impression of the demands for bandwidth from the many different services of a modern telecommunication network, one can transform the data to digital representation (in bit/sec) as in Table 1.1 which shows some services and their bitrate. The most common service so far is voice communication, and as seen it needs a rather moderate bitrate. Broadband services, however, are increasing rapidly and a common broadband network will place increasing demand on the network both in the transmission and the switching part. The possibility of offering greater bandwidth to the users has increased in time with the technological development both in transmission and in switching. At present, it can be said that the development of high capacity optical (photonic) transmission systems has developed to a commercial state while the switching systems have not yet advanced to a practical level.

Table 1.1 Some services and their bitrate

Service	Bitrate bit/s
Telex	0.05 k
Teledata – telefax	0.3 k – 4.8 k
Voice coded	10 k – 19 k
Voice uncoded	64 k
Picture phone (coded)	64 k – 384 k
High bitrate data communication	64 k – 140 M
High bitrate telefax	64 k →
Computer communication	64 k – 140 M
Video library	2 M – 140 M
TV	34 M – 140 M
HDTV	140 M →
ATM	600 M

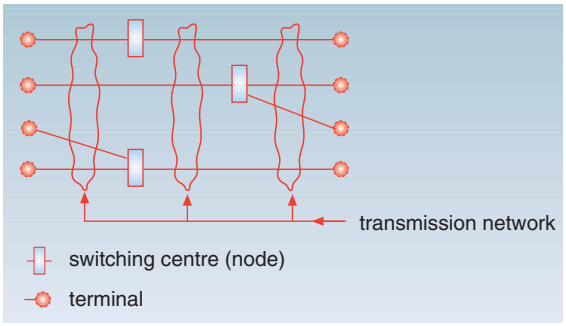


Figure 2.1 Principal structure of a telecommunication network

First, (Section 2) we will give a short overview of the principal structures and functions of the basic networks used in telecommunication. The name telecommunication network in general also includes networks where the nodes are computers and not only exchanges, as they can be viewed as specialised computers. A short notion about a mathematical description of different structures will also be given. In Section 3 a switching element and a switch is given a verbal and mathematical definition together with the basic ideas of time and space switching. Sections 4 and 5 include the basic switching parameters of a switch and various types of switching elements. Since it is expected that photonic technology is the future technology nearest in time, a more extended description has been given of photonic components than of the electronic counterparts now dominating. Section 6 gives a short classification of switches and switching networks. Section 7 gives a compressed switching theory. The mathematical description has been made brief so the reader does not need too much of mathematical background. Sections 8 and 9 are devoted to the construction of space and time switching networks, respectively. With the increasing broadband services it seems reasonable to use a section (Section 10) to deal with the special properties of these services, among them the new principle of transferring data, namely ATM. Photonic technology is expected to be the fundament for the future telecommunication system and for this reason one section (Section 11) is devoted to photonic switching. A very important function is the control of the switches. This control together with the switching network itself determines the properties of the switch. A poor control function also gives a poor total function of the switch even if the switching network itself is "excellent". However, the theme "control of switching network" is so large that a single issue is not sufficient to deal with both of these themes and it is therefore omitted.

2 The principal structure of telecommunication networks

2.1 General

The word telecommunication will be used in the sense of transmission of data (information) through a physical system called telecommunication network, or network for short. This network consists of three main components which have to cooperate in the right way to solve the different tasks:

- The transmission network
- The switching nodes
- The terminals.

The transmission part takes care of the transport of information (data) between the terminals and between the nodes. The switching part (the nodes) is responsible for the distribution to the right terminals. The terminals generate and receive the information.

Figure 2.1 shows schematically how a telecommunication network can be organised.

A modern network can be characterised by:

- A simple equipment for the user
- One channel per user to the first switching node

- Common equipment is shared by many users
- Relatively complex equipment to take care of the necessary exchange of data between the nodes.

There is, however, a trend to have more advanced equipment by the user and at the same time more services in the network. This is of course a result of the increasing number of different services.

Seen from the network, a first definition of switching can be expressed as:

> "Switching in telecommunication network sense is the function that distributes the data from a source to the right destination."

This definition must also incorporate the control function of the switches and not only the plain switch (for example in the form of relay contact) so that

> "Switching is the physical equipment and processes that make it possible to route data from a source to the right receiver."

The basic principles of switching are:

- Space switching, SSW
- Time switching, TSW
- Wavelength or frequency switching, WSW

cf. the basic principles of multiplexing in transmission networks. The first two principles are in common use and the last one is due to the development of the optical technology.

Other more functional classifications are also in use, reflecting the actual use of switching, e.g.:

- Circuit switching
  - The circuit (connection) is established for the complete duration of the connection.
- Packet switching
  - Information is divided into packets.
  - Each packet contains additional data as address, error control, etc.
- ATM-switching (see later chapter)
  - Information is divided into cells (packets), all of the same length.

Other names are cell-switching, digital switching, etc. All these "functional names" above are closely linked to the way of representing and transferring data. However, they are not new principles of switching, they all use the same basic principles: space and time switching (and perhaps wavelength switching in the future).

2.2 Structure of networks

The structures of networks can be classified into several groups (Figure 2.2):

- Mesh network
  - A node/terminal has connections to all the other nodes.
- Star network
  - The nodes/terminals have connections to a common node/terminal.
- Tree network
  - The nodes/terminals are connected like a tree.
- Bus network
  - The nodes/terminals are connected to a common bus or a system of buses.
- Ring network
  - The nodes/terminals are connected to a common ringbus.

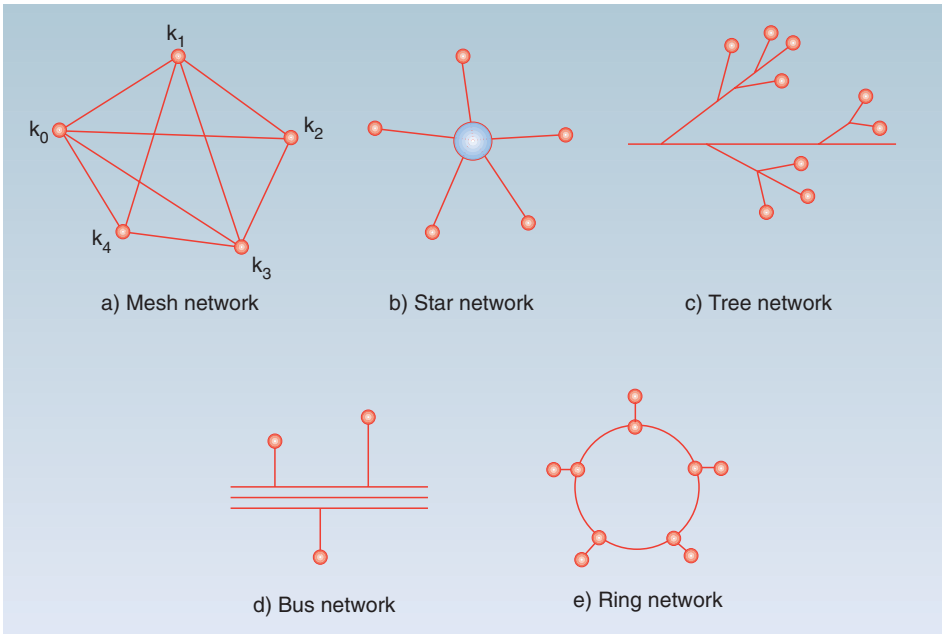


Figure 2.2 Basic structures of a network

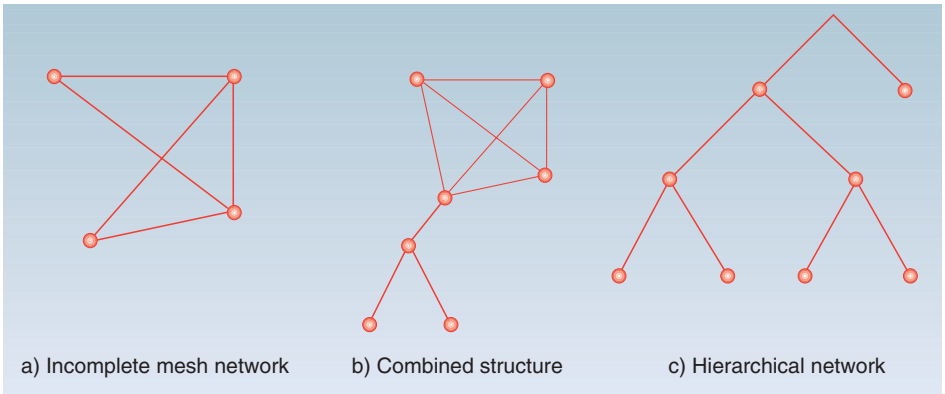


Figure 2.3 Derived network structures

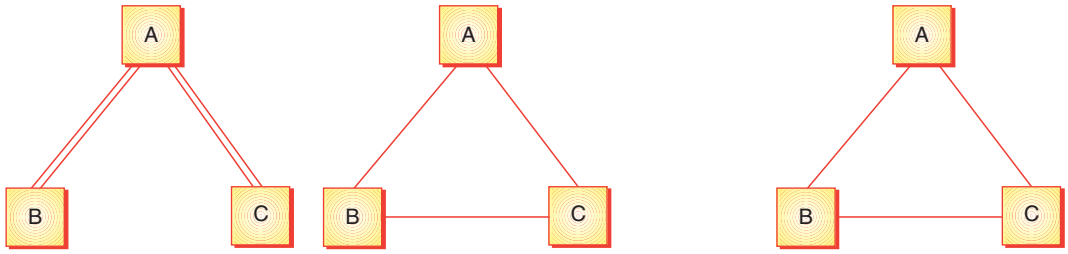


Figure 2.4 Some examples of physical and logical structures

The structures can be represented by graphs where the nodes/terminals are the vertices and the transmission lines are branches between the vertices. Mathematically, a graph $G = (V,T)$ is described as a set of vertices $V = \{v_0, v_1, v_2, ...\}$ and a set of branches defined by the ordered pairs $T = \{(v_i, v_j)\}$ of the vertices. Vertex v_i is called the initial vertex and vertex v_j the final vertex. The graph can be a directed or undirected, reflecting transmission in one or both directions. The direction is defined from the first vertex to the second vertex in the pair. $(v_i, v_j) = (v_j, v_i)$ gives duplex transmission and an undirectional graph. $(v_i, v_j) \neq (v_j, v_i)$ gives a directed graph and simplex transmission.

A graph can also be used for representation of the switching network, but here it is more common to use the terminals as nodes and the switches as branches (see later).

The basic structures in Figure 2.2 can be used to construct networks by combining the structures above and also use incomplete structures. Examples are shown in Figure 2.3. Most of the telecommunication networks are realised in this way. The networks can be arranged in a hierarchical way as a mixture of tree, mesh and star network. Often, star networks are used at the lowest level and mesh networks at higher levels. Networks named LAN, MAN and WAN also use tree and ring structures.

As illustrated in Figure 2.4, one has to distinguish between physical and logical structure. Physically, the networks can be quite different, but with respect to logical and functional behaviour, they may be similar.

The structures represented by graphs will not give the complete picture of the network. Additional data are needed on:

- The branches, e.g.:
  - The number of lines (channels) and what kind (1-,2-,4-wires, coax, optical fibre, etc.)
  - The behaviour with respect to transmission like bandwidth, crosstalk, loss, etc.
- The nodes, e.g.:
  - Switching principle
  - Technology
  - Amplification
  - Size.

2.3 Functional behaviour of a network

The basic function of a network can be defined as to enable the users to communicate with each other so that the users themselves determine:

- The time when communication has to be established
- The duration of the conversation or the amount of data to be transmitted
- The receiver's address

Most often, connections are 1:1, e.g. only two users are involved in a connection or have permission to use the same bidirectional transmission channel at the same time. As determined by the users, usage of connections varies statistically.

These functions will be much the same for all types of networks in spite of their differences with respect to actual structures and technology. The actual implementation, however, may differ in the various systems. It would be very important to define the functions in such a way that they are invariant with respect both to technology and to the actual structure of the system.

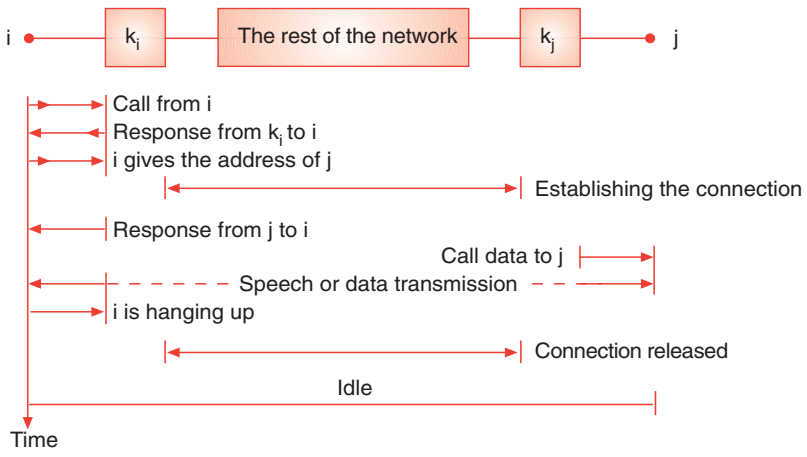


Figure 2.5 Examples of basic functions of a communication network

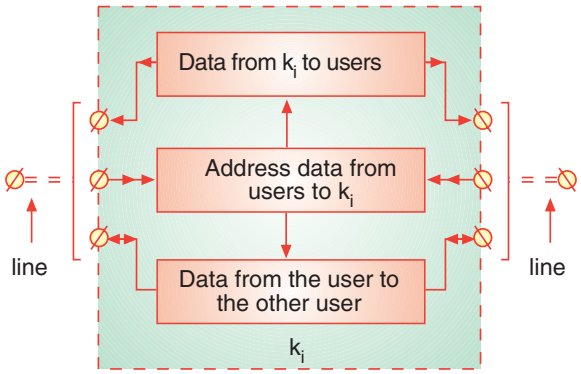


Figure 2.6 Data streams in a node

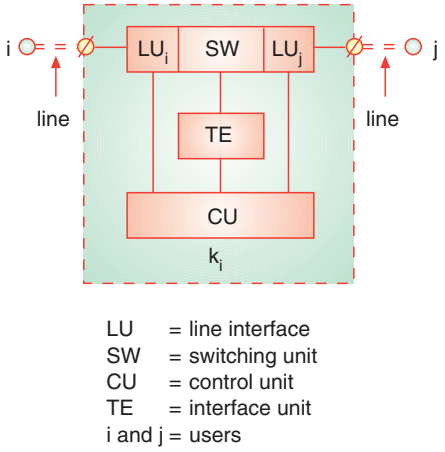


Figure 2.7 Possible functional units in a node k_i

The basic functions of a network containing nodes and transmission lines can be illustrated as in Figure 2.5, where the users *i* and *j* are connected to nodes k_i and k_j, respectively. These nodes are connected together via the rest of the network. When *i* makes a call to *j* we have the following basic phases:

- Detection phase:
  node k_i detects the call from *i*
- Addressing phase:
  k_i receives the address of *j* from *i*
- Connection phase:
  the different nodes find a way through the network and inform *i* that, in this case, *j* is idle
- "Talking" phase:
  the connection between *i* and *j* is established
- Release phase:
  the connection is released.

Each of these phases can be divided into subphases which take care of situations such as: failure, the user interrupt of the calling process, the connection cannot be established, the receiver is busy, etc.

Similar phases and processes are found in all of the nodes.

The information (data) the network is working with can be divided into the following main groups:

- User information, i.e. the information the user wishes to transmit
- Address information generated by the user giving the address of the receiving party
- Internal information generated by the system (nodes) itself to route the user information within the network.

User information may be speech, data, pictures, etc. The address information and the internal generated information are known as signalling data, and "signalling" is the process which processes those data.

2.3.1 Functional behaviour in the nodes

Within a node the data can be divided into three groups (see Figure 2.6):

- Data from user to the node (address data)
- Data from the node to the user
- Data from a user to another user.

Figure 2.7 shows a structure of a node that can be used to execute the functions, these data streams demand. The structure consists of:

- The switching unit, SW
- The control unit, CU
- Interface units, LU and TE.

The switching unit, SW, connects the users to enable data transfer between them and may also connect the various functional units to each other. The line unit, LU, adapts the users' representation of data to the representation used within the rest of the node. TE adapts SW to CU to enable CU to take care of its functions. In modern nodes, CU works like a (specialised) computer.

2.4 Formalism of functions in a network

In principle, all functions are the same regardless of which level the node is placed on. The basic functions will be the same – to establish connections between specified users (or terminals). However, it may be suitable to divide the functions into two main groups:

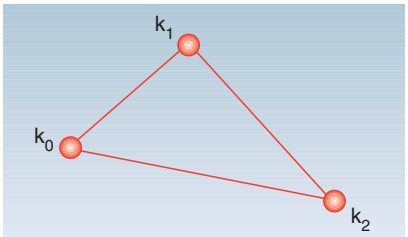
- Functions used when both users are connected to the same node
- Functions used when the node is used as transit, i.e. the node has other nodes as users.

Functions in the first group will be called local functions, and functions in the second group will be called transit functions (this also incorporates signalling functions).

A specification of functions in a network can now be expressed with matrices, giving a more compact way to describe the functional behaviour of the network. The method will be described by a few simple examples.

2.4.1 Functional matrix of a mesh network

The network shown in Figure 2.8a consists of 3 nodes with a connection matrix as given in Figure 2.8b. The following notation is used:



a) Mesh structure

	k_0	k_1	k_2
k_0	1	x	x
k_1	x	1	x
k_2	x	x	1

1 = always connection
x = possible connection

b) Connection matrix

$$F_0 = \begin{bmatrix} f_0 & \alpha_{01} & \alpha_{02} \\ \alpha_{10} & f_1 & \alpha_{12} \\ \alpha_{20} & \alpha_{21} & f_2 \end{bmatrix}$$

$$\begin{aligned}
\alpha_{01} &= f_0 f_{01} f_1 + f_0 f_{02} f_2 f_{21} f_1 \\
\alpha_{02} &= f_0 f_{02} f_2 + f_0 f_{01} f_1 f_{12} f_2 \\
\alpha_{10} &= f_1 f_{10} f_0 + f_1 f_{12} f_2 f_{20} f_0 \\
\alpha_{12} &= f_1 f_{12} f_2 + f_1 f_{10} f_0 f_{02} f_2 \\
\alpha_{20} &= f_2 f_{20} f_0 + f_2 f_{21} f_1 f_{10} f_0 \\
\alpha_{21} &= f_2 f_{21} f_1 + f_2 f_{20} f_0 f_{01} f_1
\end{aligned}$$

$$F_1 = \begin{bmatrix} f_0 & f_0 f_{01} f_1 & f_0 f_{02} f_2 \\ f_1 f_{10} f_0 & f_1 & f_1 f_{12} f_2 \\ f_2 f_{20} f_0 & f_2 f_{21} f_1 & f_2 \end{bmatrix}$$

1) Alternative routing allowed

2) No alternative routing

f_0 => functions in node k_0
f_1 => functions in node k_1
f_2 => functions in node k_2

$f_i + f_j$ => f_i or f_j
$f_i f_j$ => f_i and f_j
f_{ij} = functions to make connection from k_i to k_j

c) Functional matrices

Figure 2.8 Mesh network and associated functional matrices

f_i = internal functions in node i

f_{ij} = functions needed to establish a connection between node i and node j

$f_i + f_j$ = f_i or f_j

$f_i f_j$ = f_i and f_j

As an example, a connection from user i in k_0 to user j in k_1 first uses function f_0 in k_0, then function f_{01} to get a connection to k_1 and at last function f_1 in k_1 so that all the functions $f_0 f_{01} f_1$ are executed. If alternative routing is allowed, the user can "go" through k_2 if there is no idle path from k_0 to k_1. This gives the additional functions $f_0 f_{02} f_2 f_{21} f_1$ and the total function will be:

$$F_{01} = f_0 f_{01} f_1 + f_0 f_{02} f_2 f_{21} f_1 \qquad (2.1)$$

Without alternative routing the only connection allowed is from source node to destination node, which gives:

$$F_{01} = f_0 f_{01} f_1 \qquad (2.2)$$

Applied on the network in Figure 2.8a this gives the (functional) matrices in Figure 2.8c.

The functions must be executed from left to right i.e. $f_0 \rightarrow f_{01} \rightarrow f_1$. The function f_{ij} can be interpreted as the signalling function.

In such a functional matrix with n nodes and no restriction with respect to routing there will be n - 1 possibilities to route out of a node such that:

$$\alpha_{ij} = \sum_{x=0}^{x=n-1} e_{ij}(x)$$

$$= e_{ij}(0) + \ldots + e_{ij}(n-1) \qquad (2.3)$$

where:

$e_{ij}(x)$ = routing through x other nodes between the nodes i and j.

No alternative routing follows from eq. (2.3) with $x = 0$:

$$\begin{aligned} \alpha_{ij} &= f_i f_{ij} f_j \quad \text{for } i \neq j \\ \alpha_{ij} &= f_i \quad \text{for } i = j \end{aligned} \qquad (2.4)$$

The functional matrix can be written:

$$F = \begin{bmatrix} f_0 & \alpha_{01} & \alpha_{02} & .. & \alpha_{0n} \\ \alpha_{10} & f_1 & \alpha_{12} & .. & \alpha_{1n} \\ \alpha_{20} & \alpha_{21} & f_2 & .. & \alpha_{2n} \\ .. & .. & .. & .. & .. \\ \alpha_{n0} & \alpha_{n1} & \alpha_{n2} & .. & f_n \end{bmatrix} \qquad (2.5)$$

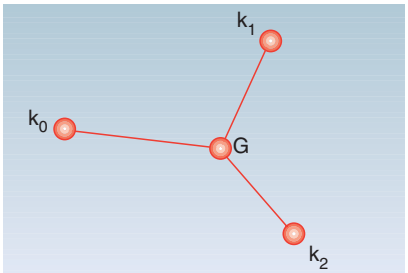
In a network with n nodes the number of ways of routing through x nodes can be written:

$$r_x = [(n-2)! / (n-2-x)!] \qquad (2.6)$$

Example:

n = 5 nodes gives: $x = 0$: $r_0 = 3!/3! = 1$;
$x = 1$: $r_1 = 3!/2! = 3$;
$x = 2$: $r_2 = 3!/1! = 6$;
$x = 3$: $r_3 = 3!/0! = 6$;
Total: 16

n = 6 nodes gives: $x = 0$: $r_0 = 4!/4! = 1$
$x = 1$: $r_1 = 4!/3! = 4$
$x = 2$: $r_2 = 4!/2! = 12$
$x = 3$: $r_3 = 4!/1! = 24$
$x = 4$: $r_4 = 4!/0! = 24$
Total: 65



a) Star network

	k_0	k_1	k_2
k_0	1	0	0
k_1	0	1	0
k_2	0	0	1

1 = always connection
0 = no direct connection

b) Connection matrix

$$F_S(1) = \begin{bmatrix} f_0 & f_0 f_c G_1 f_c f_1 & f_0 f_c G_1 f_c f_2 \\ f_1 f_c G_1 f_c f_0 & f_1 & f_1 f_c G_1 f_c f_2 \\ f_2 f_c G_1 f_c f_0 & f_2 f_c G_1 f_c f_1 & f_2 \end{bmatrix}$$

1) Full autonomy

$$F_0(0) = \begin{bmatrix} g_0 g_c G_2 g_0 & g_0 g_c G_2 g_c g_1 & g_0 g_c G_2 g_c g_2 \\ g_1 g_c G_2 g_c g_0 & g_1 g_c G_2 g_1 & g_1 g_c G_2 g_c g_2 \\ g_2 g_c G_2 g_c g_0 & g_2 g_c G_2 g_c g_1 & g_2 g_c G_2 g_2 \end{bmatrix}$$

2) No autonomy

f_i / g_i = functions in node k_i for full/no autonomy
f_c / g_c = signalling functions for full/no autonomy
G_1 / G_2 = functions in node G for full/no autonomy
f_{ij} $f_i + f_j = f_i$ "or" f_j for all $i \neq j$
$f_i f_j$ = f_i "and" f_j

c) Functional matrices

Figure 2.9 Star network and associated matrices

Following this, alternative routing can be used to increase the traffic capacity of the network or to reduce the necessary number of transmission channels between the nodes.

In a practical situation there is always some restrictions which limit the use of alternative routing. The two most important restrictions are:

- Time delay because of transferring on transmission lines between nodes
- Time delay because of processing and switching within nodes.

In a real network, the number and lengths of transmission lines (channels) can be different in the two directions. The functional matrix for alternative routing may be traffic dependent.

Assume, for all i and j:

$$f_{ii} = f_i; f_i + f_i = f_i; f_i = f_j = f_0$$
$$f_c = f_{ij} = f_{ji} = \text{the signalling function between node } k_i \text{ and } k_j \quad (2.7)$$

Used on the matrix in Figure 2.8c (3 nodes, alternative routing), the matrix can be transformed to:

$$F_0 = f_0\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & f_c & f_c \\ f_c & 0 & f_c \\ f_c & f_c & 0 \end{bmatrix} f_0 + \begin{bmatrix} 0 & f_c & f_c \\ f_c & 0 & f_c \\ f_c & f_c & 0 \end{bmatrix} f_0 f_c f_0\right) \quad (2.8)$$

$$F_0 = f_0 I + f_0 C f_0 + f_0 C f_0 f_c f_0 \quad (2.9)$$

where

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \; C = \begin{bmatrix} 0 & f_c & f_c \\ f_c & 0 & f_c \\ f_c & f_c & 0 \end{bmatrix} \quad (2.10)$$

C = "signalling matrix"

The first component in eq. (2.9) corresponds to "local functions", i.e. the functions which are related to local traffic (originating and terminating in the same node). The second component in eq. (2.9) corresponds to functions related to direct traffic between nodes. The third component is due to alternative routing.

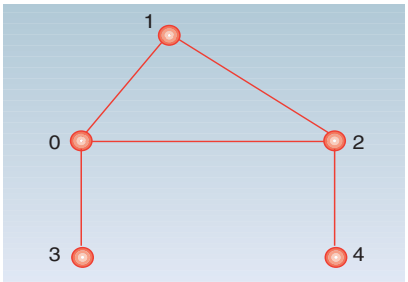
Without alternative routing the matrix in Figure 2.8c (same assumptions as for eq. (2.8)) can be written:

$$F_1 = f_0 \begin{bmatrix} 1 & f_c f_0 & f_c f_0 \\ f_c f_0 & 1 & f_c f_0 \\ f_c f_0 & f_c f_0 & 1 \end{bmatrix} = f_0\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & f_c & f_c \\ f_c & 0 & f_c \\ f_c & f_c & 0 \end{bmatrix} f_0\right) \quad (2.11)$$

$$F_1 = f_0 I + f_0 C f_0 \quad (2.12)$$

The equations above show that the total set of functions in a network are built up of functions in the nodes with additional signalling functions. This signalling function will depend on: the level in the network, services, technology, signalling system, etc.

In existing networks there are many different signalling systems, but the trend now is to standardise on one system (No. 7). The treatment of signalling systems, however, is beyond the scope of this paper.



a) Network

	0	1	2	3	4
0	1	x	x	x	0
1	x	1	x	0	0
2	x	x	1	0	x
3	x	0	0	1	0
4	0	0	x	0	1

1 = always connection
0 = no direct connection
x = possible connection

b) Connection matrix

$$F_3(r) = \begin{bmatrix} f_0 & F_{01} & F_{02} & f_{03} & F_{02}f_{24} \\ F_{10} & f_1 & F_{12} & F_{10}f_{03} & F_{12}f_{24} \\ F_{20} & F_{21} & f_2 & F_{20}f_{03} & f_{24} \\ f_{30} & f_{30}F_{01} & f_{30}F_{02} & f_3 & f_{30}F_{02}f_{24} \\ f_{42}F_{20} & f_{42}F_{21} & f_{42} & f_{42}F_{20}f_{03} & f_4 \end{bmatrix}$$

$$\begin{aligned} F_{01} &= f_0(f_{01} + f_{02}f_2f_{21})f_1 \\ F_{02} &= f_0(f_{02} + f_{01}f_1f_{12})f_2 \\ F_{10} &= f_1(f_{10} + f_{12}f_2f_{20})f_0 \\ F_{12} &= f_1(f_{12} + f_{10}f_0f_{02})f_2 \\ F_{20} &= f_2(f_{20} + f_{21}f_1f_{10})f_0 \\ F_{21} &= f_2(f_{21} + f_{20}f_0f_{01})f_1 \end{aligned}$$

1) With restricted alternative routing

$$F_3(r) = \begin{bmatrix} f_0 & F_{01} & F_{02} & f_{03} & F_{02}f_{24} \\ F_{10} & f_1 & F_{12} & F_{10}f_{03} & F_{12}f_{24} \\ F_{20} & F_{21} & f_2 & F_{20}f_{03} & f_{24} \\ f_{30} & f_{30}F_{01} & f_{30}F_{02} & f_3 & f_{30}F_{02}f_{24} \\ f_{42}F_{20} & f_{42}F_{21} & f_{42} & f_{42}F_{20}f_{03} & f_4 \end{bmatrix}$$

$$\begin{aligned} F_{01} &= f_1f_{01}f_1 \\ F_{02} &= f_0f_{02}f_2 \\ F_{10} &= f_1f_{10}f_0 \\ F_{12} &= f_1f_{12}f_2 \\ F_{20} &= f_2f_{20}f_0 \\ F_{21} &= f_2f_{21}f_1 \end{aligned}$$

2) No alternative routing

c) Functional matrices

Figure 2.10 Combined mesh and star network with associated matrices

2.4.2 Functional matrix of a star network

Figure 2.9 shows a star network with associated matrices. The structure consists of a group node, *G*, which connects several other nodes (end nodes).

We can distinguish between two cases:

- Full autonomy for the end nodes
- No autonomy for the end nodes.

A node with full autonomy is independent of other nodes with respect to establishing internal connections. A node with no autonomy has to use another node (*G* in Figure 2.9a) also for internal connections. Both degrees of autonomy will exist in a real network.

It is supposed that there is no terminating or originating traffic in *G*. If so, all terminating/originating sources are viewed as nodes. With the same assumption as before the functional matrix can be formed. The result of transforming the functional matrices shows that a star network with full autonomy for the end nodes corresponds to a mesh network with routing over the same node for all connections (except for internal connections in the nodes).

2.4.3 Functional matrices for combined mesh and star network

Combining mesh and star network is often a practical situation. The matrices for such a combined network (Figure 2.10a) are shown in Figure 2.10b–c.

2.4.4 Control data

As mentioned above data transported in a telecommunication network consist of three main types:

- Control or signalling data
- Data between nodes and users
- User data.

The system (network) has no direct control of the generation of user data.

In general, the data for control purposes will consist of:

- Data d_l assigned to line status and associated operations
- Address data:
  - d_k for address to a node k
  - d_a for address of the terminal in destination node a.

The total amount of control data, measured in bits, depends among other things on the signalling system and possibilities for rerouting.

Bibliography

1 Agrawal, D. Graph theoretical analysis and design of multistage interconnections networks. *IEEE Transactions on Computers,* 32(7), 1983.

2 Ahmadi, H, Denzel, W. A survey of modern high-performance switching techniques. *IEEE Journal on Selected Areas in Communications,* 7(7), 1989.

3 No. 1 ESS. *Bell System Technical Journal,* XLIII(5), part 1 & 2, Sept. 1964.

4 No. 2 ESS. *Bell System Technical Journal,* 48(8), Oct. 1969.

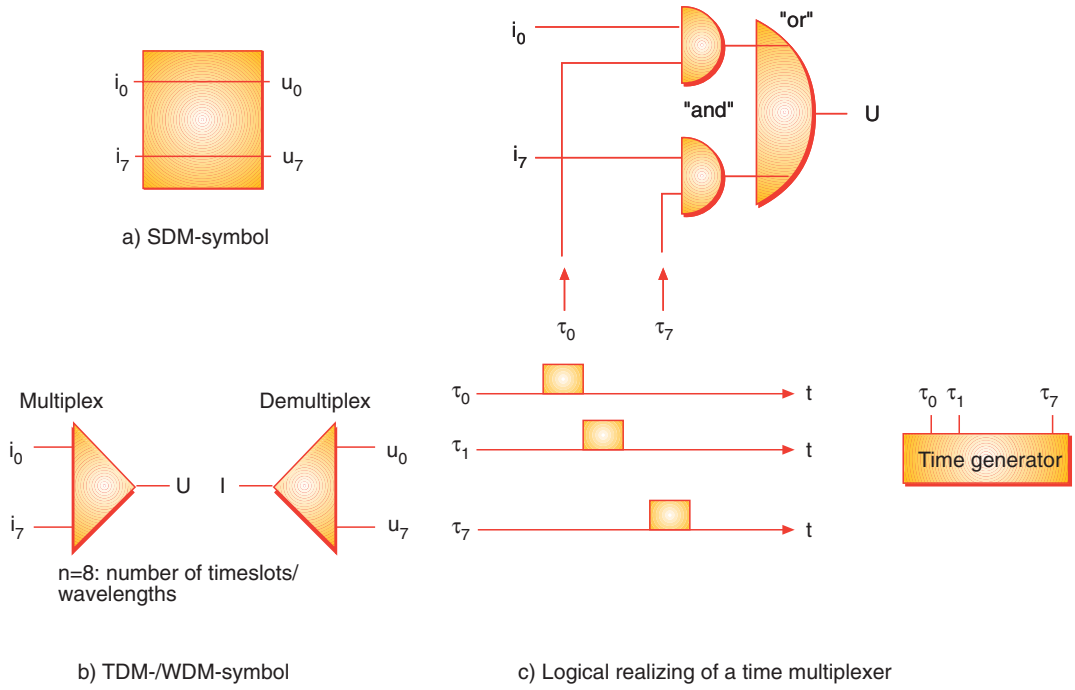


Figure 3.1 Illustration of multiplexing

3 Basic switching principles

3.1 Definitions

We will first give some definitions of the most common notions.

Terminal is the in- or outlet of an equipment.

Connection is a possible way (route) through a switching network from an inlet to an outlet (terminal). The connections are (normally) supposed to be disjunct, i.e. they have no parts in common.

A route means a path through a switching network.

Channel will be used in time multiplexed systems and denotes an assigned time slot.

Congestion means that it is not possible to make a connection between specified idle terminals. (This definition must not be mixed with the word "loss" which is related to idle and non-existing terminals.)

Simplex connection means that data can be transmitted only in one direction.

Duplex connection means that data can be transmitted in both directions.

Synchronous transmission is used in time multiplexed transmission and means that every data element has a fixed assigned time slot with fixed length and phase.

Asynchronous transmission is also used in time multiplexed transmission but in principle with no fixed length and phase of assigned time slots.

Multiplex means that information (data) from different sources are put together and transmitted through a common transmission system.

A *multiplexer* means an equipment that is able to execute multiplexing.

Multiplexing can be divided into three groups (see Figure 3.1):

- Space multiplex,
  SDM (Space-Division Multiplex) represented by a bundle of physical lines (Figure 3.1a).
- Time multiplex,
  TDM (Time-Division Multiplex) where a fixed time slot is assigned to each inlet
- Wavelength/frequency multiplex,
  WDM (Wavelength-Division Multiplex) where a fixed wavelength/frequency is assigned to each inlet.

It is important to fully realise that in the word TDM there is no dynamic allocating of time slots. The definition says that there is a fixed allocation of time slots to the various terminals. Such a TDM will be named a "static multiplexer" (short MUX). A multiplexer with dynamic or programmable assignment of time slots is said to execute switching functions (see later).

A "digital signal" is defined as a signal coded such that it can be represented by a sequence of 0 and 1.

The standard digital representation for voice transmission is called PCM (Pulse-Code-Modulation). The analogue voice signal is sampled at fixed time intervals and the samples are coded to an 8 bit digital signal, transported to the receiver where the opposite process is executed. More such analogue/digital (A/D) signals can be time multiplexed on the same transmission line (channel) by multiplexers as in Figure 3.1.c.

A PCM system of first order has the following main data (see Figure 3.2):

- Time is divided into frames of 125 µs each (sampling frequency 8000 Hz).
- Each frame is divided into 32 time slots in the 30 channel system and in

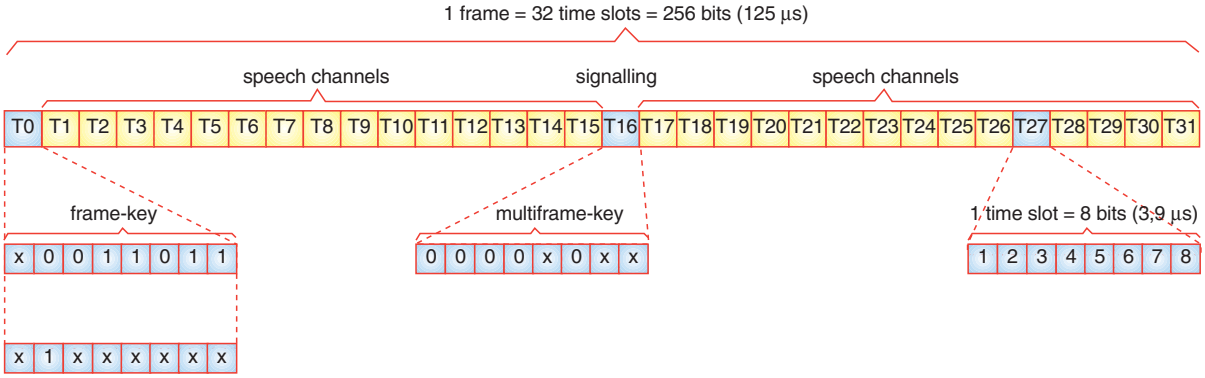


Figure 3.2 Frame structure for a 30-channel system

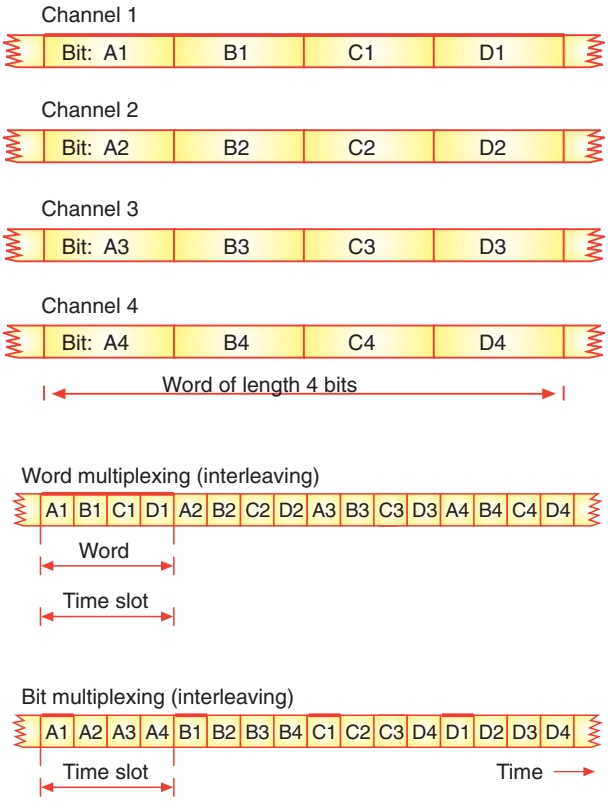


Figure 3.3 Word and bit multiplexing

24 time slots in the 24 channel system.

- 1 time slot is used for signalling purposes, 1 time slot is used for synchronization and the rest for speech or data.
- Data in each time slot is represented by an 8 bit code.

The PCM terminal (PCM multiplexer) converts the analogue signals to a 2048 kbit/s bit stream. Each channel can transmit analogue signals with a bandwidth of 300 – 3400 Hz with a bitrate of 64 kbit/s.

From the basic principle of time multiplexing the processes are "parallel to serial" and "serial to parallel" translation. A digital signal (more than one bit), to be transmitted in a time slot, can be multiplexed in two different methods (see Figure 3.3):

- Word multiplexing
- Bit multiplexing.

Word multiplexing can be used when the time slot permits serial transmission of all bits. This method is usually used in digital switching. Bit multiplexing multiplexes only one bit at a time as shown in the figure.

A first order PCM system can be used as the basic unit for the second order PCM system as illustrated in Figure 3.4, where four first order PCM systems are multiplexed on the same transmission media. The second order system can be used as a basic unit to build a third order system, and so on (see Figure 3.5).

Statistical multiplexing means that the user has assigned a transmission channel only when he has data to transmit and is disconnected when no data is to be transmitted.

Frequency or wavelength multiplexing means that signals with different frequencies or wavelengths are modulated on a common transmitting media. This method is used in transmission system and only to a small extent in switching. With the new optical technology, however, it seems that the wavelength switching can be a realistic alternative (see later chapter).

Demultiplexing means the opposite process of multiplexing.

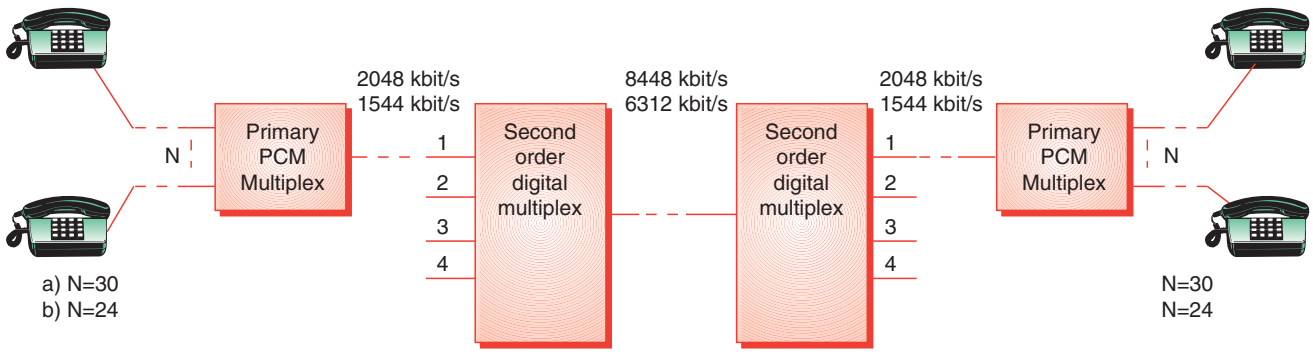


Figure 3.4 Second order digital system

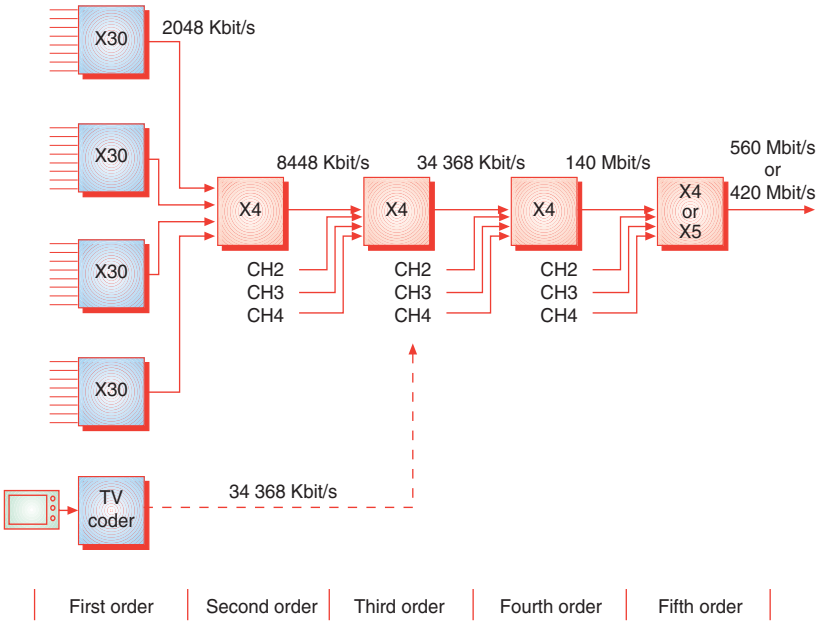


Figure 3.5 The TDM hierarchy in Europe

3.2 Definition of a switch

Switching (in telecommunication) is formerly defined as the task to steer the stream of information (data) in the right direction while transmission has to "transport" this information in the chosen direction.

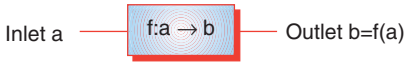
The basic element in a switch is the "switching element" (SE) which can be defined as a binary element with two states:

ON state: as much as possible of the signal is transmitted from the inlet to the outlet

OFF state: as little as possible of the signal is transmitted from the inlet to the outlet

ON state can be represented by 1 (logical 1) and OFF state by 0 (logical 0).

For further studies, however, it is necessary to give a mathematical definition of a switching element. This can be done by the following definition:



f(a) = b, domain is a, range is b
(f(a) = picture of a)

Figure 3.6 Definition of a switching element

"A switching element is an equipment or procedure which is able to execute a programmable assignment of an inlet to an outlet such that the assignment is 1-1 and ON."

A 1-1 and ON assignment means that the outlet is assigned to only one inlet and no other inlets are assigned to the same outlet (see Figure 3.6).

Programmable means that the assignment can be executed or not, depending on some conditions. This is an important difference between a switching element and an element with fixed assignment like a multiplexer.

It is also very important to be able to control a switching element which demands:

- Detecting states – implies logic
- Changing the state – implies logic
- Preserving the state – implies memory.

The memory can be represented by 1 bit where bit = 0 means OFF state and bit = 1 means ON state.

The control function can be entirely or partially integrated in the element. Often, the functions are distributed such that the memory is integrated in the switching element while the logic functions are implemented by external equipment.

A "switch" can be viewed as a collection of switching elements which are able to execute switching functions. A mathematical definition of a switch can be given as:

"An equipment or procedure which is able to execute a programmable assignment of a set of inlets $\{i_1, i_2, ..., i_n\}$ to a set of outlets $\{u_1, u_2, ..., u_m\}$ such that:

- an inlet is assigned only one outlet
- every outlet is assigned to only one inlet."

The set of inlets is called the domain and the set of outlet is called range. If $n = m$ the assignment is said to be 1-1 and ON if the conditions above are satisfied.

Alternatively, a switch can be defined as:

"A switch is an equipment or procedure which optionally may be used to make pairs of inlets and outlets such that one inlet and one outlet are found in only one pair."

A switch also needs memory and logic and so we have the main units as shown in Figure 3.7 with:

- A switch unit, SW
- A control unit, CU.

A larger switching network; SWN, can be built by putting together more switches which again are made of switching elements. This is illustrated in Figure 3.8 which can be seen as a general structure for a switching network.

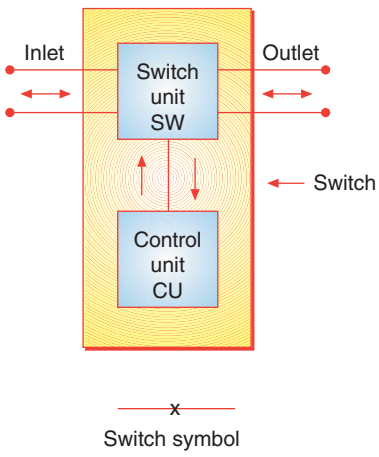


Figure 3.7 Main unit in a switch, SW

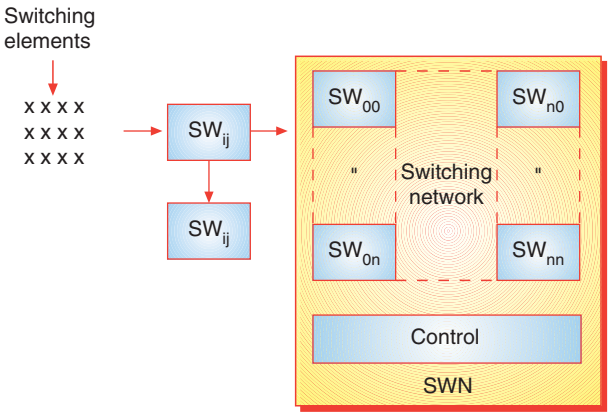


Figure 3.8 Hierarchical structure of a switching network

3.3 Space switching, SSW

In a space switch a pair of inlets and outlets are connected during the whole data transmission. This means that at the same time there must be a distinct connection for every inlet and outlet pair or the number of connections must be the same as the number of inlet/outlet pairs the switch has to connect.

A simple example of an SSW network is given in Figure 3.9a which shows a relay matrix. The matrix makes it possible to have four connections at the same time as shown for the pairs (*a,b*), (*c,d*), (*e,f*), (*g,h*) in Figure 3.9b. Note that four selectors, each with one inlet and four outlets, give the same configuration, see Figure 3.9.c.

The total number of elements in such a matrix is given by:

$K = IU$, where I and U are the number of inlets and outlets, respectively.

K will grow quadratic if $I = U$.

To decrease this quadratic growth of the number of switches it is common to construct a larger network of smaller matrices. As we shall see later this will give the number of switching elements per inlet $n_k < N$ where N = number of inlets.

Figure 3.10 shows an example of a switching network constructed from smaller matrices. The figure shows a small part of the network in NR 1 ESS (from Bell Labs, USA). (NR 1 ESS was the first computer controlled public exchange.) The switching elements are called ferreed element (reed relay with magnetic holding) and in the first stage binomial concentration is used (the elements are placed as in a binary-decimal decoder).

A characteristic property of an SSW network is that it is not (normally) necessary to store the signal to be transmitted whether this is analogue or digital. There are no limits with respect to the size of the network or the number of simultaneous connections.

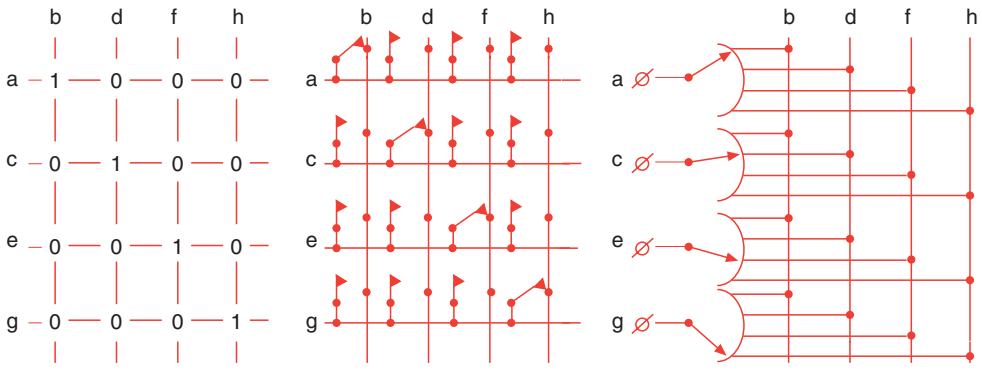


Figure 3.9 Space switching, SSW

3.4 Control operations in SSW network

The basic control operations of an SSW network are explained previously:

- Testing the states of elements
- Operating the elements
- Keeping the elements in the ordered states.

In addition to the basic control operation mentioned above, there are a lot of other functions like analysing digits, selecting connections, etc. A closer discussion of these topics are out of the scope of this paper, but these functions are of course very important to the whole operation of the switching network, especially with respect to public services. Here, we will only give a short discussion to give an impression of the complexity of controlling a switch.

Take a simple switch as in Figure 3.9a and let a possible connection between inlet i and outlet u be designated k_{iu}. Then, for correct operation, the following condition must be satisfied:

$$\left[\begin{array}{l} k_{ix} = 0 \text{ for all } x \neq u \text{ (line } i\text{)} \\ k_{yu} = 0 \text{ for all } y \neq i \text{ (column } u\text{)} \end{array}\right] \quad (3.1)$$

where

$\forall$ means for all
$\Rightarrow$ means implies.

This gives a control hardware as schematically shown in Figure 3.11. When the switching element, given by the coordinate (i = line, u = column), is ON, this element has to keep OFF all other elements in the corresponding line

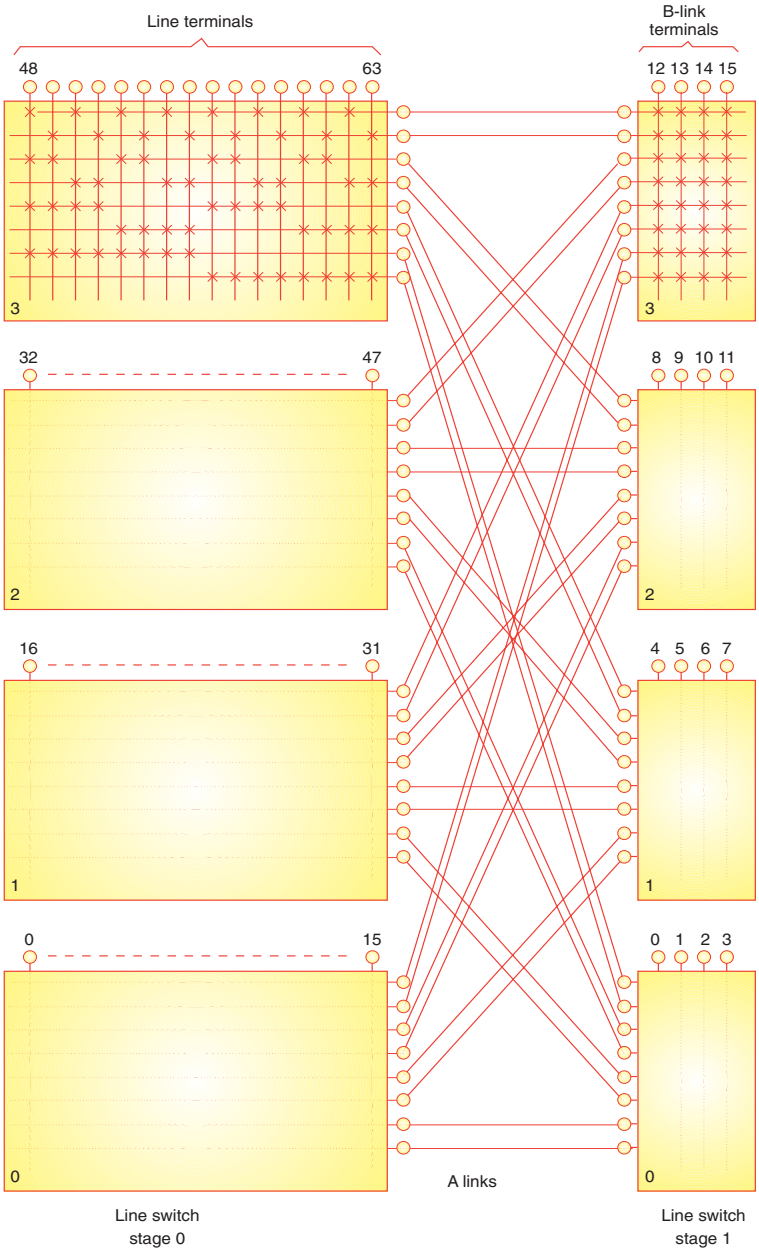


Figure 3.10 Part of the network used in NR 1 ESS from Bell Labs (USA)

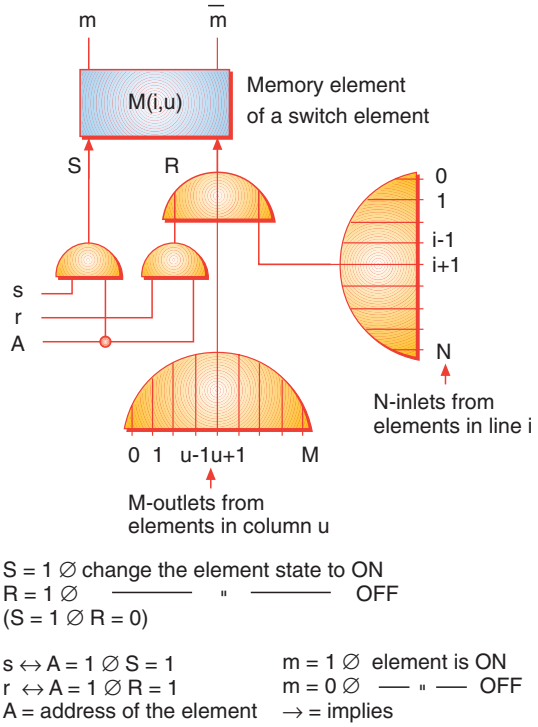


Figure 3.11 Example of hardware to control a switching element in an SSW matrix

and column. This small example gives an idea of the complexity of controlling a large switch. With the state of the semiconductor technology, however, solving the construction of such a hardware is no problem but the control will (normally) give a much more complicated construction than the bare switch.

3.5 Synchronous time switching, TSW

In time switching a pair of inlets and outlets is assigned a time slot and data are transferred only during this slot. This process repeats periodically in a synchronous system. The time interval between two successive time slots belonging to the same connection or data transfer is called a (time) frame.

Figure 3.12 shows an example of time multiplexed switching. Two terminals which are to be connected are assigned a time slot. During that time the momentary value of the signals from the terminals are transmitted over the common line to the other side. Every user (terminal) has a switch and a filter which removes unwanted frequencies. The principle is called PAM-TSW (where PAM means Pulse Amplitude Modulation).

The highest frequency which can be transmitted by such a system is given by the sampling theorem:

$$f_s = 2f$$

f_s = sampling frequency,
f = signal frequency (3.2)

In Figure 3.13 the analogue sample is coded such that the value can be represented by a binary code. The unit doing this is called encoder. This digital coded signal is transmitted by the transmission line to a decoder which translates the digital code back to analogue representation

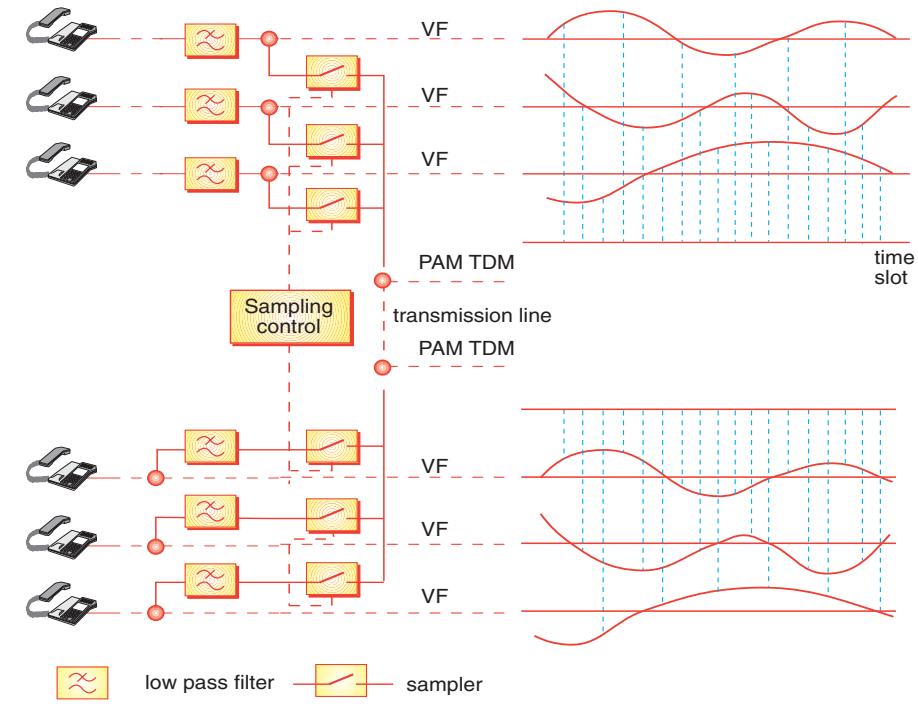


Figure 3.12 Principle of PAM-TSW

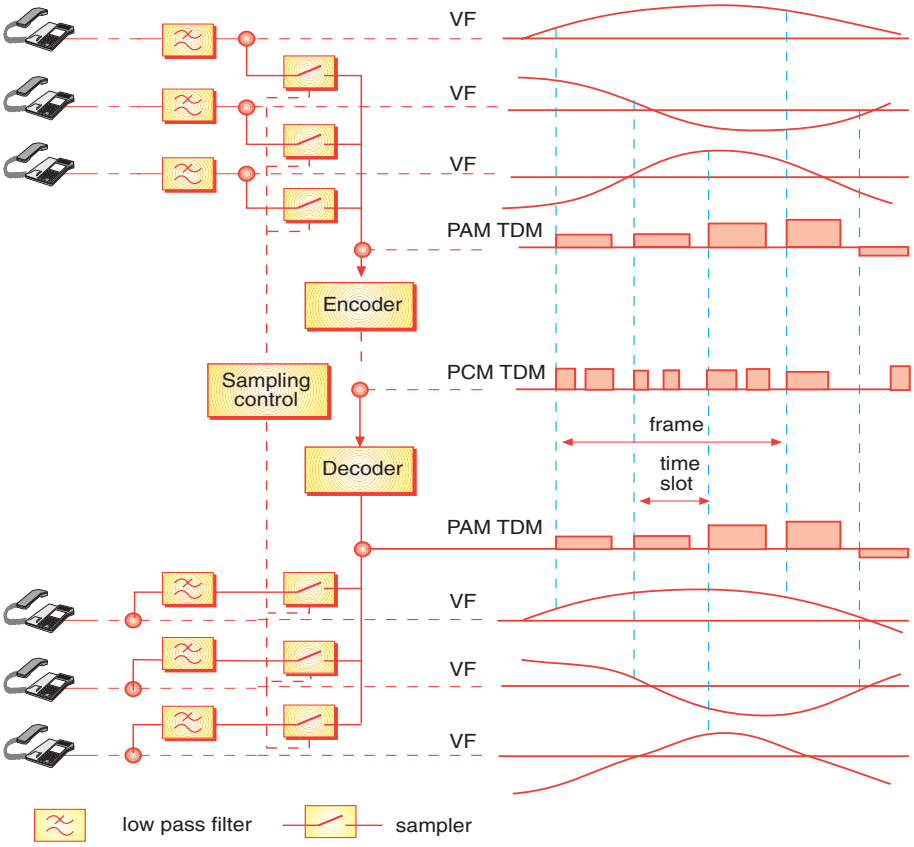


Figure 3.13 Principle of PCM-TSW

and then it is transmitted to the user. This time switching is called PCM-TSW (PCM = Pulse Code Modulation). In modern systems both the encoder and the decoder are placed in the terminal such that switching and transmission are digital.

TSW switching applying the matrix in Figure 3.9a gives a switch as shown in Figure 3.14a. The connections are established at different times so that a common line can be used. Figure 3.14b shows how the switch can be constructed with (programmable) multiplexers.

An important difference between a TSW and an SSW is the number of switching elements per inlet. The switch in Figure 3.14a has two elements per inlet while the switch Figure 3.9a has four elements per inlet. For TSW, two elements per inlet is sufficient until time and delay considerations restrict the use of only TSW principles.

The maximum number of inlets/outlets (terminals) a TSW can have is given by:

$$N = T_0 / \Delta\tau \qquad (3.3)$$

where

T_0 = time between sampling on the same terminal = 1 frame

$\Delta\tau$ = time to transmit the sample through the switch

Example: $T_o = 125$, $f_s = 8$ kHz and $\Delta\tau = 0.1$ µs gives $N = 1250$.

The possibility of implementing a relatively large switch with a small number of switches per terminal is one of the reasons why TSW is so attractive in switching. The drawback in relation to SSW is reduced bandwidth.

The time switch in Figure 3.14 indicates a space-time and time-space switching since there is switching from/to a distinct terminal in an assigned time slot. Between the units executing these operations data are transmitted in serial form. If it is possible during this transport to move data from one time slot to another arbitrary time slot and satisfy the requirements stated earlier, a switching function is achieved. To do this one can use the following basic units:

- Static multi-/demultiplexers (MUX/DMUX)
- A T-switch defined as an equipment/ procedure which programmable moves

data from one time slot to another time slot.

- An S-switch defined as an equipment/ procedure which programmably distributes time slot data in space.

Used on the situation depicted in Figure 3.14, gives a configuration shown in Figure 3.15a.

To implement a T-switch it is necessary to use time delay or storing. By using a time delay, data assigned a time slot are delayed a number of time slots, giving a new assignment between data and time slot. A shift register can be used as a digital delay mechanism as shown in Figure 3.15b. The principle is:

- Data from MUX are written in SH_i during a time frame
- A space switch SSW moves the data from SH_i to shift register SH_u and during this action decides where to place the data in SH_u
- In the next frame data are shifted out from SH_u.

Instead of delay lines one may use RAM (Random Access Memory) to store and delay data with digital representation. An implementation of a T-switch can then be done with the essential functional units shown in Figure 3.16.

- A data memory, M, with 1 cell (word) assigned to each time slot and able to store data from one time slot
- A control memory, CM, with one word (cell) assigned each word in M and containing the address to the word in M where data are to be placed
- An address counter, COUNT, which addresses M, CM and MUX/DMUX
- A pulse generator, PG, which:
  - Drives COUNT
  - Write and read pulses to M/CM.

Data from MUX assigned a time slot are stored in M. M is addressed by the address given by CM which, as MUX and DMUX, is driven cyclically by COUNT. CM gives the address to M at which data from MUX are to be written. The figure shows an example with four time slots.

In Figure 3.16 every time slot is divided in a write and read phase and the switching is achieved during the writing (write phase). The data from M are read cyclically and distributed to the outlets via

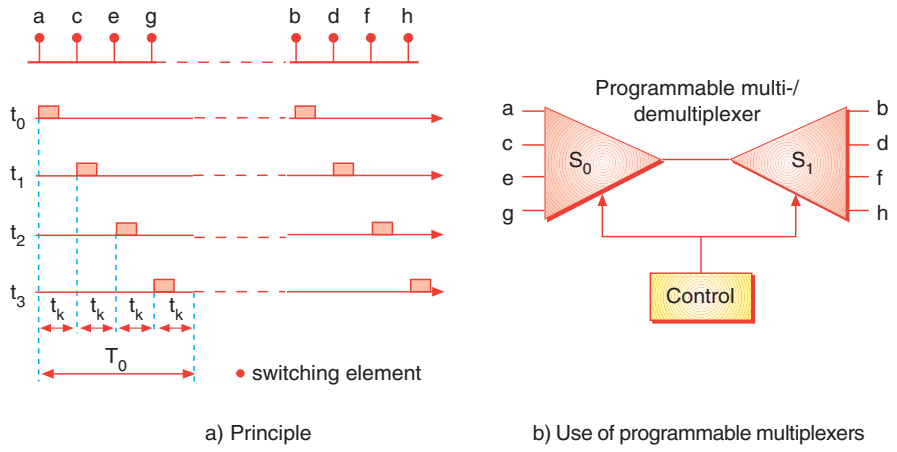


Figure 3.14 TSW implementation of a 4 x 4 SSW matrix

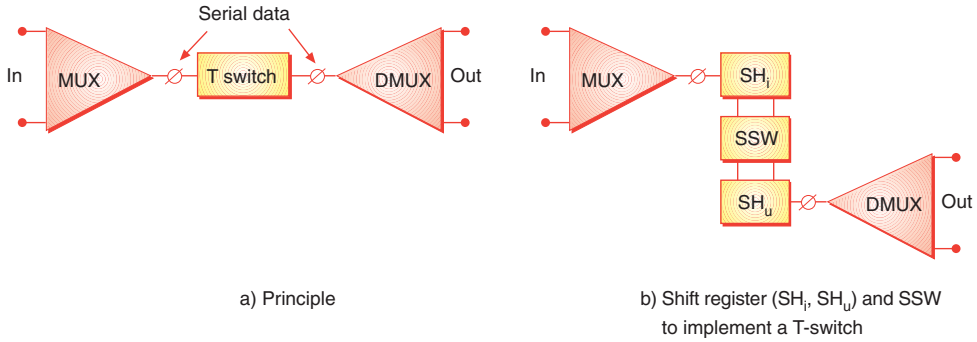


Figure 3.15 Basic units in a TSW unit

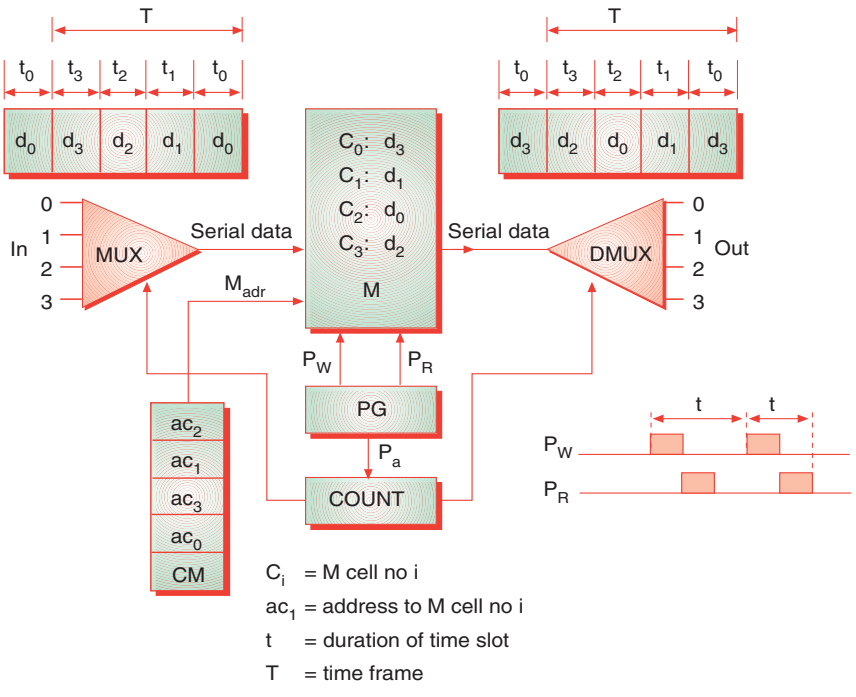


Figure 3.16 Functional units in a (write controlled) T-switch

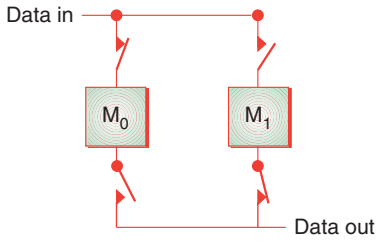


Figure 3.17 Principle of frame T-switching

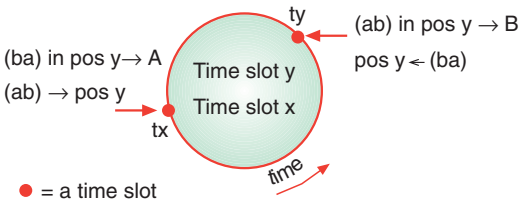


Figure 3.18 Illustration of the changing principle

DMUX. This principle is called "write controlled" T-switching. Another method is to write data cyclic in M and let CM control the reading from M. This method is called "read controlled" T-switching.

In T-switching the time delay will vary between the limits:

$$T_0 / n \leq \Delta d \leq T_0$$

where

n = number of time slots

T_0 = frame period

$T_0/n = \Delta\tau$ is the duration of a time slot.

Storing capacity is:

Data store: $N_M = nb$

Control store: $N_S = n\log_2 n$

Memory access time: $t_M \leq T_0 / 2n$

Bit per data set: b.

A first order PCM system with $b = 8$ bit, $n = 32$ channels and $T_0 = 125$ µs gives:

$N_M = 32 \cdot 8$ = 256 bit

$N_S = 32 \log_2 32 = 32 \cdot 5$ = 160 bit

Sum = 416 bit

$t_M \leq 125 / 2.32$ µs = 1.98 µs

Using two memories M_0 and M_1 (see Figure 3.17), the access time can be decreased. Data for a whole frame are written in M_0. During the next frame data are written in M_1 while reading out data from M_0. In the following frame it is shifted to M_0 and so on. The principle is called "frame" T-switching.

This principle needs:

Data store: $N_M = 2nb$

Control store: $N_S = n\log_2 n$

Memory access time: $t_M \leq T_0 / n$

The same example as above (first order PCM system) now gives:

$N_M = 2 \cdot 32 \cdot 8$ = 512 bit

$N_S = 32 \cdot \log_2 32 = 32 \cdot 5$ = 160 bit

Sum = 672 bit

$t_M \leq 125$ µs $/ 32$ = 3.96 µs

These examples give simplex connections. To get a duplex connection one has to establish a connection in the opposite direction. This can be done as above or by using a changing principle illustrated in Figure 3.18:

- In time slot t_x assigned to inlet A:
  - Data (ba) from B are read out to A
  - Data (ab) from A are written into memory position y
- In time slot t_y assigned inlet B:
  - Data (ab) in position y are read out to B
  - Data (ba) from B are written into memory position y.

Capacity and access time:

Data store: $N_M = nb / 2$

Control store: $N_S = n\log_2 n / 2$

Access time: $t_M \leq T_0 / 2n$ (same as "time slot" switching)

Usage of this exchange principle could give some problems in switching networks with many time frames. Such problems are not covered here.

3.5.1 Larger TSW switching network

The capacity (here measured in number of inlets/outlets) of a TSW network is, as stated above, given by the relations between the frame time T_0 and time duration of time slot τ_A:

$N = T_0 / \tau_A$ = number of "simultaneous connections"

Time is obviously an important parameter and N can be increased until time limits the growth.

In Figure 3.19 is shown an example of a switching network with g groups each of n channels (time slots τ_k). In each channel an 8 bit data word can be handled. The control storage, CM, is common for all groups. Each group consists of a shift register, SH, and a data storage, M, with one word (cell) per time slot. The serial data (here 8 bit) are shifted in assigned shift registers during the P_{SH} pulses from TG.

Transfer from an SH_{ij} to the corresponding data store M_{ij} is done at the flank of each time slot. The addresses of the data store are given by the time slot number (data are written cyclically in associated M_{ij}).

In the same way the (parallel) data from the outlet data store are transferred to the outlet shift register from which data are shifted out to the terminal. Switching is achieved because the control storage contains the addresses which data the input data storage should be transferred to, i.e. read controlled switching. Data are transferred as 8 bit data sets, in parallel, over a common bus to/from the data stores.

Available total time for the operations is a time slot τ_k and this is divided (see the figure) into $g + 1$ parts with one part for each group. Each of these P_g pulses are divided into 8 parts (P_{SH}).

The number of available channels (time slots) are:

$$N = n \cdot g$$

Time for each data transport is:

$$t_{ACC} = \tau_k / g = T_0 / ng$$

Example: $T_0 = 125$ µs, $n = 32$, $g = 64$ gives:

$t_{ACC} = 125$ µs $/ 32 \cdot 64$ ≈ 61 ns

$N = 32 \cdot 64$ = 2048

An access time like this is no problem with today's technology which gives access time < 1 ns. With decreasing time, however, one will realise other problems (see later chapters). It will also be shown later that the bus principle used in Figure 3.19 can be extended to a bus hierarchy.

By doubling the length of the shift registers and the words in the data stores, N will also be doubled (to 4096).

These examples show the possibility of implementing relatively large switching networks without using space switching.

In a TSW network, a time slot (or a channel) corresponds to one connection in a space network. The switching is periodic, such that after a fixed time (a time frame) an inlet/outlet pair has to be connected again. A few examples above show that the technology sets limits to the size of such a network. To overcome these limits or to better utilise existing technology, data storing is obviously an essential method. The development of the semiconductor technology has given cheap and fast components for storing digital data and so favouring digital data representation.

3.6 Control operations in a TSW network

The control of the connections in a TSW switch can be stated as (see Figure 3.9 and 3.11):

$$\forall(i,u)[\{k_i \cap k_u \cap t_{iu} = 1\} \Rightarrow \{k_x = 0 \text{ for all } x \neq i \text{ and } u\}] \qquad (3.4)$$

t_{iu} = time slot assigned to inlet number i and outlet number u

$\Rightarrow$ = "implies that"

$\forall$ = for all.

The switches k_i and k_u, assigned inlet i and outlet u, respectively, are closed at the same time while all other switches are open.

The same operations as in a space switch have to be executed. The difference is that the operations are performed during a distinct time slot. The time a switch is closed, is the time available for data transfer. If data are digital, the switching element will (normally) be designed with digital gates and the time for changing the states of the switch will be rise and fall time of the gates.

In a TSW unit all data concerning the status of the switch will be found in the associated control store. We have a memory map of the network.

Since the operations in a TSW switching network are controlled by clocks, it is obvious that some synchronising problems will arise. One has to take into account such parameters as tolerances,

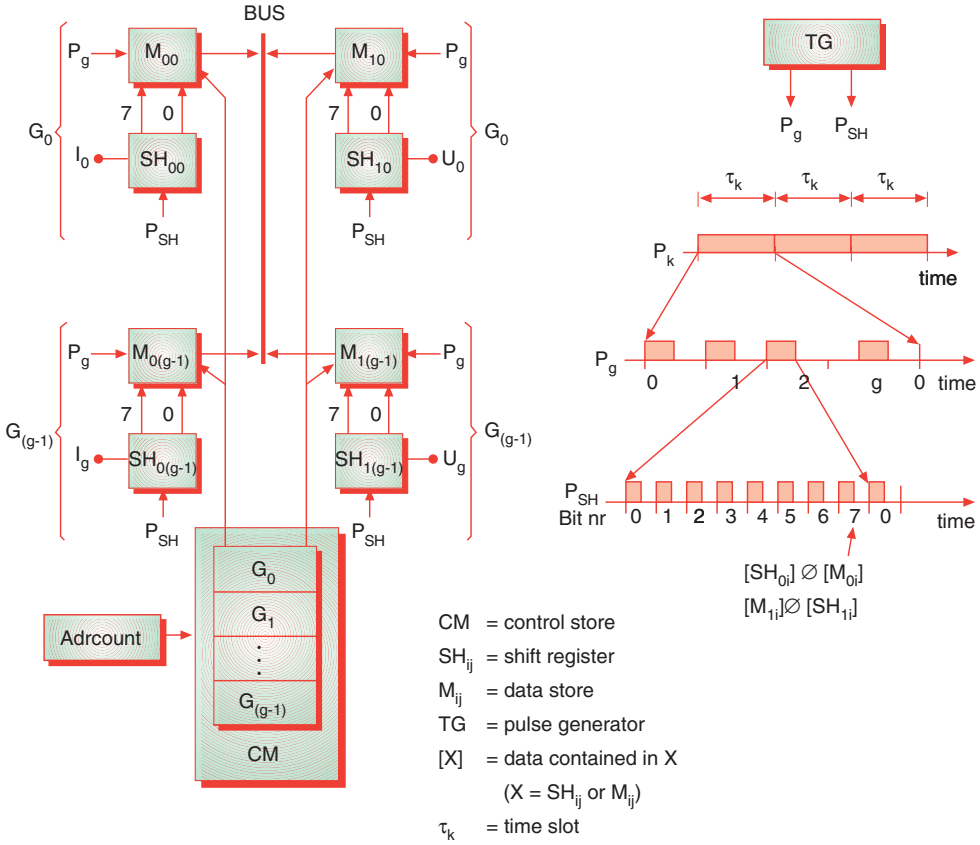


Figure 3.19 Implementation of a large TSW network by using T-switches and a bus

transmission delays, etc. Problems concerning synchronising are discussed in many textbooks and therefore not dwelled on here. Inside a switching node, however, it is felt that those problems are more easy to handle.

Bibliography

1 Ahmadi, H et al. A high-performance switch fabric for integrated circuit and packet switching. *Infocom '88,* New Orleans, March 1988.

2 Birdsall, T G, Ristenbatt, M P, Weinstein, S B. Analysis of asynchronous time multiplexing of speech sources. *IRE Transactions on Communications Systems,* December 1962.

3 Boissier, A, Glatigny, A, Thepaut, L. Space division optical switching system of medium capacities. *Proceedings SPIE,* 585, 65–70, 1985.

4 Bonatti, M, Fratta, L, Tomasi, C. A call set-up architecture for very large switching systems. *Globecom '86,* Houston, December 1986.

5 Braugenhardt, S, Bjurel, G. Transition to a digital network. *International Switching Symposium,* Kyoto, 1976.

6 Braugenhardt, S, Jönsson, I, Nilsson, B. Introduction of the AXE 10 switching system in the telephone network. *Ericsson Review,* 53(2), 60–69, 1976.

7 Braugenhardt, S. *Digitala telefonväxlar : en jämförande studie av tidsuppdelade kopplingsnät.* Stockholm, Department of Telecommunication Networks and Systems, Royal University of Technology, March 1972. (In Swedish.)

8 Buchner, R B, Milort, W. The switching network of Philips' PDX system. *Philips Telecommunication Review,* 31(3), 1973.

9 Cattermole, K W. Principles of pulse code modulation. *Iliffe,* London, 1969.

10 *Digital telephony : an introduction.* Stockholm, Telefonaktiebolaget L M Ericsson, Jan. 1977.

11 Duerdoth, W T, Cheeseman, D S. Techniques for an integrated digital network for telephony. *International Switching Symposium,* Boston, 1974.

12 Fraleigh, J B. *A first course in abstract algebra.* London, Addison-Wesley, 1969.

13 Geissler, H. Planning of a PCM system hierarchy. *NTZ-Report 8,* Berlin, VDE-Verlag, 1971.

14 Gerke, P et al. The time division multiplex switching of digital speech channels. *IEEE Transactions on Communications,* vol. 22, 1721–1275.

15 Granello, G. Switching networks for PCM time-division exchanges. *IEEE Proceedings ISS 1972,* 81, 1972. (72 CH 0 617-1-COM.)

16 Harrison, S R, Andresen, S. TDM strictly nonblocking connecting networks employing fixed delay elements. *Proceedings ICC 72,* Ch. 40-13, 1972.

17 Hartley, G C et al. *Techniques of pulse code modulation in communication networks.* Cambridge University Press, 1967.

18 McDonald, J C (ed.). *Fundamentals of digital switching.* New York, Plenum Press, 1983.

19 No. 2 ESS. *Bell System Technical Journal,* 48(8), Oct. 1969.

20 Pan, J W. Synchronizing and multiplexing in a digital communications network. *Proceedings IEEE,* 60(5), 594–601, 1972.

21 Rekiere, B J et al. An introduction to PCM switching. *Automatic Electric Journal,* April 1971.

22 Schindler, H R. Delta modulation. *IEEE Spectrum,* Oct. 1970.

23 Slabon, R. Introduction and realization of the PCM switching network in the EWS 1 electronic switching system, *NTZ 28,* H.3, 102–107, 1975.

24 Streitlien, B. The switching network in ITT 1240. *Telektronikk,* 80(2/3), 1984.

25 Vaughan, H E. An introduction to No. 4 ESS. *International Switching Symposium,* Boston, 1972.

26 Widl, W. PCM transmission. *Ericsson Review,* 49(1), 1972.

27 Wragge, H S. The effect of digital switching techniques on the form of telecommunication networks. *Proceedings I.R.E.E. Australia,* Nov. 1970.

Table 4.1 Switching parameters

Parameter	RE	LE	Remark
Transmission parameters	x		An LE regenerates the signal
Operation speed	x	x	
Bandwidth	x	x	
Linearity	x		LE is very “unlinear”
Lifetime	x	x	
Power consumption	x	x	
Stability	x	x	

4 Parameters of switching elements

The switching elements can be divided into two main groups:

- Relation elements, RE
- Logical elements, LE.

All elements are able to execute a 1-1 and ON assignment. The main differences between those two types are:

- A relation element has undefined inlet and outlet levels i.e. the outlet signal follows the input signal more or less exactly.
- A logical element has defined levels (normally two) and is also called a threshold element since the output level is a fixed level if the input level has a certain minimum.

The transfer function of a switching element can be described by:

$$u = k \cdot i$$

where

u = output level
i = input level
k = proportionality constant (= 1 for a logical element)

An RE is in many situations able to carry both analogue and digital signals (data) while a logical element is only able to carry digital signals. Besides the functional property that an element can have, two states (ON or OFF), there are a lot of other parameters that characterise an element and thereby also the whole switching network. The most important parameters are listed in Table 4.1.

A switch, and thereby the whole switching network, will also be characterised by the same set of parameters. In addition, there is another important parameter:

n_k = number of switching elements per inlet (input)

This is not a real switching parameter but important for both economical and functional reasons. A small n_k means an economical network but can also mean a much more functionally complicated network. The switching principle also has a great impact on n_k as indicated earlier.

For a space switching network we may, within certain limits, state that nk is reduced by:

- Increasing the number of switching stages
- Decreasing the size of the basic switching matrix.

Implemented space switching networks often have an n_k in the range:

$10 - 16 < n_k < M$;
M = number of outlets

Time switching in principle gives a much smaller number of switching elements.

In general, however, it is difficult to characterise a network by this parameter

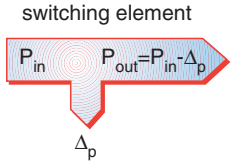


Figure 4.1 Power relation in a switching element

because there are a lot of other parameters like:

- The control functions with test of status, path selection, etc.
- Demands for modularity
- The need for growth of the network.

4.1 Transmission parameters

Figure 4.1 shows the power relation in a switching element. The input power, P_{in}, is divided into a part out P_{out} and a loss part Δ_p.

The transmission parameters are:

- Loss, a_c
- Crosstalk, a_B

Both parameters are defined as the logarithm of a power relation:

$a = 10 \cdot \log[P_{out} / P_{in}]$ dB,
$a = a_c$ for loss
$a = a_B$ for crosstalk

where

P_{out} = power from the outlet of the element

P_{in} = power to the inlet of the element

dB = decibel.

In practice, a voltage (current) relation is used which gives:

$a = 20 \cdot \log[V_{out} / V_{in}]$ dB.

Used on the element in Figure 4.1 we have:

$a = 10 \cdot \log(P_{out} / P_{in})$
$= 10 \cdot \log(1 - \Delta p / P_{in})$ dB

A practical relation element will have $\Delta p \neq 0$. A logical element will have $\Delta p = 0$ (with respect to transmission or transfer function) because the input signal is regenerated in the element.

An ideal element can be defined as an element with no loss ($\Delta p = 0$) when ON and with infinite crosstalk ($\Delta p = P_i$) when OFF. (Higher crosstalk means a better element.)

One also has to take into account that by increasing the size of the switch the aggregated parameters can change in an unfavourable direction. For this reason an element ought to have better parameters than the whole switching network.

The demands on the parameters depend on many factors with the most important ones being:

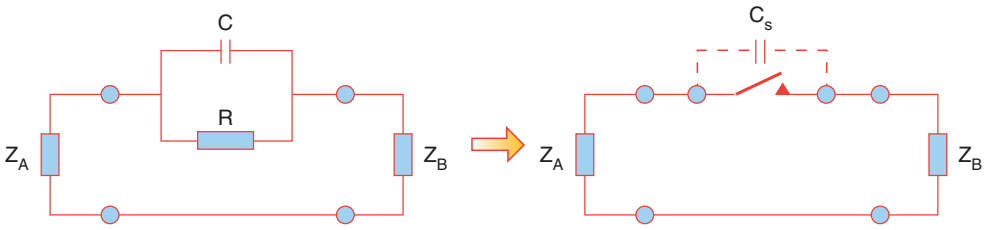


Figure 4.2 Simplified equivalent of an (electrical) switching element

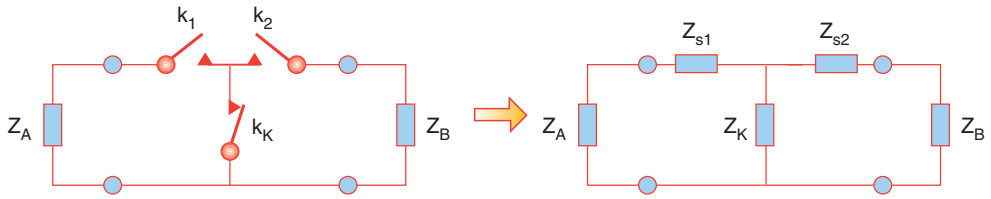


Figure 4.3 Configuration to improve the crosstalk conditions

- Analogue or digital signal on the inlet
- The possibility of amplifying.

With digital signals $a_B \geq$ -30 dBm will give a satisfactory signal/noise ratio, while analogue signals need to have $a_B \geq$ -(60 - 70) dBm to give a satisfactory signal/noise ratio. (dBm is related to a specified input level.) The signal/noise ratio is here defined as the ratio between the input power through an ON element and the power from all other elements transmitted through the ON element.

An element can be represented by its equivalent which, for an electrical element, is shown in Figure 4.2. Higher frequencies and OFF element gives $R << Z_s$ (C) and the influence from R will be negligible.

With this assumption aB can be written:

$a_0 = -20 \cdot \log [1 + 1 / (2 \cdot \pi \cdot f \cdot C \cdot Z_0)]$,
$Z_0 = Z_A + Z_B$

Specifying a lower value a_{min} gives:

$\log C_{max} = -a_{min} / [20 - \log(4 \cdot \pi \cdot f \cdot Z_0)]$

If a_{min} = 60 dBm and Z_0 = 75 ohm this gives:

C_{max} = 250 pF for f = 4000 Hz

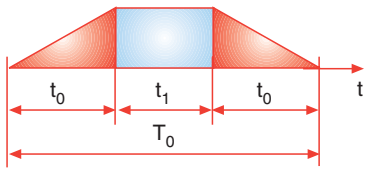
C_{max} = 0.2 pF for f = 5 MHz

At higher frequencies the requirement on the element is so high that more than one element has to be used. A common configuration is shown in Figure 4.3. Here, the two elements k_1 and k_2 are serial connected and a third element k_K is used as a shunt. The shunting element is closed when the two other elements are open.

If the shunt element impedance in ON state is R (pure ohmic), the configuration in Figure 4.3 gives with $Z_A = Z_B = Z_0$ = 75 Ω and a_{min} = 60 dBm:

C_{max} = 28 pF for R = 10 ohm, f = 5 MHz

C_{max} = 9 pF for R = 100 ohm, f = 5 MHz



t_0 = ON-OFF-time

t_1 = hold time

Figure 4.4 Definition of times for a switching element

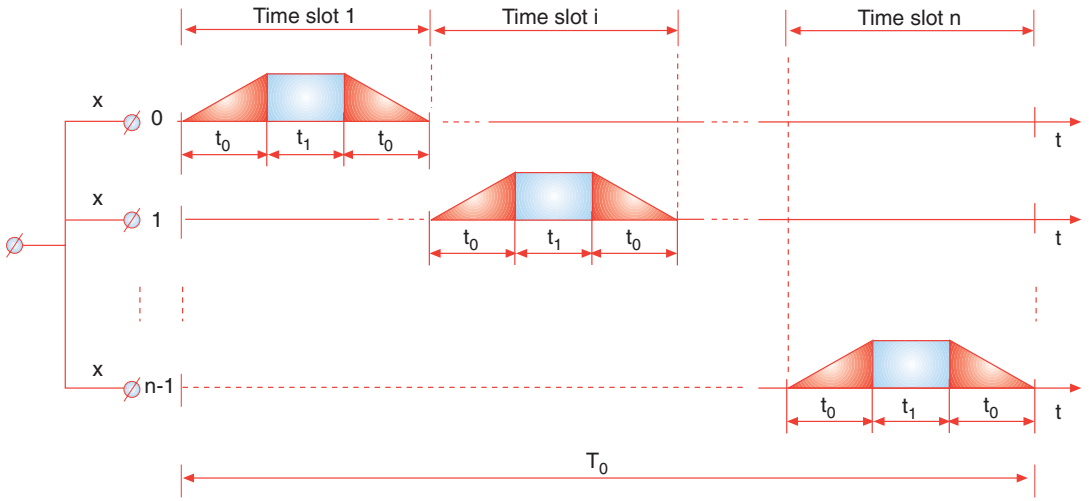


Figure 4.5 Demultiplexing of first order PCM system (schematic)

Metallic (relay) elements (contacts) give $R << 1$ ohm so that such a configuration gives sufficient crosstalk conditions even at very high frequencies.

Electronic (semiconductor) elements will give a somewhat higher R, but still small enough to give elements with very high crosstalk.

A drawback with configurations like the one in Figure 4.3 is that the losses also increase since more elements are serially connected. To overcome this problem one can use amplification.

4.2 Operation speed

The operation speed or switching time is here defined as the time to change from one state to the other, i.e. from ON to OFF or from OFF to ON (see Figure 4.4). In general, these two times will be different.

The demands on the switching times depend on the actual use and results in large time variations (from sec to ns).

Figure 4.4 gives following relation:

$$t_0 = (T_0 - t_1) / 2$$

where

T_0 = the time the element is in use

t_1 = time to transfer data.

A space switch in a public network with voice data has a T_0 (Figure 4.4) measured in seconds (> 180 seconds = mean conversation time) while the operation times of all actual elements are from µs to 3 – 4 sec. In this situation the switching times are no critical parameters.

In an SSW or TSW network with digital data transfer, however, the situation can be quite different. As an example, a first order PCM system is illustrated in Figure 4.5. This gives the following relation:

$$t_0 = (T_0 - nt_1) / (2n)$$

where

T_0 and t_1 are as above

n = number of time slots (channels).

Actual parameters (T_0 = 125 µs and n = 32) for this system give:

t_0 = 1.302 µs for $t_1 = t_0$

t_0 = 1.953 µs for $t_1 = 0$.

Decreasing T_0 corresponding to a sampling frequency of 10 MHz gives:

$t_0 \approx 1.5$ ns

The demands placed on the switching times are seen to increase drastically.

If the input data are digital, the demand on t_0 differs for an LE and an RE element.

For an RE t_0 is given as above. In this situation an LE has to follow changes of the input data. With b bits per transfer this gives:

$$t_0 = (T_0 - nt_1) / 2nb$$

When $b = 8$, $n = 32$ and $t_1 = 0$ we get (1 ps = 10^{-12} sec):

t_0 = 244 ns if T_0 = 125 µs

$t_0 \approx$ 195 ps if T_0 = 100 ns

The last figure is on the limit of today's commercial technology.

After those examples it should be clear that the operating speed of the switching elements is very important for several applications.

4.3 Bandwidth

Bandwidth for analogue data transfer is defined as the frequency range where the signal is attenuated 3 dBm.

In digital transmission the bandwidth is defined as the highest bitrate measured in

bit/s which can be transmitted such that the bit can be isolated.

In both cases the demand for allowed crosstalk gives the higher limit.

The bandwidth and bitrate are therefore referring to a given crosstalk level.

4.4 Linearity

The requirement on linearity (with respect to frequency and amplitude) is of the most importance when data are of analogue type. The parameter is measured as per cent and another word is clatter. In a network composed of several switching elements, the output signal depends not only on the parameters of the involved switching elements, but also on the parameters of the surrounding elements.

Metallic (relay) switches normally have very linear parameters.

Electronic switches are not linear since many of the parameters are dependent on the voltage used in the switch. This unlinearity can be compensated by various means, but on the price of:

- Greater complexity
- Greater power consumption
- Greater physical size.

4.5 Lifetime and reliability

Lifetime and reliability are often in close relation to each other since a short lifetime means poor reliability. A measure of reliability is MTBF (Mean Time Between Failure).

Expected lifetimes of telecommunications systems have so far been in the range of 30 – 50 years which is a very long time span for technical systems. The trend now is lifetimes of 5 – 10 years.

In practice, we can take advantage of the fact that, in general, the failure of components follows an exponential curve with respect to time, so that the bad components are filtered out after a certain "burn in time".

Electromechanical components, such as relays, have shown a long lifetime and high reliability. The same can be said about most of the electromechanical selectors like the Strowger, 500-point and coordinate selectors.

Electronic components also seem to have a very long lifetime. One drawback when employed in telecommunication equipment, is the sensibility to overvoltages which can arise in such networks. Advantages are small size and the fact that it is easy to "burn in" the components to give high reliability and great flexibility.

The new optical technology will overcome problems with electrical noise.

4.6 Power consumption

There is a close relationship between switching time and the necessary energy which has to be supplied to get the switching time. Increasing switching speed (decreasing switching time) also increases the energy consumption. The changing of state needs energy and an expression for this is:

$$W = \text{switching power} = E_w / \tau$$

where

E_w = switching energy (energy to change the state)

τ = switching time

For example $E_w = 10^{-6}$ Ws $= 10^{-6}$ J will give:

$\tau = 100\ \mu s : W = 10$ mW

$\tau = 10\ \mu s : W = 100$ mW

$\tau = 1\ \mu s : W = 1000$ mW

Changing E_w from 10^{-6} Ws to 10^{-12} Ws gives:

$\tau = 1\ \mu s : W = 1\ \mu W$

$\tau = 1$ ns $: W = 1$ mW

$\tau = 1$ ps $: W = 1$ W

This shows that the switching time has a major impact on the power consumption such that:

- High switching speed (small τ) needs high power
- Power to the element means heating and increasing temperature
- The heat has to be taken away.

The limit of how much energy can be consumed by an element is normally set by the temperature. This means that one has a limit on the energy which an element can receive without being damaged.

If the energy is decreased, one gradually reaches a limit which gives "spontaneous switching", i.e. a random change of energy changes the state. So there is also a lower limit on the energy.

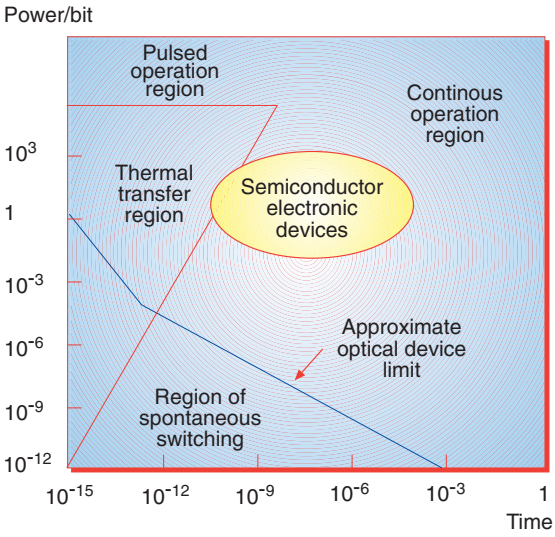


Figure 4.6 Example of characteristics of a semiconductor junction

The above can be expressed as:

$$E_{spontaneous} < E_S < E_{heat}$$

where

E_S = energy to the element

$E_{spontaneous}$ = energy to give "spontaneous switching"

E_{heat} = energy which gives excess heating

Figure 4.6 shows the relationship between switching time and power/bit.

The power can be divided into two parts:

- Power consumed during the state changes
- Power consumed to keep an element in a defined state.

The first part is a product of switching speed and power consumption for each change of state.

In a large switching network with many elements it will be important to keep the power of each element as low as possible.

The power to keep the element operated is independent of switching speed. Some elements like the relay need no power in one state (in the unoperational state). An electronic element needs power in both states but normally this will be much lower than the switching power.

Bibliography

1 Goodman, J W. Fan-in and fan-out with optical interconnections. *Optica Acta,* 32(12), 1489–1496, 1985.

2 *Digital telephony : an introduction.* Stockholm, L.M. Ericsson, Jan. 1977.

3 Hinton, H S. Photonic switching technology applications. *AT&T Technical Journal,* 66(3), 1987.

4 No. 1 ESS. *Bell System Technical Journal,* XLIII(5), part 1 & 2, Sept. 1964.

5 Smith, P W. On the physical limits of digital optical switching and logic elements. *Bell System Technical Journal,* 61(8), 1975–1993, 1982.

5 Switching components

The necessary main components in switching are, of course, the components to execute the switching functions. In addition, it is necessary to have components or equipment to control the switching. Depending on the technology one also needs other components, for example components to join fibre or other optical equipment.

Beside the functional grouping in chapter 2 we also use a technological grouping of the switching elements. The elements can be grouped into two main sets:

- Mechanical elements, i.e. elements with moving parts
- Solid state elements, i.e. elements with no moving parts.

These main groups can be further divided into smaller groups. The first group can be divided into:

- Electromagnetic relays and switching matrices made up of such elements
- Electro-mechanical or electromotor driven selectors like the Strowger selector, the 500-point selector, and the code selector
- Opto-mechanical elements.

The second main group can be divided into:

- Electronic elements
- Optical elements
- Supra conducting elements.

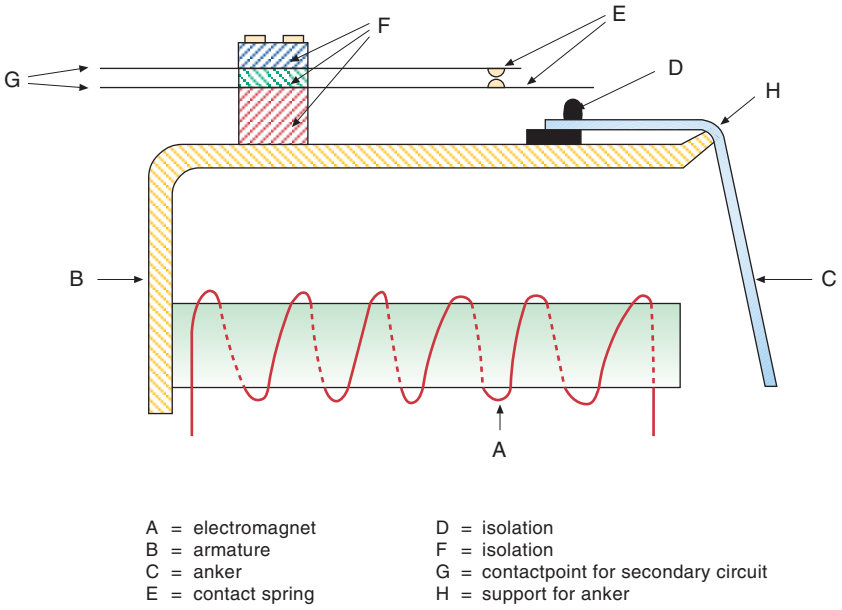


Figure 5.1 Electro-mechanical relay

5.1 Electro-mechanical elements

The switching in elements of the first main group, is achieved by some form of mechanical movement. Those were the first elements used in telecommunications and the group has grown to be large and somewhat heterogeneous. The elements are now being replaced by elements in the second group. In spite of the many drawbacks of these elements they have contributed enormously to the evolution of modern telecommunications and in fact, they are in use in large parts of today's network.

The basic component in these elements is the electromagnetic relay with a construction as shown in Figure 5.1. Electrical current through the coil gives the movement of the armature, *C*, and thereby the contacts, *E*. Another type is the reed relay which is used in the 10-C exchange system (Figure 5.2). Today, there are types which look much like a standard IC.

Operating time of a commercial relay is in the ms (millisecond) area. However, relays in Si (silicon) have been constructed with an operating time in the µs area.

All these selectors can be characterised by:

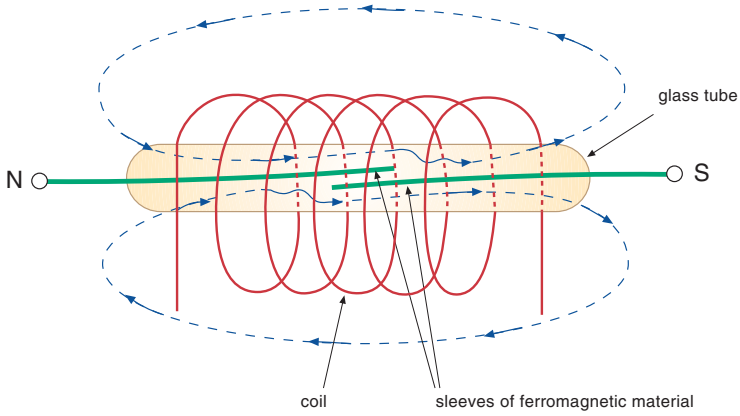


Figure 5.2 Reed relay

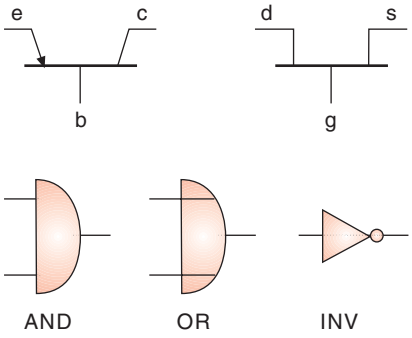


Figure 5.3 Basic electronic elements

- Low loss
- High crosstalk attenuation
- Can be used only in space switching
- Mechanical wear
- Rugged against overvoltage
- Need for manual maintainance
- Long lifetime
- Linearity
- No power consumption when un-operated
- Relatively large physical size
- Too long operating time to be used in time switching.

All these selectors have shown a surprisingly long life time (up to 50 years).

5.2 Electronic elements

These elements are based on semiconductor components and its development is still going on. The electronic elements enable the use of the TDM principle in switching.

Implemented electronic switching elements can be characterised by:

- Transmission parameters which are somewhat poorer than the electromechanical elements
- Much shorter operating time
- Small physical size
- Same technology in all units in a switching system
- The elements can be used in all the types of switching principle
- Long lifetime
- Need of special action to prevent the circuits from being damaged by overvoltage (EMP pulses)
- Possibility to integrate amplification in the elements
- No mechanical wear.

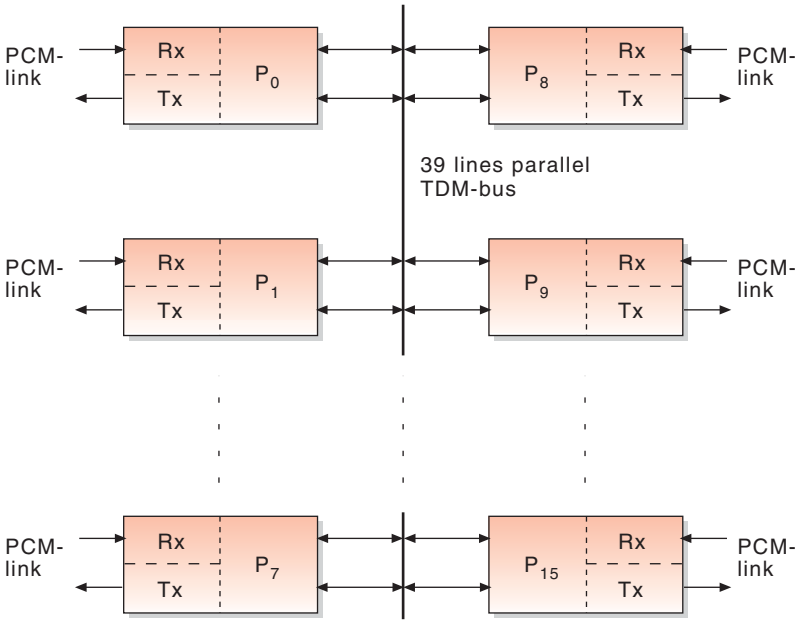


Figure 5.4 The digital switching unit in system ITT 1240

One basic component in electronic switches is the transistor (bipolar/unipolar) as depicted in Figure 5.3. A voltage at the base (b) makes the transistor conduct, thus making a small impedance. No voltage at the base gives a non-conducting state and a large impedance. The transistor is also the basic component in the logic circuits with the "AND" gate, "OR" gate and "INVERTER" as basic logical building components.

The logical components can be used to build switching networks and this supports using digital representation of data. It is also very important that the semiconductor technology (VLSI) has given cheap means to store data.

In digital switching network special switching units have been developed to take better advantage of the semiconductor integration. An example of such a switch is given in Figure 5.4 which shows a simplified switching unit in system ITT 1240.

The unit consists of 16 switches with access to a TDM bus. Each switch contains a T-switch (see Section 3), an LSI (large scale integration chip with about 11,500 transistors), circuitry for communication with the bus, receiving circuit R_x and transmitting circuit T_x for communication with other units via PCM links.

5.3 Optical elements

Since the introduction of the optical fibre in the transmission part of the telecommunication networks, it has been desirable to use the same technology in the rest of the network. The reasons for this can briefly be summarised as:

- Same technology in the whole network
- No need for optical/electrical converting giving:
  - Lower cost
  - No degrading of transmission parameters because of the converting
- A better utilisation of the large bandwidth of optical components.

Optical technology demands some special components to implement a complete switching system. These are components to connect the single switches, switching matrices or other units together, to connect optical effect in and out, terminate fibre and other components, etc.

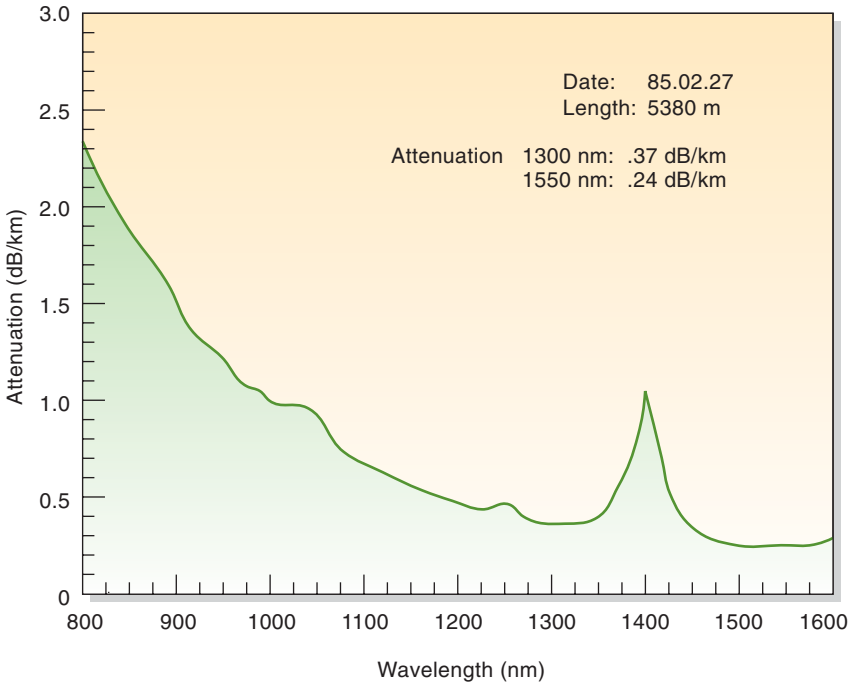


Figure 5.5 Spectral attenuation of a commercial mono-mode fibre

The optical components can be classified in two main groups:

- Passive components
- Active components.

The first group consists of (linear/non-linear) components not changing the optical signal (transport) by using external energy.

The second group consists of components which change some optical parameters, generate or detect optical power by using some form of external energy like electrical, magnetical, acoustical, thermical or optical energy or a combination of two or more of these.

The components can also be divided into two structural groups:

- Components with periodic structure
- Components with non-periodic structure.

The structure of active components could be both periodic and non-periodic.

5.3.1 Passive optical components

Examples of passive components with non-periodic structure are:

- The fibre and integrated wave guides
- Connectors
- Fan-out/fan-in elements (splitters/combiners)
- Lenses
- Terminators/absorbers.

The best known optical component is the (optical) fibre which is now in increasing use in all modern telecommunication networks. With respect to transmission this means access to an economical and attractive transmission medium with a very large bandwidth.

An optical fibre can be characterised by:

- Loss down to 0.2 dB/km
- Polarisation dependent
- Dispersion
- Immunity for electrical disturbances
- Very large bandwidth.

Figure 5.5 shows the spectral attenuation of a mono-mode fibre.

The bandwidth given by two wavelengths λ_{i+1} and λ_i can be expressed as:

$$\Delta f = f_{i+1} - f_i = (c / n)\,(1 / \lambda_{i+1} - 1 / \lambda_i)$$

where

c = speed of the light (in vacuum)

n = refraction index of the fibre.

$\lambda_{i+1} = 1.0$ µm, $\lambda_i = 1.3$ µm, $c = 3.10^8$ m/s and $n = 1.5$ give:

$$\Delta f \approx 46 \text{ THz} = 46 \cdot 10^{12} \text{ Hz}$$

This illustrates the very large bandwidth of the fibre.

The power to detect a bit will limit the usable bandwidth since one has the following (simplified) relation between bandwidth and power:

$B = P_0 / E$
P_0 = input power, E = energy to detect 1 bit

With $P_0 = 1$ mW and $E = 500$ photons/bit, this gives:

B = ca 10 Tb/s = 10^{12} b/s.

Effective bandwidth is reduced but still very large compared to the metallic cables in use.

"Integrated" waveguides, made by doping a substrate material with another suitable material, have about the same transmission parameters as the fibre.

Connectors to connect fibres have been developed and are characterised by:

- A typical loss of 0.1 – 0.2 dB
- Possibility for connecting both mono-mode and multi-mode fibres.

Fibres could also be connected by welding them together. This gives a connection with lower loss.

Combiners and splitters can be characterised by:

- Wavelength dependence
- Coupling loss
- Loss due to polarisation
- Coupling ratio
- Non-uniform distribution of power amongst the outlets
- Crosstalk due to the amount of input power connected back to the inlet
- Number of inlets and outlets (in practice restricted to about 32 x 32).

In a splitter the input power is distributed more or less uniformly to the outlets:

$$p_o = P_0 / N \pm \Delta p$$

where

P_0 = power into an inlet

Δp = variation in power distribution

N = number of outlets.

A combiner has the opposite function of a splitter, i.e. to multiplex power from more inlets to a common fibre, waveguide, etc. Splitters can also be used the other way around to give combiners. For a mono-mode combiner with N inlets, the "constant radiance" theorem gives the output power due to one inlet p_0:

$$p_{out} \leq p_0 / N$$

and total power:

$$p_{out} = \sum_{i=1}^{i=N} p_i$$

The above gives for fan-in/fan-out components an effective bitrate:

$$b = P_0 / N \cdot E$$

where

P_0 = input power

N = number of inlets

E = energy to detect 1 bit.

Some methods to realise splitters are shown in Figure 5.6. The most relevant components are fibre and integrated splitters.

A fibre splitter can be realised by:

- "Placing" the fibres close enough so that the optical fields in the fibres interact
- Fusing the fibres together so that the optical fields in the fibres interact.

A splitter/combiner with the same number of inlets and outlets is known as a "star coupler".

In Figure 5.7 is shown how a star coupler can be realised by connecting more 2 x 2 splitters/combiners. Integrated splitters can also be realised by making waveguides on a suitable substrate.

As indicated above a customary lens can be used to get fan-in and fan-out functions. Lens function can also be made by changing the diffraction index or geometry in a suitable way as shown in Figure 5.8. The change in diffraction is achieved by diffusing a material to increase the diffraction index such that the wavefront is disturbed in the right way (Figure 5.8a). In Figure 5.8b the geometry of the waveguide is changed giving a so-called "geodetic" lens which is achromatic.

One technique which better lends itself to integrated technology, however, is to use (optical) structures or periodic gratings. This can be used:

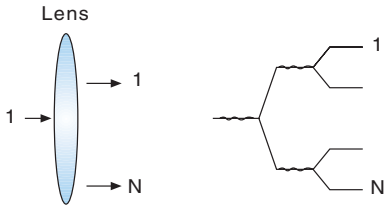


Figure 5.6 Fan-in/fan-out (splitter/combiner) principles

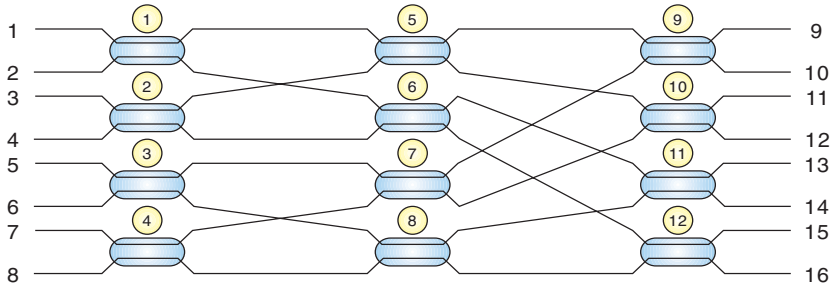


Figure 5.7 Example of an 8 x 8 star coupler based on "fused" fibres

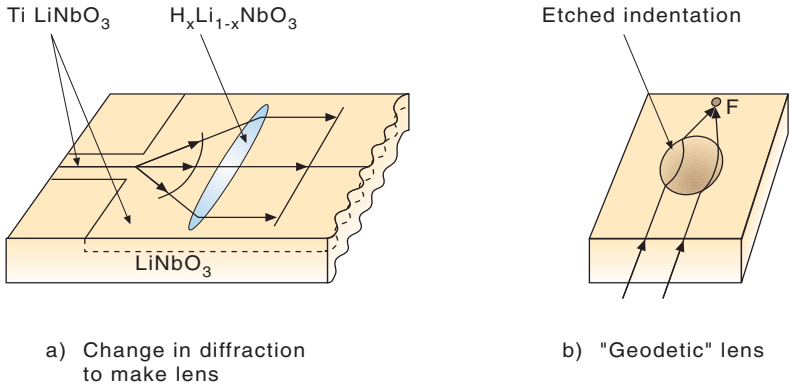


Figure 5.8 Possible integrated lens structures

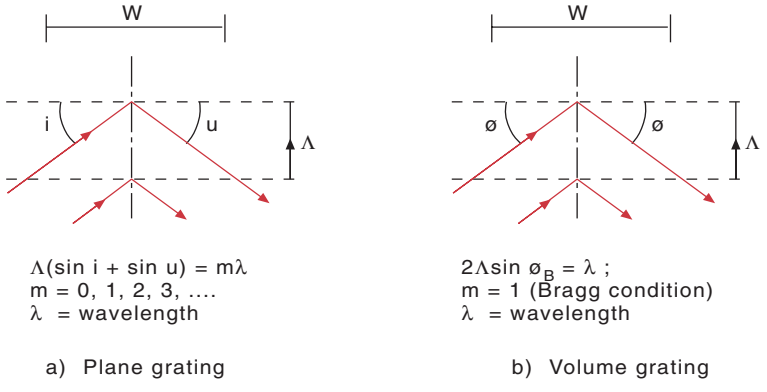


Figure 5.9 Methods of grating refraction

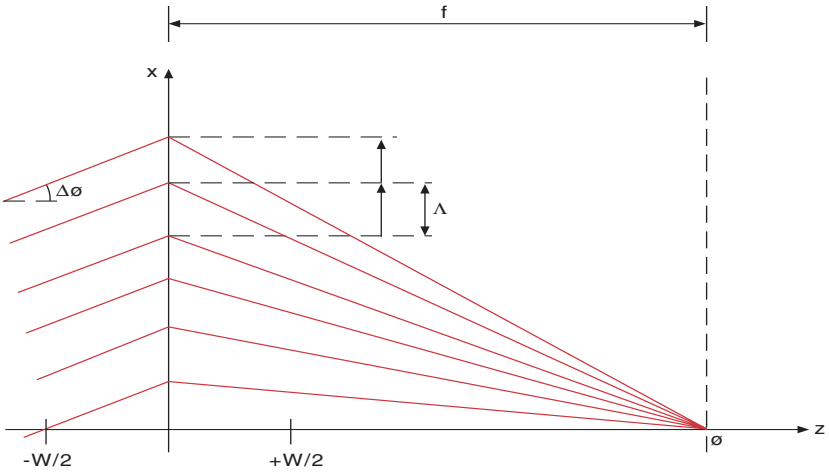


Figure 5.10 Variation of Δφ and Λ to make a grating lens

- In phase converting:
  - To connect or disconnect optical power
  - As reflectors, terminators, polarisators
  - As filters
- In wavefront converting:
  - To execute lens functions
- In wavelength dispersion:
  - For multiplexing.

Grating structures have many advantages:

- They can be implemented in many materials.
- They can be implemented with the same tolerances as in semiconductor technology.
- Other optical components can be integrated on the same substrate.
- They use processes like those used to make waveguides.
- Important parameters are relatively insensitive to failure in the grating.

The structures can be "passive" or "active", the latter being important components to get tuneable filters and lasers.

All types (transmission/reflection/absorption) structures are based on plane grating (Figure 5.9a), or volume grating (Figure 5.9b).

Figure 5.10 illustrates how wavefront converting can be used to get lens function by varying the parameters in the first order grating equation (Figure 5.9a).

Grating with varying grating distances are called chirped gratings (Figure 5.11).

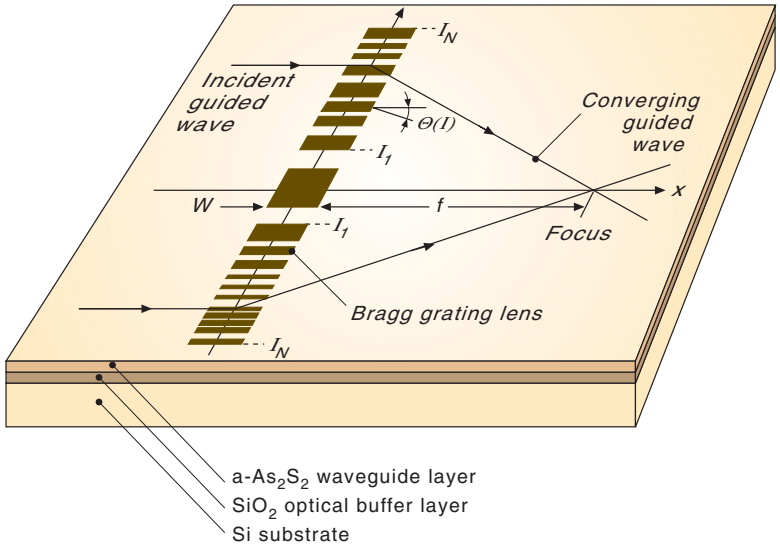


Figure 5.11 Chirped grating lens

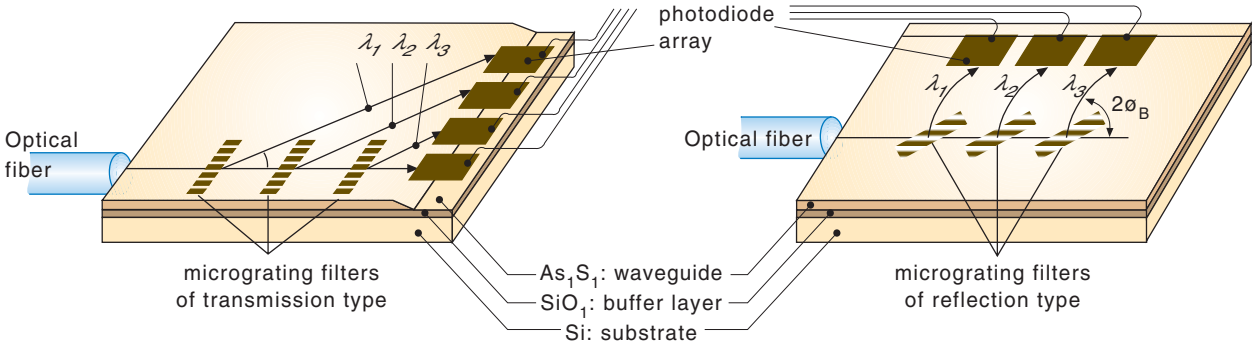


Figure 5.12 Wavelength demultiplexing by use of passive gratings

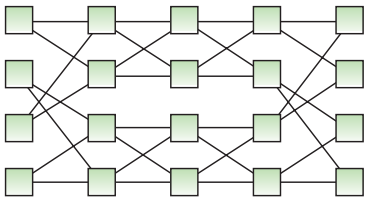


Figure 5.13 Multistage network realised by β-elements

The grating is made by changing the refraction index at the grating places.

Periodic structures can also be used to make splitters and combiners, i.e. to connect or disconnect optical power.

An important use of periodic structures is to make wavelength multiplexers and filters since the gratings are wavelength dependent (see Bragg condition in Figure 5.9b). Figure 5.12 shows some (passive) implemented gratings for wavelength demultiplexing purposes.

Even if the connections of optical fibres by using connectors or welding today are routine operations, the connection of units with a large number of inlets and outlets is not a trivial problem. Such a problem arises when more switching units are to be connected to make larger units. The number of inlets and outlets will soon be large. As an example, the switching unit in Figure 5.13 implemented by β-elements or optical direction couplers (see later chapters) will need a number of links between each stage given by:

$L = N\,(2 \log_2 N - 2)$;
N = number of inlets.

Example:

$N = 1024$ gives: $L = 1024 \cdot 18 = 18{,}432$

($N = 1024$ is small in public switching networks.)

Many methods have been proposed and tried to overcome this serious problem. Two of the most promising methods are to use the shuffle and the crossover principle to connect the stages. The functional behaviour of the switching network will be the same with both methods. Optical shuffling can be realised by using bulk optics (lenses, prisms).

Other methods are holography and the use of arrays of microlenses. Microlenses can be made at the end of a fibre by using a melting process.

Table 5.1 Parameters of some tuneable lasers

Configuration	Method of tuning	Range of tuning	Line width	
External grating cavity	Mechanical	55 nm	10 kHz	Step-by-step tuning
Electro-optical	Electronic	7 nm	60 kHz	
Acoustic optical	Electronic	70 nm		
2-sections DFB	Electronic	3.3 nm	15 MHz	Continuous tuning
3-sections DFB	Electronic	2 nm	500 kHz	
3-sections DBR	Electronic	8 – 10 nm 4.4 nm	20 – 100 MHz 1.9 MHz	Step-by-step/ continuous tuning
PIC laser	Electronic	21 nm		

5.3.2 Active optical components

Examples of active optical components are:

- Lasers
- Detectors
- Filters
- Switches/modulators
- Logic
- Memory
- Amplifiers.

Important parameters of a laser will be:

- Stability with respect to wavelength and power
- Tolerances
- Lifetime
- Linearity.

If the laser is tuneable, more parameters are also important:

- The number of wavelengths the (tuneable) laser is able to produce
- Distances between the wavelengths
- The power distribution of the wavelengths.

A tuneable laser is a laser which can be tuned to several wavelengths by using electrical power or by some other means. The development of tuneable lasers with a large number of wavelengths will have a major impact on the future networks. Table 5.1 shows achieved parameters for some tuneable lasers.

The tuning range of the lasers seems to be quite small. However, one has to take into account that 1 nm is equivalent to a bandwidth of about 180 GHz. When used in multi-wavelength systems a critical parameter will be the shortest distance between the wavelengths. Within the range of wavelengths of the laser, this will give an upper limit on the number of wavelengths. Output power is sufficient for most applications in local and urban transmission networks.

The detectors make it possible to detect levels down to -40 dB and coherent detection makes it possible to detect levels down to near the quantum limit.

The output power from the laser and the detection level of the detector are key parameters in the effect budget in an optical system.

Tuneable optical filters (periodic structures) can be made by using acoustical or electrical power to change the parameters. Usage of an acoustical surface wave to make a tuneable grating is depicted in Figure 5.14. A surface wave from the transducer SAW compresses and decompresses the substrate and thus makes a grating. The distance between the grating lines is determined by the acoustic wavelength. By using acoustic frequency and angle of incidence satisfying the Bragg condition, the optical power is deflected to the right position. A grating, set up by using an electrical grating, is depicted in Figure 5.15.

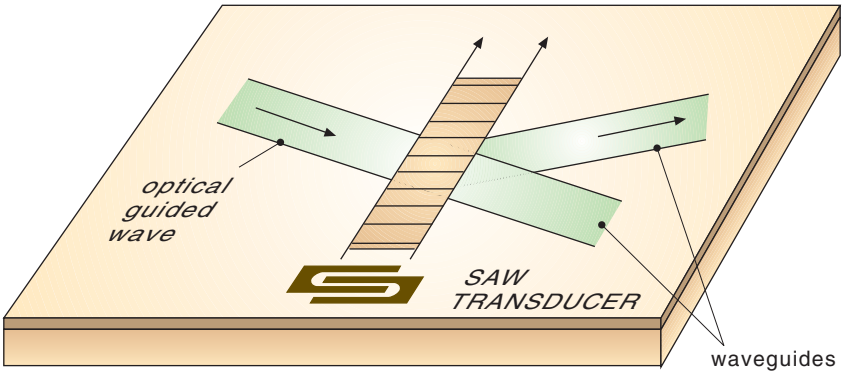


Figure 5.14 Grating with acoustic control

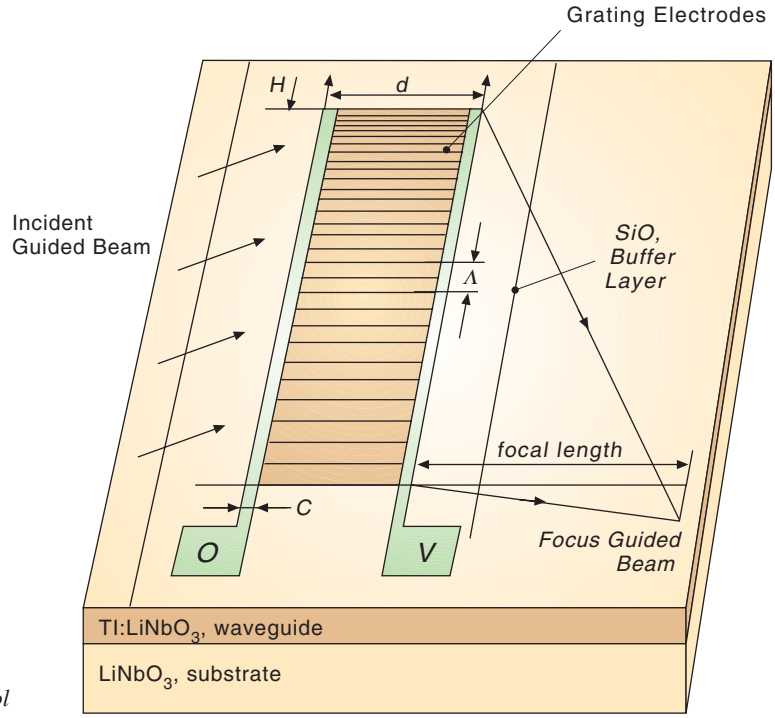


Figure 5.15 Grating with electrical control

Table 5.2 Some implemented opto-mechanical switches

Technology	Size	Crosstalk dB	Loss dB	Switching time ms
Movement of pentaprisms	Up to 10 x 10	60	0.7–1.3	40
V-grooves in metal	1 x 93	60	1.4	2–200
V-grooves in silicon	1 x 8	60	0.2	
Reed-relay type	1 x 2	50	1	10
Micromovement in a tube	1 x 4	60	0.25	
Electrowetting	1 x 2	23-51	0.5-2	20

The last two components are examples of what is called SLM – Spatial Light Modulator – which is used in many applications to deflect or steer light in a preferred direction.

Terminators and absorbers are used to absorb unwanted optical power.

With respect to switching the optical switch is of course a very important component. The optical switches can be classified in the following main groups:

- Opto-mechanical switches
- Solid state optical switches
- Electro-optical switches.

The first group consists of elements based on movement as illustrated in Figure 5.16a. A fibre is moved until axial alignment with one of the fixed fibres. Examples of realised optical switches are shown in Figure 5.16b–c, and Table 5.2 shows obtained parameters of realised switches.

The mechanical switches can be used both for mono- and multi-mode fibres, and another favourable property of these elements is that they are bilateral and independent of wavelength and polarisation. Because of the long switching times they will not be suitable in a TDM system.

Elements in the second main group – the solid state group – can be divided into two subgroups:

- Elements without moving parts and based on using optical properties in certain crystals
- Thin film elements based on optical properties in some crystals.

The optical properties used in both groups are:

- The electro-optic effect
- The magneto-optic effect
- The acoustic-optic effect
- The thermo-optic effect.

All these effects change the refraction coefficient of the material and this is the base of the function.

Thermo switches use the temperature to change the refraction index while polarisation switches use the change of direction of the polarisation in certain crystals when a voltage is applied. It is supposed

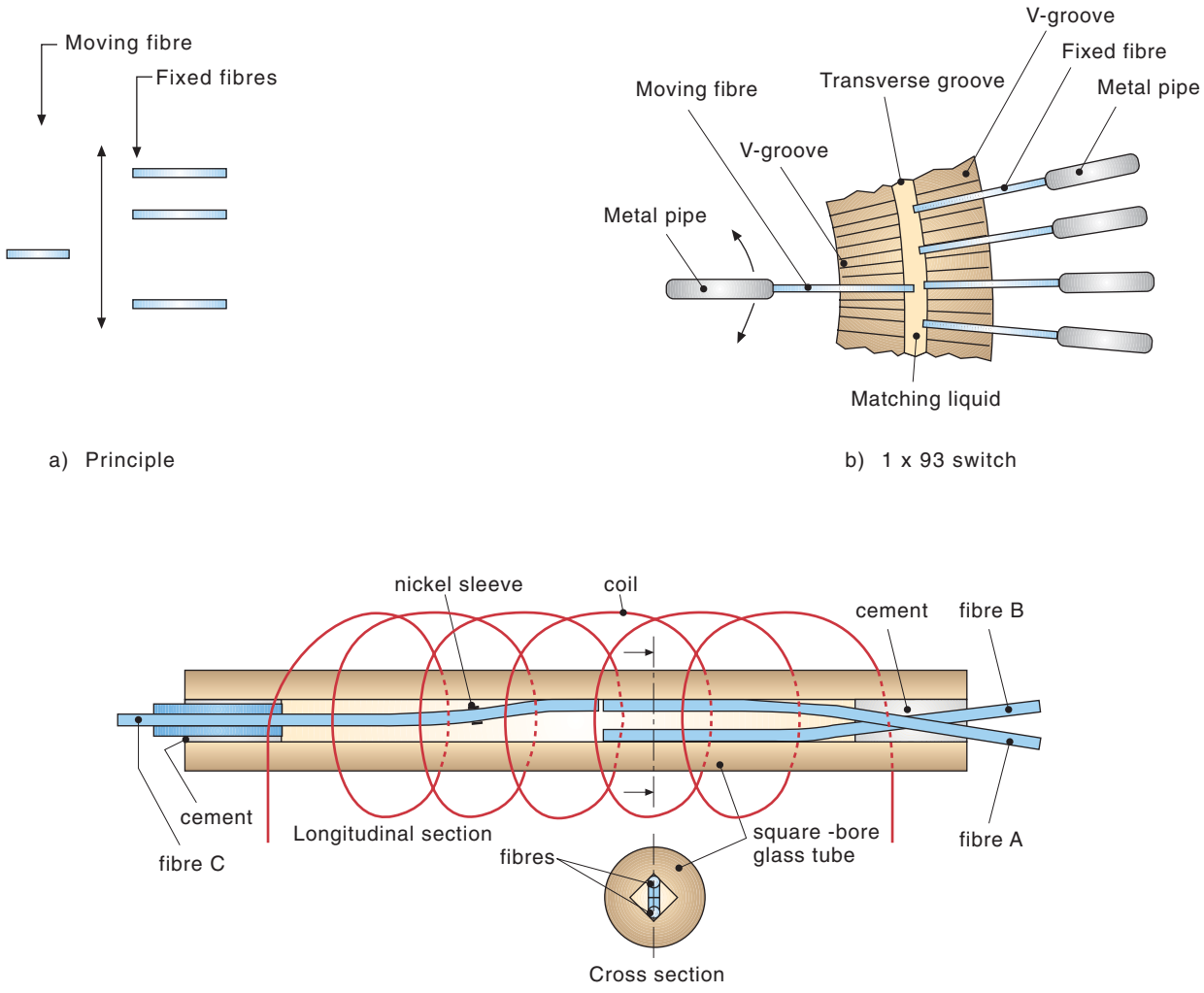


Figure 5.16 Opto-mechanical switches

that none of these types of switches be used in future switching networks.

The most important group of solid state switches is the integrated switches. These elements are implemented in thin film technology by making waveguides on a substrate with lower refraction index. In addition, optical properties of the waveguide material must be changeable in a suitable way. Other optical components like splitters, combiners, lenses, etc. can be implemented by this technology.

The switching is based on the following principles:

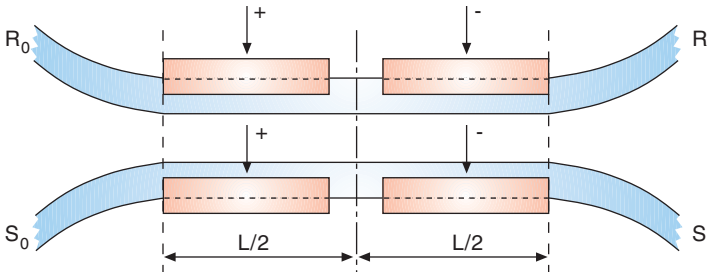
- Two (possibly more) waveguides are placed so close that the optical fields overlap each other and the degree of this overlapping can be changed by an outer stimulus
- The fields in two (or more) waveguides are interfering with each other in a common region such that the optical power can be steered to different outlets by using an outer stimulus.

Figure 5.17 illustrates two switches based on these two principles. Figure 5.17a shows the “directional coupler”. This type is based on the fact that optical power can be transferred from one waveguide to another if they are so close to each other that the fields interact. The switch can be designed such that with zero voltage the power into one waveguide is transferred to the other waveguide. With a difference in the voltage levels the refraction index is changed such that no (optical) power in the waveguide is transferred to the other waveguide and leaves at the outlet of the same waveguide.

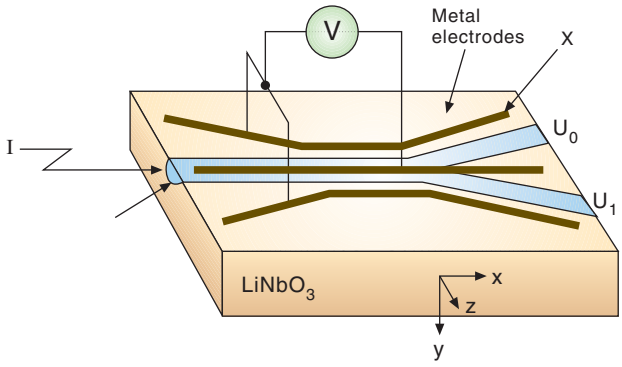
In Figure 5.17b the power is steered more directly to the output by changing the refraction index in the region where the incoming waveguide is split. This type is often called a “Y-coupler”.

The advantages of the Y-coupler compared to the directional coupler are briefly

- Shorter switch (less place in transmission direction)



a) Directional coupler



b) Y-coupler

Figure 5.17 Two basic principles of integrated switches

- The demand to tolerances is somewhat less.

The main drawback is the higher voltage needed in order to perform the switching.

Common properties of optical switching elements for the two types above are:

- Very large bandwidth
- Higher loss than the counterparts (electromechanical and electronic)
- Greater physical size compared to electronic switches
- Relative low crosstalk
- Low power consumption.

Another drawback is that they (normally) can be used with only mono-mode signals.

Figure 5.18 shows an acousto-optical switch. This is based on the principle that an acoustic wave in a crystal works as a grating, see above. This grating can then be used to deflect the incoming light.

Another group of switching elements is based on converting optical power to electrical power, switching and then converting back to optical power. The switch can be realised by commercial components (lasers, detectors). It is also possible to implement such switches by integration of laser and detector and also to integrate amplification in the same chip. This, of course, makes it possible to make switches with none of the drawbacks mentioned above for the directional and Y-coupler switches.

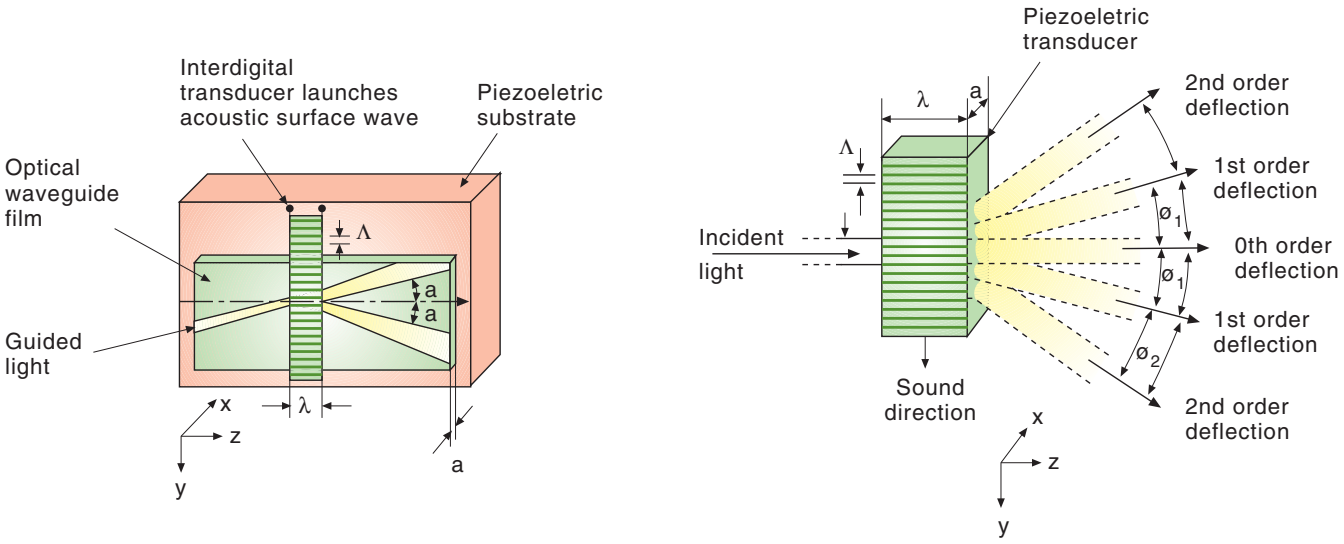


Figure 5.18 Acousto-optical switch

The development of optical switches has not reached a level where it is possible to use them on a large scale. However, this is expected to be solved in a matter of time.

The most common substrate $LiNbO_3$ limits the size of switching matrices to about 16 x 16. Larger switches have to be made by connecting more such matrices.

The best switches have crosstalk parameters of about 30 dB and losses 3 – 5 dB. The crosstalk is acceptable for switching of digital data but not for switching of analogue data. Applying a number of switches in serial also demands amplification. For control functions it is necessary to have logical elements such as AND, OR and NEGATION gates. The OR function can be executed by a combiner. To execute the AND and the NEGATION functions active components have to be used. It is supposed that usage of lasers will be important in these components.

It is also necessary to have components to store data. The lack of practical memory elements is a serious drawback for making use of all advantages of optical technology in the telecommunication network. Today, the actual memory components are:

- Fibre delay element
- Element based on combination of laser and detector.

The first type is characterised by:

- Storing by using delay in a fibre
- Both analogue and digital data can be stored
- Need for amplification to compensate loss
- Storing capacity determined by length of the fibre
- Access time in principle depending on total delay time
- Simple structure.

The second type gives the possibility to make optical flip-flop which:

- Needs no refreshing of stored data
- Has short (and constant) access time
- Lends itself to integration of larger memories.

Optical amplifiers are also essential in order to implement an optical telecommunication network. They will compensate for the losses in various parts of the network. Usage of lasers and detectors are seen to be the most promising type. In later years the "fibre amplifier" has reached a commercial level, yet it is not a cheap component. This type is based on doping a fibre with erbium and "pumping" the erbium ions with a shorter wavelength. This "pumping" lifts the ions to a higher level and when the signal light is passing, the ions return to their prior level giving the "pump energy" to the light ions and thus amplify the signal.

Bibliography

1 Arrathoon, R, Hassoum, M H. Optical threshold logic elements for digital computation. *Optics Letters,* 10(4), 1984.

2 Delavaux, J M P, Betts, G E, Chang, W S C. Electrooptic linear chirped grating lenses on planar optical Ti:$LiNbO_3$ waveguides. *Journal of Lightwave Technology,* LT-3(3), 1985.

3 Digonnet, M, Shaw, H J. Wavelength multiplexing in single mode fiber couplers. *Applied Optics,* 22(3), 1983.

4 Forouhar, S, Chang, W S C, Yao, S-K. Performance and limitations of chirped grating lenses of Ti-indiff-fused $LiNbO_3$ planar waveguides. *Journal of Lightwave Technology,* LT-2(4), 1984.

5 Gaylord, T K, Moharam, M G. Analysis and applications of optical diffraction by gratings. *Proceedings of the IEEE,* 73(5), 1985.

6 Gibbs, H M. *Optical bistability : controlling light with light.* New York, Academic Press, 1985.

7 Goodman, J W. Fan-in and fan-out with optical interconnections. *Optica Acta,* 32(12), 1489–1496, 1985.

8 Goodwin, M, Stewart, C. Proton-exchanged optical waveguides in Y-cut lithium niobate. *Electronics Letters,* 19(6), 1983.

9 Haus, H A, Whitaker, N A Jr. Elimination of cross talk in optical directional couplers. *Applied Physics Letter,* 46(1), 1985.

10 Hinton, H S. Photonic switching using directioinal couplers. *IEEE Communications Magazine,* 25(5), 16–26, 1985.

11 Hunsperger. *Integrated optics : theory and technology.* Springer, 1982.

12 Jackel, J L, Rice, C E, Veselka, J J. Proton exchange for high-index waveguides in $LiNbO_3$. *Applied Physics Letter,* 41(7), 1982.

13 Jahns, J, Murdocca, M J. Crossover networks and their optical implementation. *Applied Optics,* 27, 3155, 1988.

14 Kahn, A H, Nejib, U R. Optical logic gates employing liquid crystal optical switches. *Applied Optics,* 26, 270, 1987.

15 Kilvik, K. Optiske komponenter i fremtidig automatikk. *Telektronikk,* 75(3), 1979. (In Norwegian.)

16 Kogelnik, H. Coupled wave theory for thick hologram gratings. *Bell System Technical Journal,* 48(9), 1969.

17 Kogelnik, H. *Review of integrated optics.* Holmdel, New Jersey, Bell Telephone Laboratories.

18 Lea, C. A directional-coupler-based broadcast switching network. *Globecom '87,* 3, 2044–2048, 1987.

19 Lentine, A L et al. The symmetric self electrooptic effect device. *Conference on Lasers and Electrooptics, Technical Digest Series.* Washington, DC, Optical Society of America, 249, 1987.

20 Mandel, P, Smith, S D, Wherrett, B S (eds.). *From optical bistability towards optical computing : the EJOB project.* The Netherlands, North Holland, 1987.

21 Marcatili, E A J. Dielectric rectangular waveguide and directional coupler for integrated optics. *Bell System Technical Journal,* 2071–2102, Sept. 1969.

22 Midwinter, J E, Smith, P W (eds.). Photonic switching. *IEEE Journal on Selected Areas in Communications,* SAC-6(7), 1988.

23 Miller, D A B et al. The quantum well self-electrooptic effect device : optoelectronic bistability and oscillation and self-linearized modulation. *IEEE Journal on Quantum Electronics,* QE-21, 1462, 1985.

24 Miller, S E. Integrated optics : an introduction. *Bell System Technical Journal,* 2059–2069, Sept. 1969.

25 Mortimore, D B. Wavelength-flattened 8 x 8 single-mode star coupler. *Electronics Letter,* 22(22), 1986.

26 Murphy, E J, Rice, T C. Low-loss coupling of multiple fibre arrays to single-mode waveguides. *Journal of Lightwave Technology,* LT-1(3), 1983.

27 Murdocca, M J et al. Optical design of programmable logic arrays. *Applied Optics,* 27, 1651, 1988.

28 Murdocca, M. *A digital design methodology for optical computing.* The MIT Press, 1990.

29 Ogawa, Y, Ito, H, Inaba, H. Bistable optical device using a light emitting diode. *Applied Optics,* 21(11), 1982.

30 Okumura, K et al. Optical logic inverter and elements using laser or light-emitting diodes and photo detectors in bistable system. *Optics Letters,* 9(11), 1984.

31 Prise, M E et al. Module for optical logic circuits using symmetric self-electrooptic effect devices. *Applied Optics,* 29, 2164, 1990.

32 *Proceedings IEEE 72,* 758–974, 1984. (Special issue on Optical Computing.)

33 *Proceedings of the Royal Society's Topical Meeting on Optical Bistability, Dynamic Nonlinearities and Photonic Logic, Phil. Trans. Roy. Soc. Lond.,* A313, 187, 1984.

34 Prucnal, P, Blumenthal, D, Perrier, P. Photonic switch with optically self-routed bit switching. *IEEE Communications Magazine,* 25(5), 50–55, 1987.

35 Reviewed in Peyghambarian, N and Koch, S W. Semiconductor nonlinear materials. In: Gibbs, H M, Khitrova, G, Peyghambarian, N (eds.). *Nonlinear photonics,* 7–60. New York, Springer-Verlag, 1990.

36 Smith, P W. Hybrid bistable optical devices. *Optical engineering,* 19(4), 1980.

37 Smith, P W, Tomlinson, W J. Bistable optical devices promise subpicosecond switching. *IEEE Spectrum,* June 1981.

38 Smith, S D. Optical bistability, photonic logic and optical computation. *Applied Optics,* vol. 25, 1550–1564, May 15, 1986.

39 Sohler, W. Bauelemente der integrierten Optik : eine Einführung. *Laser und Optoelektronik,* 4, 1986. (In German.)

40 Sottini, S, Russo, V, Giorgetti, E. A new family of geodesic lenses for waveguide optical devices. *Journal of Lightwave Technology,* LT-5(8), 1987.

41 Suhara, T, Nishihara, H. Integrated optics components and devices using periodic structures. *IEEE Journal of Quantum Electronics,* QE-22(6), 1986.

42 Tamir, T. *Integrated optics.* Berlin, Springer Verlag, 1975.

43 Walker, A C, Wherrett, B S, Smith, S D. First implementation of optical digital computing circuits using nonlinear devices. In: Gibbs, H M, Khitrova, G, Peyghambarian, N (eds.). *Nonlinear Photonics,* New York, Springer Verlag, 91–149, 1990.

44 Warren, C, Forouhar, S, Chang, W S C. Double ion exchanged chirp grating lens in lithium niobate waveguides. *Applied Physics Letter,* 43(5), 1983.

45 Yariv, A, Nakamura, M. Periodic structures for integrated optics. *Journal of Quantum Electronics,* QE-13(4), 1977.

6 Classification of switching networks

The main classification of switching networks is given by the switching principles mentioned previously.

In general, all types have to execute the same tasks, i.e. to establish connection between a set of inlet and outlet pairs. The main parameters as number of inlets and outlets, capacity measured in, for example, the number of simultaneous connections, and same congestion probability, can all be made equal or approximately equal for all of the switching principles. The representation of the data often determines the choice of switching principle. Digital data representation makes it natural to use time switching while analogue data representation can favour space switching. Of course this ought not to be a rule since space switches can be used for both representations.

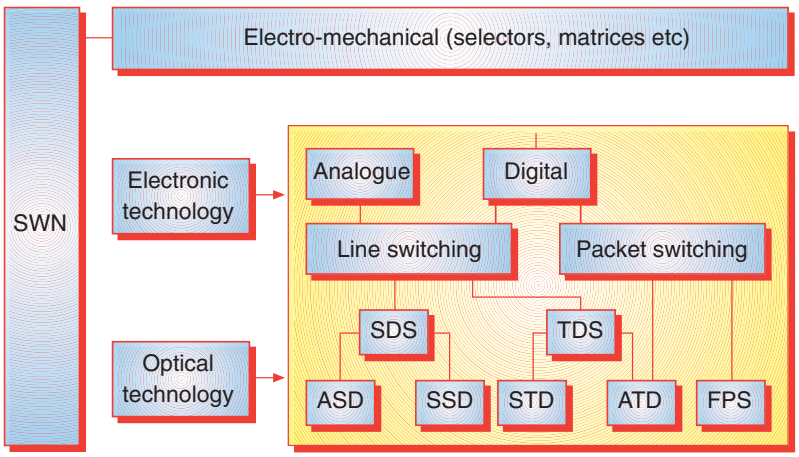
The classification will, in the following, be given for space and time switching networks separately.

6.1 Classification of space networks

Space switches in the form of selectors were the first type to be developed – the patent of Strowger selector goes back to the 1890s – and this group has grown to a variety of types and sizes for different employment. One way to group these switching networks is based on the basic element in the network, i.e. relay, selector, step-by-step or motor driven, electronic switch, etc. Each type of element gives a network with special properties with respect to transmission, data representation, operating speed, etc. Such a technological classification is shown in Figure 6.1.

Another way of classification which gives a better view, with respect to the structural and functional properties, is shown in Figure 6.2. This classification gives a better possibility to compare the various switching networks. The grouping is made after

- Overall structure
- Topological geometry
- Switching qualities
- Basic switching unit employed in the network.



ASD = Asynchronous Space Division
ATD = Asynchronous Time Division
FPS = Fast Packet Switching

SDS = Space Division Switching
SSD = Synchronous Space Division
STD = Synchronous Time Division
TDS = Time Switching

Figure 6.1 Technological classification of space switching networks (SWNs)

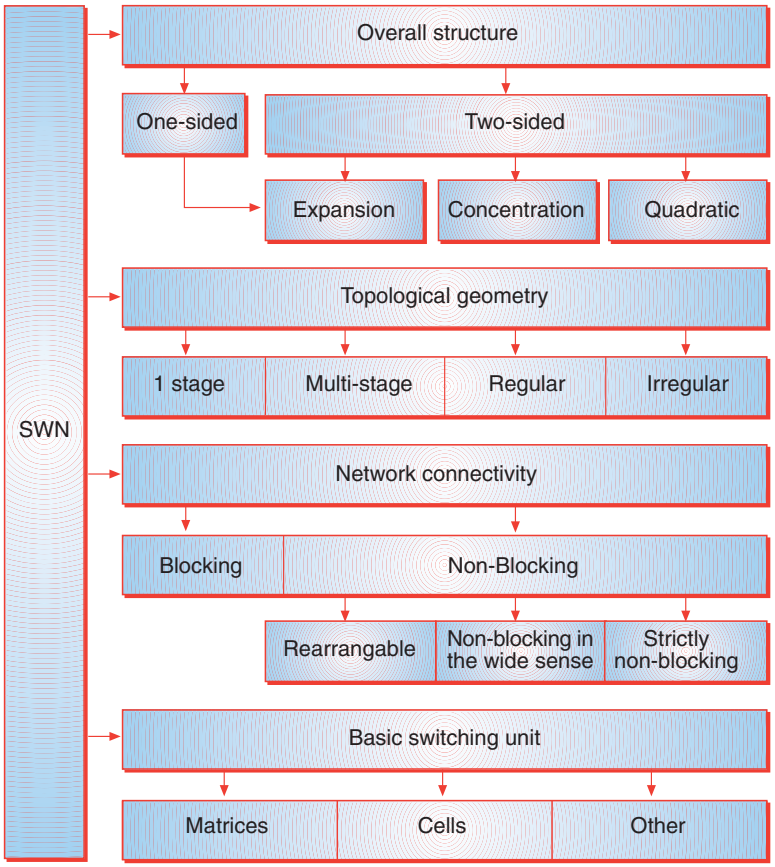


Figure 6.2 Classification of space switching networks

Overall structure tells what the whole switching network looks like seen from the outside or in a block-schematic way. The main grouping can be

- One-sided networks, i.e. networks with the terminals connected to one side (Figure 6.3)
- Two-sided networks, i.e. the terminals are connected to both sides.

A network in a "conventional" public exchange is illustrated in Figure 6.4 for two types – one-way networks and "reflected" networks. In the first type the outlet is connected back to the inlets with the same position. In the second type the call is reflected through the network back to the same side. In Figure 6.4a one has the following grouping:

- A concentrator part, *K*
- An expansion part, *E*
- A distribution part, *F*.

One-sided connection is illustrated in Figure 6.3. All terminals are equal so that a terminal can be both sender and receiver. The switching network must be of the bilateral type.

Two-sided switching networks are shown in Figure 6.5. The terminals are divided into two groups connected on each side of the network. The switching network in a public exchange is an example of such a network. Such networks could also be viewed as one-sided because of the two groups and since the terminals of the output group are connected back to the corresponding terminals in the input group. So the distinction between the two groups is rather unclear.

With a notion as in Figure 6.5 we can introduce a quantity f defined by:

$$f = M / N$$

This gives:

$f < 1$ means concentration, *K*
$f = 1$ means square network
$f > 1$ means expansion, *E*.

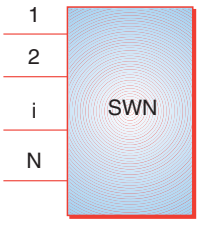
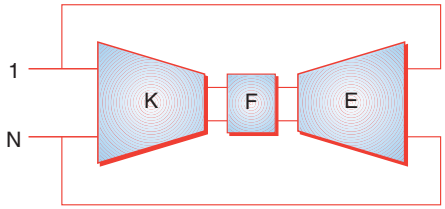
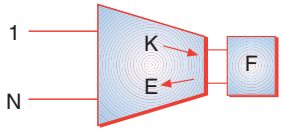


Figure 6.3 One-sided switching network



a) Unilateral switching network



b) Bilateral switching network

Figure 6.4 Switching network for public exchanges

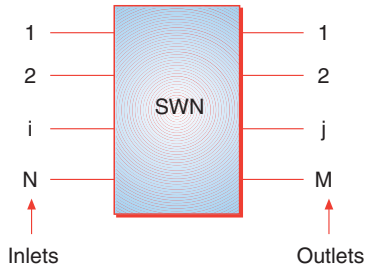


Figure 6.5 Two-sided space switching network

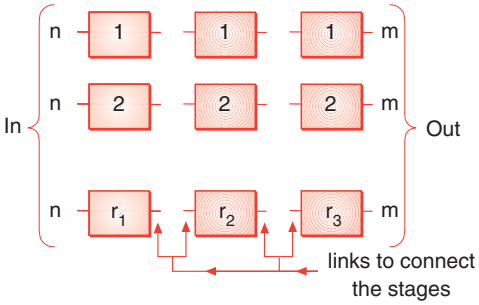


Figure 6.6 3-stage switching network

In telecommunication networks these structures are important. Basic for the concentration/expansion is that the number of simultaneous connections is much smaller than the number of inlets and outlets. In a practical case the figure $f = 0.1$ for concentration (i.e. $f = 10$ for expansion) and with $N = 10{,}000$, in Figure 6.4a, this gives (without the distribution part) the number of switching elements per inlet:

$n_i = 10000 \cdot 1000 \cdot 2/10000$
$= 2000$ with concentration/expansion

$n_i = 10000 \cdot 10000/10000$
$= 10000$ without concentration/expansion

This is a considerable reduction and a yet greater reduction is achieved by using more stages in the switching network.

The next main group in the classification scheme is based on the topological structure of the network, i.e. the parts of the network and the connection of these parts. One can have the following parts:

- One-stage network
- Multi-stage network
- Regular network
- Irregular network.

A one-stage network can be represented by a matrix of size N x M with one switching element for every combination of inlet/outlet pair. This gives:

$n_k = N \cdot M$ = total number of switches

$n_i = n_k / N = M$ = number of switches per inlet.

$N = M = 20000$ in public exchanges can be a reasonable size which gives:

$n_k = 400 \cdot 10^6$ = total number of switches

$n_i = 20 \cdot 10^3$ = number of switches per inlet.

To reduce the amount of switches it is common to use multi-stage networks as illustrated in Figure 6.6 (3-stage network).

A drawback of multi-stage network is the more complicated control of the network and the increasing delay will in some circumstances also be a drawback.

A network can also be classified as regular and irregular groups. The definition is based on whether there are vertical or horizontal symmetry lines or not. Symmetry lines defines a regular network while lack of such lines defines an irregular network. Another view is to look at the length of the connection, i.e. the number of switches in the connection. A regular network is said to have the same lengths while the irregular network have different lengths.

Another important way of classifying a network is to use the connection properties, i.e. the possibilities to connect inlet/outlet pairs. This gives two groups:

- Network without blocking, i.e. a connection can always be established
- Network with blocking, i.e. a connection cannot always be established.

Network in the first group can always establish a specified connection and with respect to the definition of a switch this gives:

- The network can realise all assignments of inlets to an outlet such that the assignment is 1-1 and ON (see Section 2)
- All assignments can be implemented by disjunct connections.

An example of a non-blocking network is a switching matrix which can implement all assignments.

Non-blocking networks are an important group in switching systems. This group can be divided into three subgroups:

- Strictly non-blocking
- Non-blocking in the wide sense
- Rearrangeable.

A strictly non-blocking network can be defined as above. All assignments can be realised.

A network blocking in the wide sense is also able to establish all assignments but now with the restriction that the connections have to be established in a particular order (see later chapter). This order or rule varies for the different structures but in multi-stage networks, like the one in Figure 6.7b, the rule is to fill a middle matrix before the next middle matrix is taken into use.

A rearrangeable network has the same possibility with respect to assignments as the two others but the connections could be rearranged in order to avoid blocking.

Figure 6.7 shows some examples of this type of network. The main reason for these networks being that interesting is that the number of switching elements can be greatly reduced.

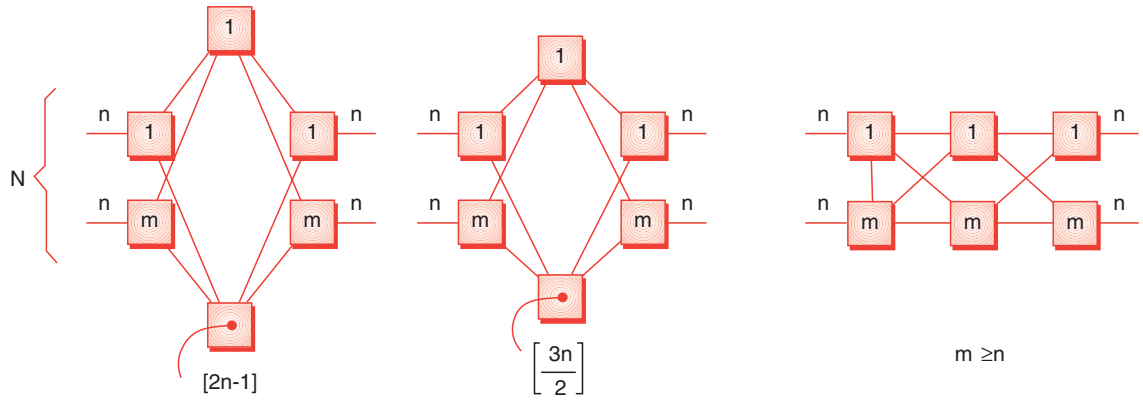


Figure 6.7 Example of a (3-stage) non-blocking network

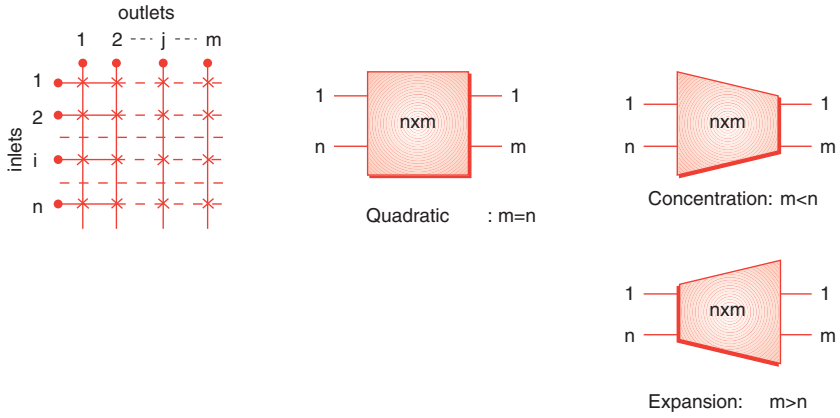


Figure 6.8 Two-sided matrices

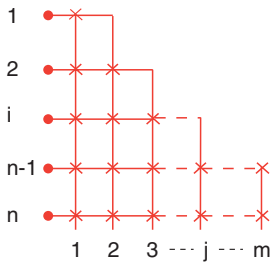


Figure 6.9 One-sided triangular network

As shown in Figure 6.7 the main condition that has to be fulfilled is that the number of middle matrices must have a minimum value. The total number of switching elements in the networks in Figure 6.7 is:

Network in Figure 6.7a:

$n_1 = [2(n(2n - 1)m)$
$+ m \cdot m(2n - 1)] / (nm)$
$= (2n - 1)(2n + m) / n$

Network in Figure 6.7b:

$$n_2 = 2(n \cdot \lfloor 3n/2 \rfloor m) + m \cdot m \lfloor 3n/2 \rfloor / (nm)$$
$$= \lfloor 3n/2 \rfloor (2n + m)/n$$

Network in Figure 6.7c:

$n_3 = [2 (n \cdot m \cdot m) + m \cdot m \cdot m] / (nm)$
$= m \cdot (2n + m) / n$

One can see that the three alternatives give decreasing number of elements, but there is also an increasing complexity of the control functions. A drawback of a rearrangeable network can be disturbances in transmission parameters when rearranging the connections.

Comparison between the number of elements for the three types above with the number of elements in a quadratic structure shows that a quadratic matrix is more favourable for N less than a certain limit:

Strictly non-blocking network:
$N \leq (2n - 1)(2 + m / n)$ (Figure 6.7a)

Non-blocking in the wide sense network:

$N \leq (2 + m/n)\lfloor 3n/2 \rfloor$ (Figure 6.7b)

Rearrangeable network:
$N \leq m \cdot (2 + m / n)$ (Figure 6.7c)

The size of the switch can also be used to group a network. The modern network can then be grouped in:

- Network based on matrices
- Network based on cells
- Network based on other types.

In principle, all networks based on matrices could be classified as a cell network. However, this name is here reserved for networks where each element has a restricted number of inlets and outlets and also has a reduced number of states (or reduced number of assignments) in each cell. An example of such a cell is the β-element shown in Figure 6.11. This element can be viewed as a 2 x 2-matrix but with only the two states indicated in the figure.

A β-element has the following functional properties:

- There is no OFF-state, i.e. there is always a connection between some inlet/outlet pair
- It needs only 1 control bit.

Figure 6.12 shows examples of switches based on β-elements.

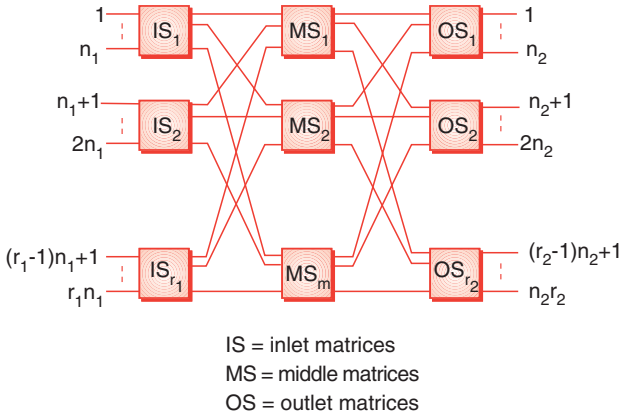


Figure 6.10 Example of a matrix configuration

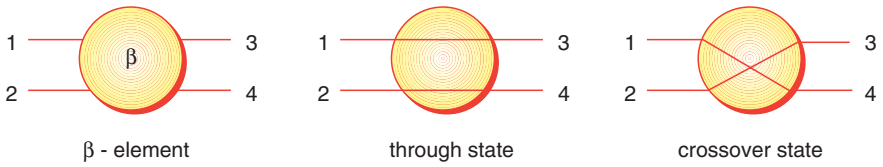


Figure 6.11 β-element

Number of elements can be measured in number of β-elements since the β-element can be viewed as an independent element. Figure 6.13 shows a large network using such elements. It uses 19 stages giving 9.5 β-elements per inlet.

Switching matrices with n inlets and m outlets are illustrated in Figure 6.8.

Figure 6.9 illustrates a one-sided triangular network and Figure 6.10 illustrates an example of a configuration with expansion from left to right and concentration in the other direction (bilaterally, switching matrices are assumed). Such a network is often denoted as:

$$v(m, n_1, r_1, n_2, r_2)$$

where:

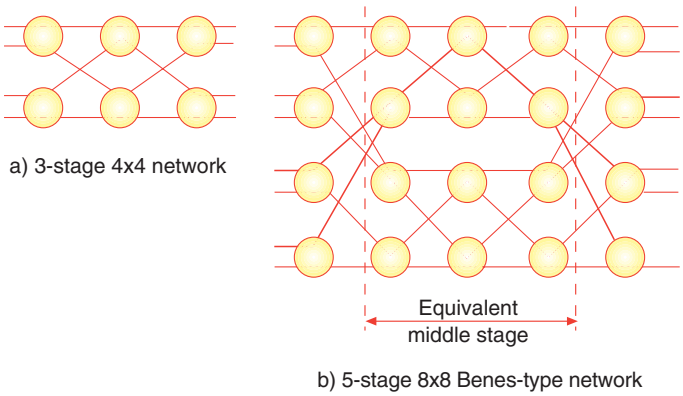


Figure 6.12 Examples of a network based on β-elements

r_1 = number of inlet matrices with size = (n_1 x m)

m = number of middle matrices with size = (r_1 x r_2)

r_2 = number of outlet matrices with size = (m x n_2).

A more ordinary matrix network is shown in Figure 6.14 where 2 x 2 switches are used.

Another type of cell elements is indicated in Figure 6.15. This element is called a Kreuz element and has a broadcasting state in addition to the two states in a β-element.

6.2 Classification of time switching networks

The time and space network are both able to realise the same assignment and can be classified in the same way. There are, however, some differences which make it more favourable to use a somewhat different classification.

The size of a time switching network (measured in number of inlets and outlets) has a much higher "lower limit". For example, when bitrates are low, there is no sense in implementing an 8 x 8 network in time switching. In fact, the building block of a time switching network for low bitrates will be much greater than this.

A possible classification of time switching network could be:

- Synchronous network
- Asynchronous network.

In a synchronous network there are fixed times between the events. The result from two successive samplings will arrive at the receiver (terminal) within the same time interval. Asynchronous networks do not have this restriction.

The difference between the two methods, however, is more functional than structural or technological, so such a division is not suitable.

Time switching can be divided into two basic functions:

- Switching in a distinct time (phase)
- Switching to a distinct terminal (user).

The first function is designated pure time switching or T-switching for short, and consists of changing the time sequence of the samples.

The second function can be designated time divided space switching and con-

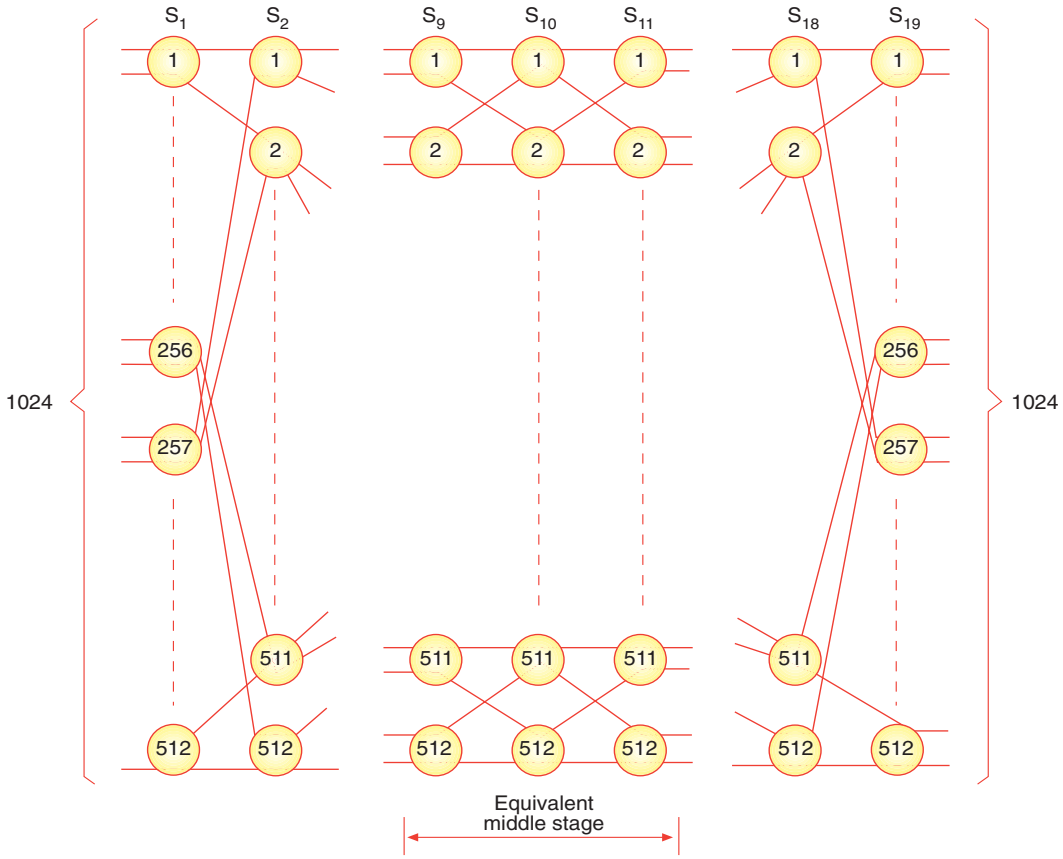


Figure 6.13 A 1024 x 1024 network based on β-elements

sists of distributing the samples to the right space position. The network is here called an S-network to distinguish it from a space network where time is not involved in the same manner.

Since, in a real network, both types could be present, it is relevant to make a division into:

- Networks based on T-switching
- Networks based on a combination of S- and T-switches (T-S-T or S-T-S).

In Section 2 an example is given of a time switch consisting of a multiplexer and a demultiplexer, one or both being programmable. This is illustrated in Figure 6.16. The data are transmitted on a common medium. If this point is "stretched" to a line, it can be viewed as a bus on which data are transported in a fixed time sequence. Such a bus is often designated as "fixed coupled bus" or a "pure transmission bus". The whole system in Figure 6.16 is called a "switching bus" and can be classified as a group of switching. It is, however, difficult to set a limit of what can be called a bus, so the designation is therefore reserved to real buses with a significant distance between functional units generating/receiving data.

The above gives a classification scheme shown in Figure 6.17. Comparison with the scheme in Figure 6.2 shows that there is no classification following basic switching unit. A reasonable equivalent size of a matrix in time switching is a 32 x 32 instead of a 2 x 2 matrix. Such a format could of course also be called a cell.

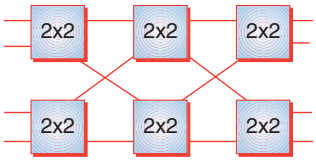


Figure 6.14 Benes-type network based on 2 x 2 matrices

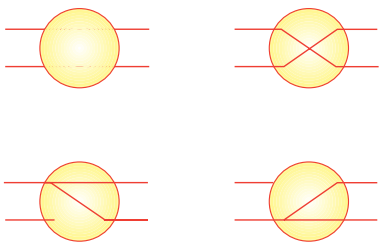


Figure 6.15 Cell element with broadcasting states

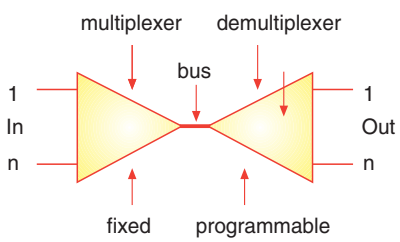


Figure 6.16 Illustration of a "switching bus"

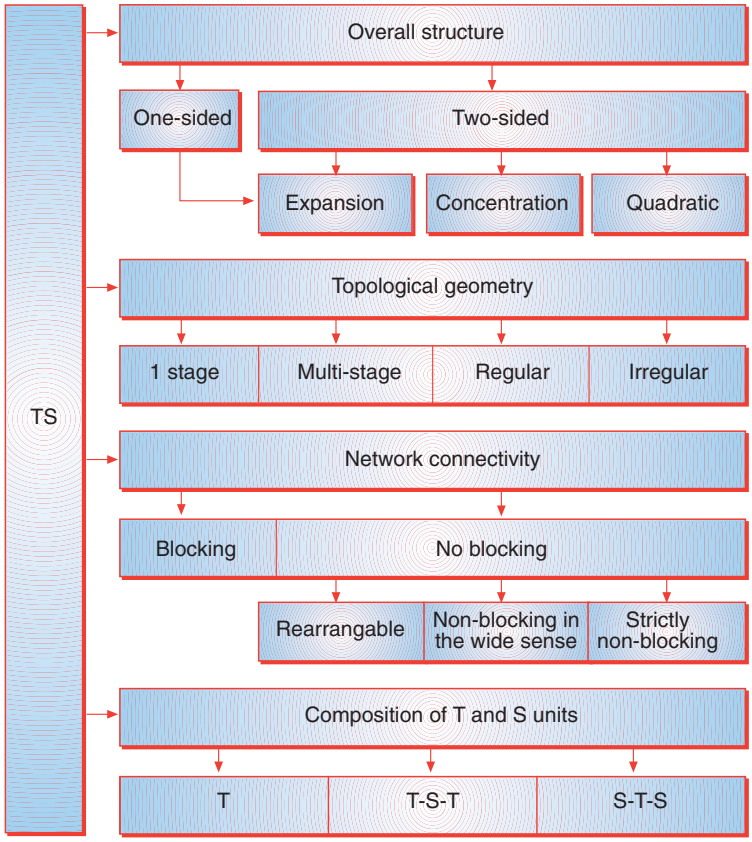


Figure 6.17 Classification of time switching networks

Bibliography

1 Ahmadi, H, Denzel, W. A survey of modern high-performance switching techniques. *IEEE Journal on Selected Areas in Communications,* 7(7), 1989.

2 Bassalugo, L A et al. The structure of one-sided connecting networks. *6th International Teletraffic Conference,* Munich, paper number 245, equation (10), 1970.

3 Broomel, G, Heath, J R. Classification categories and historical development of circuit switching topologies. *Computing Surveys,* 15(2), 95–133, 1983.

4 Feng, T. A survey of interconnection networks. *Computer,* 14(12), 12–27, 1981.

5 McMillen, R J. A survey of interconnection networks. *Proceedings ISS '84,* Florence, 7–11 May, 5.1.1–5.1.9, 1984. Also in *Globecom '84,* Atlanta, 1984.

7 Mathematical model of switching networks

In the preceding sections different types and configurations of switching networks are described. To be able to construct (calculate) networks, it is necessary to have one or more models to describe mathematically the behaviour of the networks. In the course of time much work has been done to give a mathematical description of switching networks. A pioneer in this field was Charles Clos who established the mathematics to construct multi-stage networks. Many others (Benes, Joell, among others) have continued this work and now there exists a relatively rich literature about mathematical switching theory.

This chapter gives a short description of the switching theory. The theory referred to is the one by Benes, and his work is based on group theory. The other theory referred to is that of Mirsky and is based on transversal theory. In both cases no mathematical proof is given of the theorems since this is beyond the scope of this paper.

7.1 Combinatory properties of switching networks

A switching network can be visualised by a "graph". There are two ways to do this (see examples in Figure 7.1):

- The switching element is the *node* and the inlets and outlets are *branches*
- The switching element is the *branch* and the inlet and outlets are *nodes.*

In the following the second method will be used.

Figure 7.2b shows the graph (or structure), *G*, of the 2 x 2-matrix in Figure 7.2a. The branches can be marked with 0 or 1 indicating the state of the switching element with 0 to indicate OFF-state and 1 to indicate ON-state. The state of the switch will then be given by a combination of zeros and ones. With marked branches of this kind the graph is called a "named graph" (Figure 7.2c) and this graph shows both the structure and the state of the switching network. Two nodes/terminals are connected when there is a chain of elements in the ON-state between the nodes/terminals.

Figure 7.3 shows the graph for a 2-stage switch.

To distinguish nodes of inlet and outlet types and nodes such as an exchange in a telecommunication network, we will from now on use the word terminal instead of node. The terminal symbol from Figure 7.1 will be the preferred indication for the node in a graph.

It is common to distinguish the following types of terminals:

- Inlet and outlet terminals
- Links or connections between switching stages.

Connections are always being set up between inlets and outlets.

In a telecommunication system the following connections are found:

- The terminals are connected pairwise together, i.e. only two terminals are involved in a connection
- More than two terminals are connected together
- One terminal is connected to several other terminals.

The most common type is the first and such connections are said to be "disjunct". When more than two terminals participate in a connection we speak about a "conference connection". The last type of connection is a typical example of broadcasting. Internally in a switching node this type of connection is used when the various signals are distributed to the terminals.

Let the graph, G, represent the structure of the network and let $V = \{g_0, g_1, ..., g_n\}$ be the set of all named graphs, i.e. those graphs got by assigning 0 and 1 to the branches (there is a named graph for each combination of ones and zeros). If the graph, G, only has to contain connections that are disjunct, the number of named graphs is restricted, i.e. there is a restricted number of named graphs that can represent a useful network.

We will now give a more precise definition of a graph that can be used to represent a switching network. Let T be the set of terminals in the network. Define G to be a subset of the Cartesian product T x T, i.e.:

$$G \subset T \text{ x } T = \{(i,u) : i \in T, u \in T\} \quad (7.1)$$

This Cartesian product consists of a set of pairs of elements. The graph G is now defined by taking elements from T x T which satisfy the following conditions:

- (i,i) and $(j,j) \notin G$, i.e. no switching element between the same terminal
- $(i,j) \in G$ iff:
  - $(j,i) \in G$
  - (i,j) is a branch in G
  - Terminal i and terminal j have no terminal between themselves on the same branch
  - There is a switching element between terminal i and terminal j.

Decompose T in the subsets I, U and L such that:

$$(I \cup U) \cap L = \phi = \text{(the empty set)} \quad (7.2)$$

where

I = the set of all inlets

U = the set of all outlets

L = the set of all links.

Figure 7.4 visualises the decomposition by showing the connection of more stages by using links.

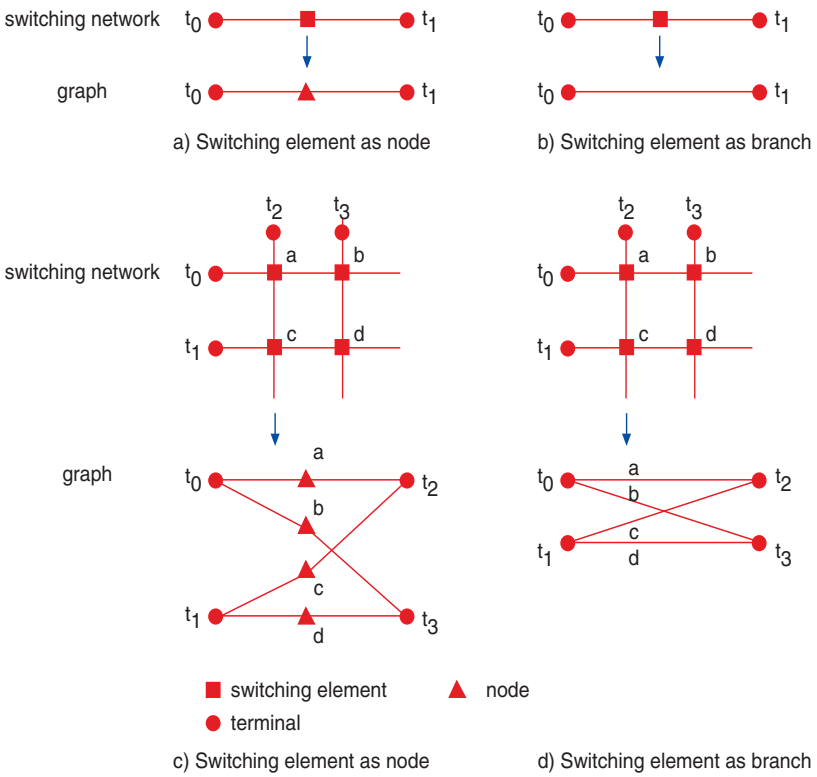


Figure 7.1 Different methods of graph representation of a switching network

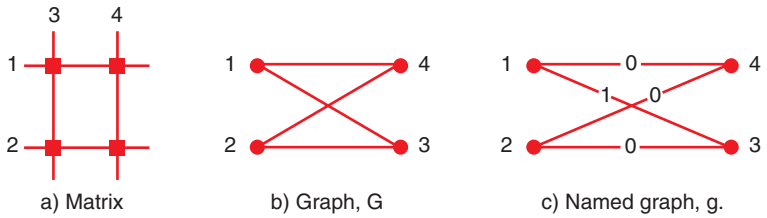


Figure 7.2 Graph representation of a 2 x 2 matrix

A chain, p, of length n between two terminals, i and j, is defined as a sequence of elements $z_k \in T$ for all $k = 0, 1, 2, ..., n$ such that:

$$z_0 = i;\ z_n = j;\ z_i \neq z_j \text{ for } i \neq j;$$
$$(z_k, z_{k+1}) \in G \subseteq T \text{ x } T$$
$$\text{for all } k = 0, 1, ..., n - 1 \quad (7.3)$$

Two paths (chains) p_1 and p_2 are said to be disjunct if they have no common element terminal or:

$$p_1 \cap p_2 = \phi = \text{empty set} \quad (7.4)$$

Let K denote the set of possible paths (p_i) (chains or connections) such that:

$$K = \{p_i\} \quad (7.5)$$

From K we can find the subset E consisting of all (disjunct) paths of the same length, i.e. chains (paths) not going back in the switching network. E will then consist of a set of elements each representing a single path or connection and all chains have the same length:

$$E = \{E_0, E_1, E_2, ..., E_m\} \quad (7.6)$$

Each of these paths will represent an elementary state and a collection of such states can be used to define a class of states of the network by setting:

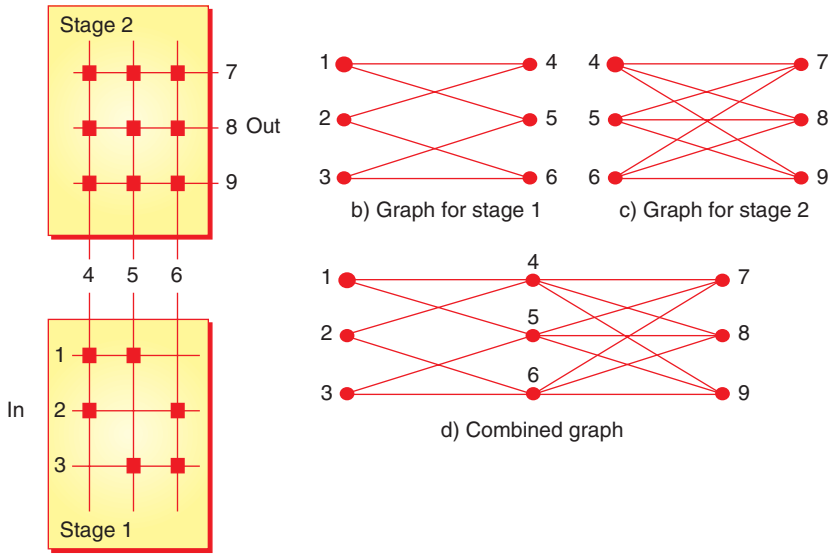


Figure 7.3 Combined graph representation

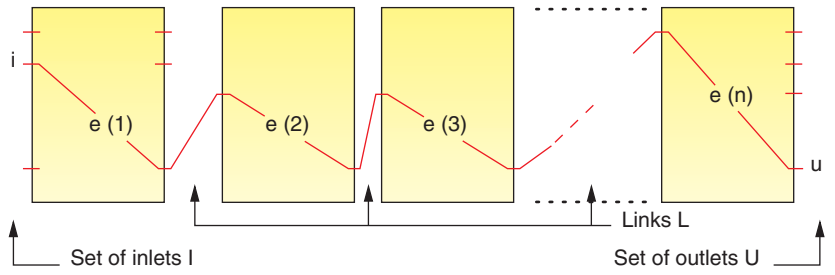


Figure 7.4 Decomposition of T

$$p \in E \Rightarrow p \in S$$
$$x,y \in S \Rightarrow x \cap y \in S \qquad (7.7)$$

if $x,y \in S$ and $p_1 \in x$; $p_2 \in y \Rightarrow p_1 \cap p_2 \in \phi$ then $x \cup y \in S$.

A switching network can be specified by:

- Structure represented by the graph G
- Inlets I
- Outlets U
- Elementary states or single connections E.

A common notation is:

$$\nu(G, I, U, E) \qquad (7.8)$$

Network states can be visualised as shown in Figure 7.5b for a network like the one in Figure 7.5a.

In the zero or ground state there are no connections and the network is on level L_0. From a level L_i the network goes to the next level L_{i+1} by adding 1 elementary state to level L_i.

Following this, a state consists of a number of elementary states and the number of these in a state x is denoted by:

$$n_L = |x|$$

where

$|x|$ = the cardinality of X.

The distance between two levels L_x and L_y can be defined by:

$$L_y - L_x = |y| - |x| = \text{difference in number of elementary states in } y \text{ and } x \qquad (7.9)$$

This also means that:

$y - x \in E$; i.e. the difference is again elementary states belonging to E (7.10)

A state specifies uniquely which terminals are connected. Such a specification is called an "assignment", function or picture such that:

$f : I \rightarrow U$; I = domain, U = range

A "maximal assignment" is an assignment that uses the whole domain or range.

All assignments a network can have is designated by T.

All maximal assignments are denoted by T_m ($T_m \subset T$).

In the following it is assumed that an assignment is 1-1 and ON, which means that each inlet is assigned only one outlet, and no two outlets are assigned the same inlet.

An assignment function can now be defined as:

$\phi(x) = \{(i,u) \in I \times U$: i and u are connected in state $x\}$ (7.11)

$\phi(x)$ gives all assignments that can be implemented in state x.

A state x implements exactly one assignment (or function).

Two terminals i and u are said to be connected in state x only if $p_{iu} \in x$ is a chain between the terminals, i.e. p_{iu} is an elementary state E_{iu} such that:

$\phi(E_{iu} = \{(i,u) \in IxU$: i and u are connected in $x\}$ (7.12)

$\phi(E_{iu})$ defines a unit assignment $f : i \rightarrow u$.

The set of all unit assignments is called Ω. A "call" can now be defined as a unit assignment that is not implemented.

The inverse function of $\phi(x)$ is denoted $\phi^{-1}(x)$ and defines the state giving the assignment $\phi(x)$.

"Equivalence" between two states is denoted "$x \sim y$" and is defined by:

$$\phi(x) = \phi(y) \qquad (7.13)$$

Equivalent states implement the same assignment but use different elementary states.

The "distance between two states" x and y is measured in the difference in number of elementary states and can be defined by:

$$d(x,y) = |\, \phi(x) \,\Delta\, \phi(y) \,| \qquad (7.14)$$

where

$\phi(x)$ = number of assignments in state x = number of E-states in x

$\phi(y)$ = number of assignments in state y = number of E-states in y

Δ = symmetric set difference

$|\,.\,|$ = cardinality

Example:

$d(x,0) = |\, x \,|$

$d(x,y) = 0 \Rightarrow x \sim y$, i.e. same number E.

Suppose a state x and there is a call in this state, i.e. the terminals i and u have to be connected. There must then exist a state y that can implement the old assignment $\phi(x)$ and the new unit assignment. This condition can be written:

$$\phi(y) = \phi(x) + \phi(E_{iu}) \qquad (7.15)$$

or

$$y = x \cup c$$

where

c = the elementary state that can implement the call.

This new state y is a "non-blocking" state.

Following this, a non-blocking state is a state where there is a path disjunct from all other paths between i and u so that:

$$x \cup v \in S \Rightarrow y = x \cup v > x$$

where v is a possible state.

If there is no state that fulfills this condition one can write:

$$\phi(y) \neq \phi(x) + \phi(E_{iu}) \qquad (7.16)$$

or

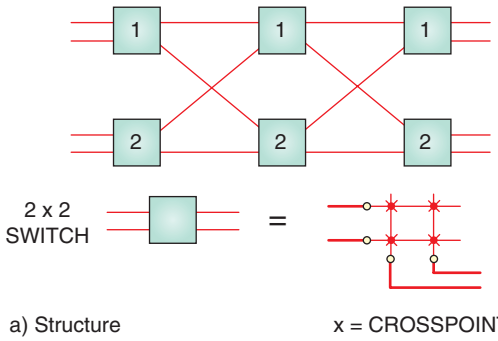
$$y \neq x \cup c$$

This new state y is then a "blocking" state.

The set of all non-blocking states is denoted by B'.

A state reflecting T_m has no idle inlet/outlet pair and is a non-blocking state.

From the above one can define a "strictly non-blocking" network, SNN, as a network that fulfills the condition:



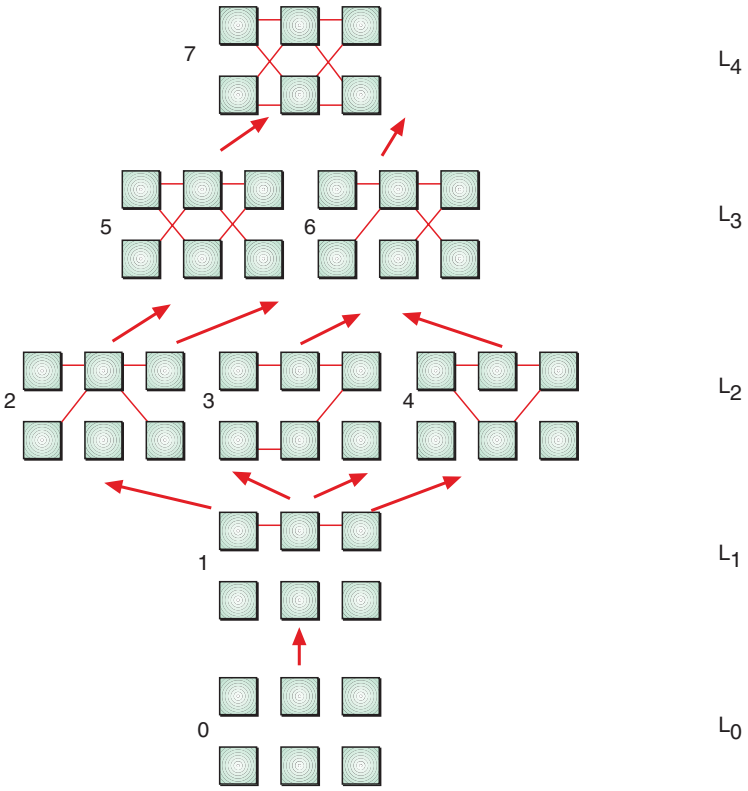


Figure 7.5 Illustration of some states in a network

$$B' = S \qquad (7.17)$$

This network has no blocking state and it is possible to find a subset X , B' such that:

1) X is sufficient

2) $X = X(d)$, i.e. X = d-closed.

A state is said to be sufficient if $\phi(X) = T$, i.e. every assignment can be realised by state X.

The closure of a set (here a state) X consists in this case of all y that for some $x \in X$ have $d(x,y) = 0 \Rightarrow y \sim x$. d-closure of X is the set of all states which are equivalent to a state in X, i.e.:

$$X(d) = \{y \in S : y \sim x \text{ for some } x \in X\} \qquad (7.18)$$

Suppose that there is at least one state y that can realise state x and the unit assignment c. This can be written:

$$\phi(y) = \phi(x) + c \qquad (7.19)$$

If this requirement can be met for all T_{max} by using some rule during the selection of elementary states to establish the calls, then the network is denoted as a

"non-blocking in the wide sense" or "widely non-blocking" network, WNN.

If X is a class of states, X is said to be "preservable" if equation (7.19) is satisfied for all states $y \in X$. This means that there will always be a state y such that y reflects both the new unit assignment and the assignment reflected by state x. If X is preservable, $0 \in X$ and $X \subseteq B'$, then X is sufficient. It is then possible to start in the zero state and set up calls (connections) one by one and still be in states belonging to the class X and all states are non-blocking. The conditions of WNN can then be written:

. X is preservable

. $X \subseteq B'$

. $X \leq$ d-closed $\Rightarrow X = \underline{X}$ (7.20)

where

$$\underline{X} = \{y \in S : y \leq x \text{ for some } x \in X\}$$

Assume it is possible to find a state y such that:

$$\phi(y) = \phi(z \sim x) + c \quad (7.21)$$

This means that by rearranging state x it is possible to find a state z equivalent to x so that the state $y > x$ reflects the assignment $\phi(x) + c$. A network avoiding blocking states by rearranging existing states is denoted a "rearrangable" network, RN.

Following the above, with respect to the properties of connection, one finds that a switching network belongs to one of the following groups:

- Strictly non-blocking network
- Non-blocking in the wide sense network
- Rearrangeable network.

7.2 Structures of networks

We will now take a look at some structures that implement the switching properties discussed above.

A basic structure to implement the switching properties discussed above is an N x M-matrix where there is a switching element in every cross-point in the matrix. Such a matrix can also be characterised by the "capacity" C which gives the number of simultaneous connections. ($N > M$ gives $C = M$, $N < M$ gives $C = N$.) A matrix like this satisfies the condition of a strictly non-blocking network. All maximal assignments can be accepted, i.e. assignments having the whole domain or range possible. To distinguish such a matrix from other types it will be called a "complete" N x M- or M x N-matrix. The number of switching elements is given by:

$$K = N \cdot M = N^2 \text{ if } N = M \quad (7.22)$$

To reduce K it can be shown that it is favourable to construct the network by using smaller matrices and more stages. Such a basic 3-stage structure is shown in Figure 7.6. The parameters n_1, n_2, r_1, r_2 and m can be chosen to get the combinatorial properties above. The structure consists of:

- An inlet stage consisting of r_1 matrices of size n_1 x m, i.e. $N = n_1 r_1$
- Middle or central stage consisting of m matrices of size r_1 x r_2
- An output stage consisting of r_2 matrices of size m x n_2, i.e. $M = n_2 r_2$.

Such a network is designated by:

$$\nu(m, n_1, n_2, r_1, r_2) \quad (7.23)$$

Another name is 3-stage "Clos network", after Charles Clos, being the first to study such structures and showed their special combinatorial properties.

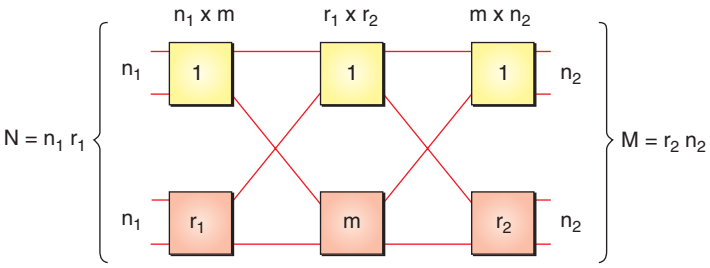


Figure 7.6 3-stage (Clos) network

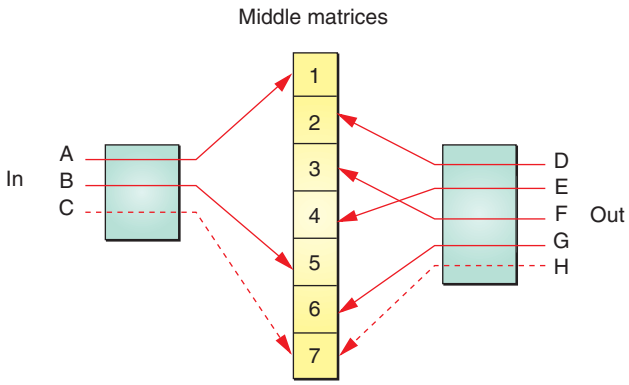


Figure 7.7 Illustration of a condition to make a 3-stage network strictly non-blocking

The network in Figure 7.6 is strictly non-blocking if $m = n_1 + n_2 - 1$. The proof of this is illustrated in Figure 7.7 where it is assumed that $n_1 = 3$, $n_2 = 5$, the inlets A, B and C are all connected to the same input matrix while D, E, F, G and H are connected to the same output matrix. As a worst case we can have the situation that all connections except C and H go over each middle matrix. To connect C and H it is necessary to add a new middle matrix giving:

$$m = (n_1 - 1) + (n_2 - 1) + 1 = n_1 + n_2 - 1 \quad (7.24)$$

$n_1 = n_2 = n$ and $r_1 = r_2 = r$ gives:

$$m = 2n - 1 \quad (7.25)$$

A 3-stage network can be transformed into a 5-stage network by dividing each middle matrix in a 3-stage network. This 5-stage network is also strictly non-blocking if the transformation satisfies Eq. (7.26) (or Eq. (7.27)).

The network in Figure 7.6 with $n_1 = n_2 = n$ and $r_1 = r_2 = r$ is said to be non-blocking in the wide sense if the number of middle matrices is given by:

$$m \geq \lfloor 3n/2 \rfloor \qquad (7.26)$$

where

$\lfloor x \rfloor$ = greatest integer less or equal to x

To avoid congestion it is necessary to follow some rules during set-up of the calls. Such a rule can be to fill up a middle matrix before using a new one.

A 5-stage network non-blocking in the wide sense can be derived from a 3-stage network as in Figure 7.6 by dividing the middle matrices in a 3-stage subnetwork in such a way that the condition of a non-blocking in the wide sense given by Eq. 7.26 is satisfied.

To get a rearrangable 3-stage network it is sufficient that the number of middle matrices is given by:

$$m \geq n_1 \text{ when } n_1 r_1 > n_2 r_2$$
$$m \geq n_2 \text{ when } n_2 r_2 > n_1 r_1 \qquad (7.27)$$

In most cases we have $n_1 = n_2 = n$ and $r_1 = r_2 = r$ such that:

"$\nu(m, n, r)$ is rearrangable iff $m \geq n$" (7.28)

This theorem is called the Slepian-Duguid theorem and the proof is based on Hall's theorem:

"Let A be any set and let $A_1, A_2, ..., A_r$ be any subsets of A. Let also $a_1 \in \boldsymbol{A}_1$, $a_2 \in \boldsymbol{A}_2, ..., a_r \in \boldsymbol{A}_r$ be distinct elements such that:

$a_i \neq a_j$ for $i \neq j$.

A necessary and sufficient condition that there exist r such elements is that, for each k in the range $1 \leq k \leq r$, the union of the sets $A_1, A_2, ..., A_r$ have at least k elements."

Another important theorem for rearrangeable networks is:

"In a rearrangeable network $\nu(m,n,r)$ with $m = n = r$ at most n - 1 existing connections have to be moved (rearranged) to avoid congestion."

However, we must keep in mind that in a greater network the number of such rearrangements can be large even if n is small.

Figure 7.8 illustrates the three configurations.

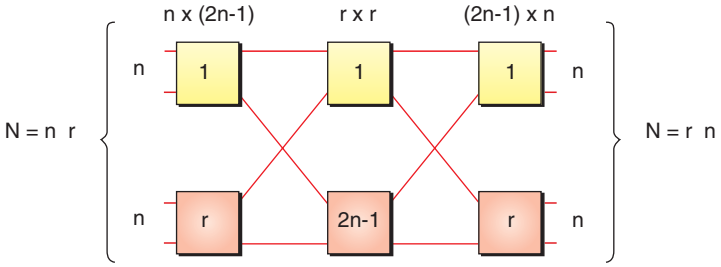
7.3 Transversal theory and switching networks

Instead of the mathematical theory above it is possible to use transversal theory to describe switching networks. This theory gives a more direct relation between the mathematical model and the physical switching network. The following gives a short description of the use of this theory.

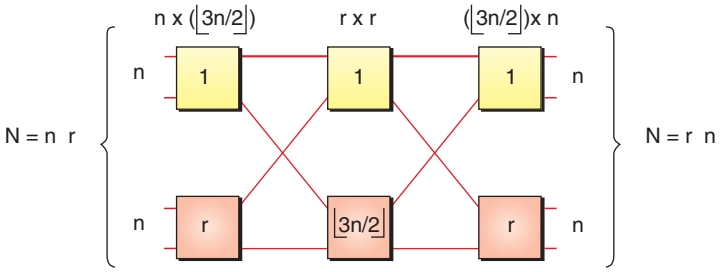
7.3.1 Definitions

To use this theory some definitions are necessary:

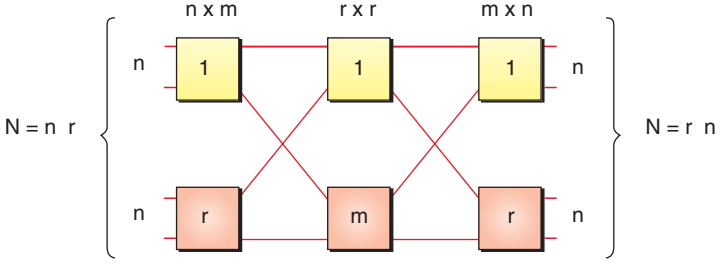
- A "map" or function means that elements in a set A are mapped on elements in a set B.
- A "bijective mapping" is a "1-1 and onto" mapping and is a map where every element in the set A is mapped on one and only one element in the set B, and every element in the set B is the map of only one element in the set A (i.e. the number of elements are equal in both sets).
- A "bijection" is a bijective mapping, i.e. a 1-1 and onto assignment.
- A "permutation" is a bijection which maps the set A on the set A.
- A "permutator" is an operator, program or equipment that executes a permutation.
- A "complete permutator" is a permutator that realises all possible permutations.



a) Strictly non-blocking network



b) Non-blocking network in the wide sense



c) Rearrangeable network

Figure 7.8 3-stage networks with different properties

- A “programmable permutator” is a permutator which eligibly can execute the desired permutations.
- A “fixed permutator” executes the same permutation all the time.
- A “bijector” is a system which can execute 1-1 and onto mapping of a set of inlets on a set of outlets, i.e. it can execute a bijection.
- A “symmetric bijector” is a bijector which can generate all functions of inlets on outlets and outlets on inlets (e.g. a quadratic matrix).
- “Imprimitive bijectors” are bijectors with mutual disjunct range and domain.
- An “assignment list” is a list assigned to the inlets and specifies which outlets have to be assigned to the inlets. The list has one element for every inlet.
- A “group assignment list” is a list for a system where inlets and outlets are divided into groups. An element in an inlet group gives the address to an outlet group.
- A “maximal assignment” means an assignment with every inlet assigned one outlet.
- A “switching stage” is a set of variables or changeable imprimitive bijectors.
- A “link stage” is a set of fixed or unchangeable imprimitive permutators.
- A “transversal” $T = \{a_0, a_1, a_2, ..., a_r\}$ is defined by:
  - from a family A consisting of r subsets $A_0, A_1, ..., A_r$, one element is taken such that:
  - exactly 1 element is taken from each subset
  - the elements are distinct, i.e. no element is taken two or more times.
- A “subtransversal” is a transversal not having elements from all members of the family.
- A transversal is said to have defect d if its d elements are missing.
- A set is said to have the cardinality n if it has n members.

7.3.2 Implementing a transversal

A system with N inlets and N outlets is divided into r groups with n elements in each group giving a family with r members on the inlet side and r members on the outlet side. Then, an assignment list is made for each inlet group. This means that every inlet in the r inlet groups is assigned one outlet among the outlets of the outlet groups such that the assignment is 1 - 1 and onto. The list contains group number and individual number within the group and looks like:

$$L = \{(g_0, i_0), (g_1, i_1), ..., (g_{(r-1)}, i_{(r-1)})\} \quad (7.29)$$

L contains only distinct elements. Another assignment will give a list with other elements.

A group assignment list l_g is split out from the list L. This list contains the number of the outlet group for each inlet member (group), giving:

$$l_g = \{g_0, g_1, ..., g_{(r-1)}\} \quad (7.30)$$

Contrary to the L-list the list l_g does not contain only disjunct elements.

All such lists, i.e. all possible assignments can be written as a family:

$$L_g = \{l_{g0}, l_{g1}, ..., l_{g(r-1)}\} \quad (7.31)$$

where every l_{gi}-list is a member of the L_g-family and contains one element from each of the r groups.

A transversal is now made by choosing elements from the L_g-list such that:

- Only one element is chosen from each of the members l_{gi} in the L_g-list.
- No elements are chosen twice or more.

It is not sure that a transversal can be found but if the assignment list is constructed from an 1-1 and onto assignment, i.e. from a permutation, a transversal can always be found.

Example:

Assume the following assignment list for $N = 16$ and $r = 4$:

Inlets: I = {0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15}

Outlets: O = {14,5,0,2,1,15,9,13,12,3,4,6,8,10,11,7}

The L-list is divided into 4 groups:

L_1 = (14,5,0,2), (1,15,9,13), (12,3,4,6), (8,10,11,7)

The L_g-list is found by taking the integer part after the elements in the L_1-list are divided by 4:

L_g = {(3,1,0,0), (0,3,2,3), (3,0,1,1), (2,2,2,1)}

The transversal is constructed by taking one element from each subset in the L_g-list:

T = {(3,0,1,2), (1,3,0,2), (0,2,3,1), (0,3,1,2)}

The transversal T consists of 4 subtransversals T_0 = (3,0,1,2), T_1 = (1,3,0,2), T_2 = (0,2,3,1) and T_3 = (0,3,1,2). In Figure 7.9 the subtransversal is illustrated, and the whole transversal is illustrated in Figure 7.10.

A transversal can be thought of as a connection between a pair of inlet/outlet groups and this pair will be disjunct. Each inlet group has a connection to the transversal and so has the output group.

If each of the subtransversals is replaced by a (symmetric) programmable bijector and each of the inlet and outlet groups is replaced by imprimitive programmable bijectors then this will be equivalent to a 3-stage switching network.

It is obvious that if all possible assignments have to be implemented, transversals reflecting all these assignments must be defined, so that an assignment is always found in one of the transversals.

A system of transversals reflecting all possible assignments is called a “complete system of transversals”, *CST*. A *CST* consisting of n transversals, each with cardinality r, is denoted $CST(n,r)$.

The condition to find at least one $CST(n,r)$ is given by the Slepian-Duguid theorem which can be stated as:

> “A bijection used on a system with N elements grouped in r groups of n elements such that it constitutes a family of r members will have at least one complete transversal $CST(n,r)$.”

An example of such a complete transversal is the one shown in Figure 7.10. By definition the inlet and outlet bijectors are imprimitive and if they are programmable as well, they satisfy the condition of a switching stage. The middle stage is represented by the transversal so the relation between a transversal used on a set of inlets and a switching network, in this case a 3-stage network, is recognised. Since $m = n$, the network is rearrangeable.

There are different methods to make a *CST*. The following example shows one of the methods. It is based on successive dividing. Assume the following permutation with $N = 16$ and $r = 8$ (this gives $n = 2$):

Inlet: $I = \{0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\}$

Outlet: $O = \{15,4,1,3,14,0,8,13,12,2,5,7,11,9,6,10\}$

O corresponds to a family with sub-families of cardinality 2. Every outlet is given an identity as a pair, group- and individ-number, by dividing the outlet number by 2. This gives the following list when grouping with one group for each bijector:

$L_f = \{[(7,1),(2,0)]; [(0,1),(1,1)];$
$[(7,0),(0,0)]; [(4,0),(6,1)];$
$[(6,0),(1,0)]; [(2,1),(3,1)];$
$[(5,1),(4,1)]; [(3,0),(5,0)]\}$

The group list is:

$l_g = \{7,2,0,1,7,0,4,6,6,1,2,3,5,4,3,5\}$

Because $n = 2$, this list can be rearranged to indicate two inlets for each bijector:

$l_g = \{(7,2),(0,1),(7,0),(4,6),(6,1),(2,3),$
$(5,4),(3,5)\}$

This will also give two transversals.

For this simple example the two transversals can, as above, be constructed by taking one element from each pair such that the same element is not taken two or more times. One finds the transversals:

$T_0 = \{7,1,0,4,6,2,5,3\}$

$T_1 = \{2,0,7,6,1,3,4,5\}$

A more systematic method is first to divide the list l_g into the same number of parts as the number of transversals. This gives the two parts:

$l_0 = \{7,0,7,4,6,2,5,3\}$

$l_1 = \{2,1,0,6,1,3,4,5\}$

Then the following procedure is used:

1 Start by marking the first pair of elements (in l_0 and l_1) which in our case is the pair (7,2).

2 Search from left to right to find an unmarked pair (an element in same position in l_0 and l_1. If no unmarked pair is found, the process is terminated.

3 When an unmarked pair is found, the "l_0-element" in this pair is compared to the "l_0-element" in the last marked pair:

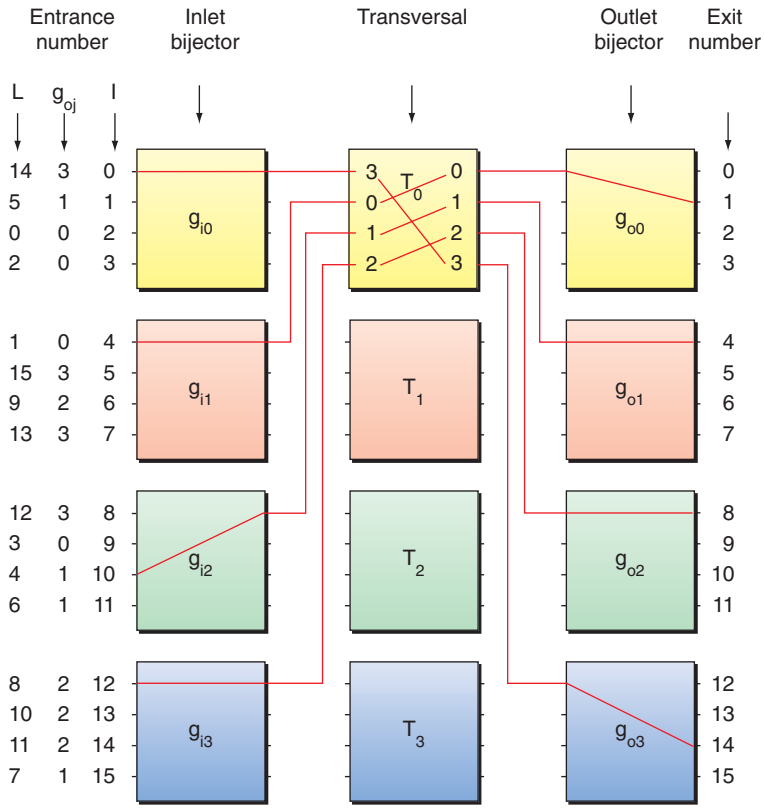


Figure 7.9 Illustration of a transversal

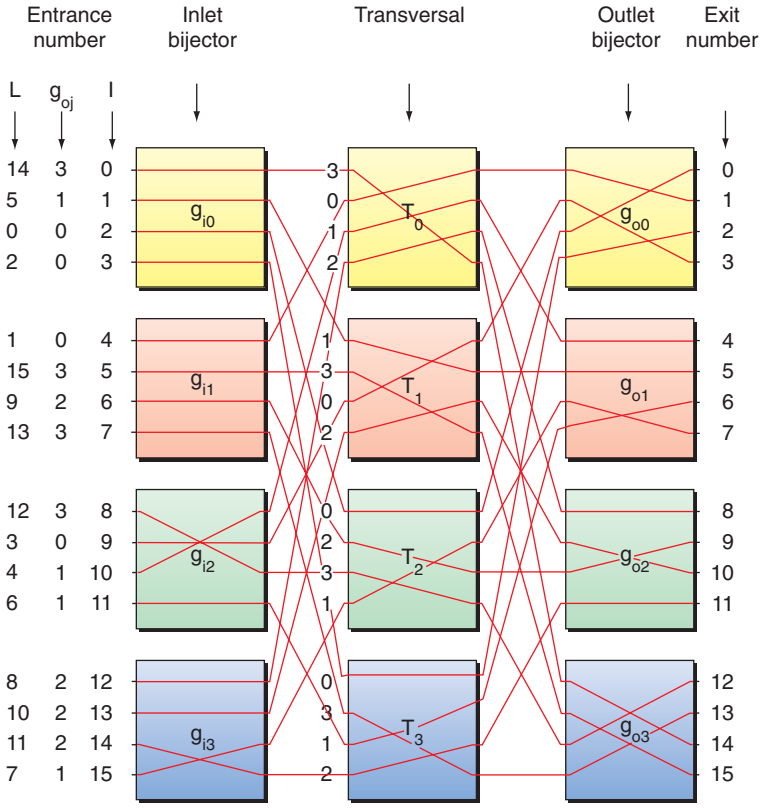


Figure 7.10 A complete transversal

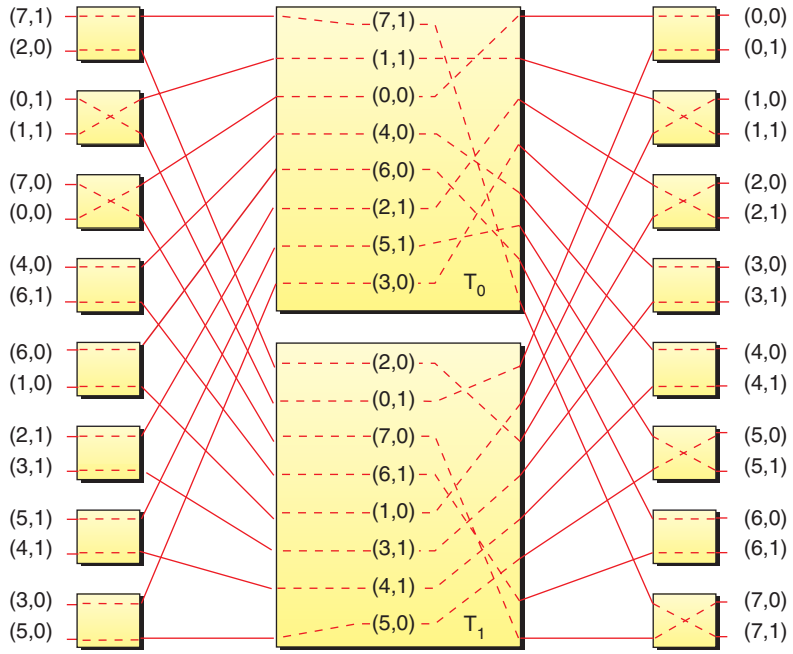


Figure 7.11 3-stage bijector system with cardinality 2

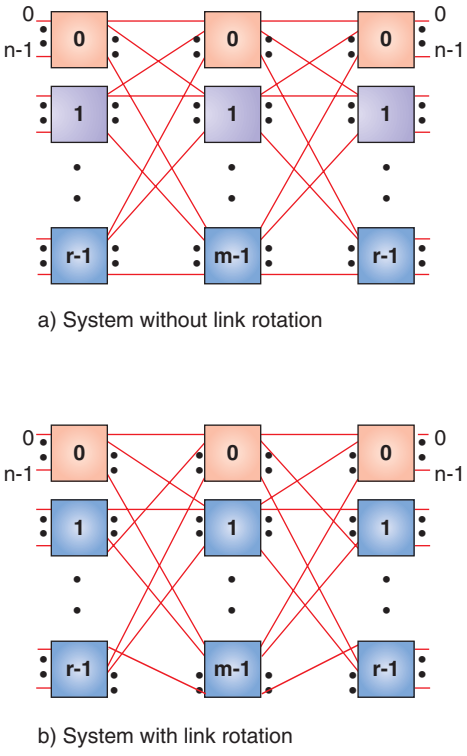


Figure 7.12 Examples of isomorphic structures

- If the elements are equal, the "l_1-element" and the "l_0-element" (in the unmarked pair) change place and the pair is marked. Then the process returns to step 2
- If the elements are not equal, the pair is marked and the process returns to step 2.

The method secures two transversals and cut and try is avoided. Figure 7.11 shows the transversal. As previously stated, this system is equivalent to a 3-stage switching network.

In the examples above there is no demand on how the (fixed) links are connecting the various bijectors (or switches). The only demand is that there should be one link between each input or output bijector and the subtransversals. The place where the terminals are terminated is not a matter of consideration as long as every bijector is symmetric. Such systems which differ only in the link assignment are called "isomorphic systems". The link assignment can be done in many different ways and two ways are shown in Figure 7.12.

It is important to note that two isomorphic systems implement the same assignment, even if the bijectors (switching matrices) are changing place, as long as the links attached to a bijector are the same ones for the two systems. Such a rearranging can be used to split out substructures (or subswitching network) or to compare different structures (networks). Figures 7.14 and 7.15 show two switching networks which both are a restructuring of the network in Figure 7.13. In Figure 7.14 a subnetwork of size 4 x 4 is sorted out and in Figure 7.15 a subnetwork of size 8 x 8 is sorted out.

A systematical method to split out subnetworks of size 2^c of a given network without link rotation and based on input groups with cardinality of 2 ($n = 2$) is:

1 The bijectors are numbered as in Figure 7.13. The input bijectors are then grouped in groups of 2^{c-1} bijectors. This gives input groups of size 2^c.

2 In the stages $t - 1$, $t - 2$, ..., $t - i + 1$ the bijectors belonging to a subsystem are numbered with the same number as numbers given from the input stage and such that the least number is placed on top.

3 Move the matrices within each stage $t - 1$, $t - 2$, ..., $t - i + 1$ such that the numbers of the bijectors are in the same order as in the input stage. During this process the links have to follow the respective bijectors so that the functional behaviour is not changed.

Subsystems as those in Figures 7.14 and 7.15 are called "imprimitive subsystems".

In a system with link rotation it is not easy to sort out subsystems.

The method to split out subsystems can also be used in systems with both concentration or expansion since such systems can be separated into one concentration/expansion part and one quadratic part.

7.4 Permutations in switching networks

A switching network executing a permutation is called a "permutation network". The name "complete permutation network" is used if all $N!$ permutations of the N inlets on the N outlets can be executed.

Any permutation can be represented by a "permutation matrix" which is a quadratic matrix with $p_{ij} = 1$ if inlet i is connected to outlet j and $p_{ij} = 0$ if i and j are not connected.

Example:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

P specifies which inlets and outlets are connected in a network with size 3 x 3.

Put:

$\underline{b}$ = column vector for the N inlets

$\underline{c}$ = column vector for the N outlets

Then the column vector $\underline{c}$ is given by:

$$\underline{c} = P \cdot \underline{b} \qquad (7.32)$$

Example:

$$\underline{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$$

and P as above gives

$$\underline{c} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_2 \\ b_1 \end{bmatrix}$$

The following inlet/outlet pairs are connected: (c_0, b_0), (c_1, b_2), (c_2, b_1).

If the inlet vector consists of m bits the inlet will be given by a matrix B with m columns and the outlet vector C by a matrix with m columns such that:

$C = P \cdot B$; P = the permutation matrix.

Permutation matrices are orthogonal, i.e.:

$$P^T \cdot P = I_N \qquad (7.33)$$

where

I_N = the identity matrix (size N x N)

P^T = the transposed P-matrix (changing columns and rows in P).

In a multistage switching network a permutation matrix can be used to indicate the link distribution between the various stages. This can be written as:

$$\underline{k}_{i+1} = P_i \cdot \underline{k}_i \qquad (7.34)$$

where:

- vector $\underline{k}_i$ specifies the outlets from switching stage sw_i
- vector $\underline{k}_{i+1}$ specifies the inlets to switching stage sw_{i+1}
- matrix P_i specifies which sw_i-outlets are connected to the sw_{i+1}-inlets.

The topology of a multistage switching network can then be described by the relation:

$$L = P_k \cdot (...(P_2 \cdot (P_1 \cdot (P_0 \cdot \underline{k}_0)))...) \qquad (7.35)$$

where the operations are to be executed in succession from right to left (first $P_0 \cdot \underline{k}_0$, then $P_1 \cdot (P_0 \cdot \underline{k}_0)$ and so on).

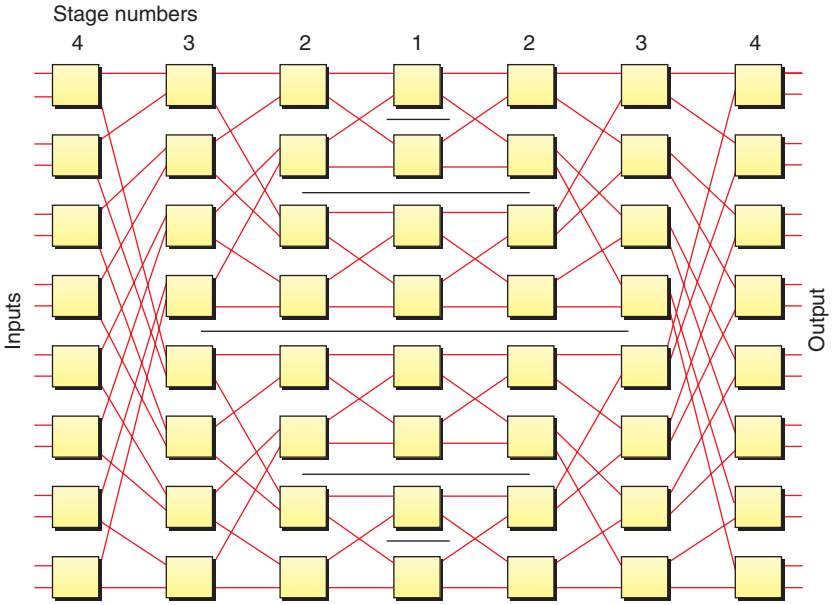


Figure 7.13 Switching network of size 16 x 16 based on β-elements

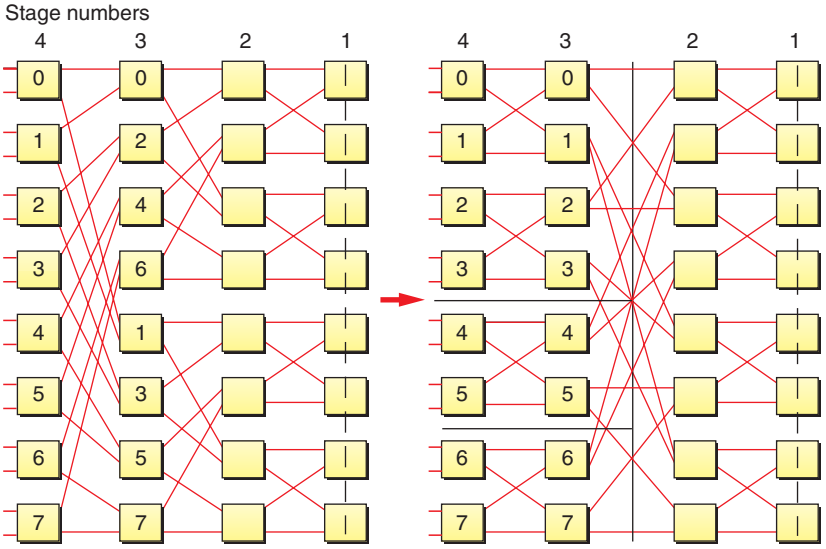


Figure 7.14 Separation of (outer) imprimitive subnetworks with size 4 x 4

This is illustrated in Figure 7.16. The switching stages, which are programmable bijectors, are the connections between the link stages.

Each switching stage consists of imprimitive (programmable) bijectors (switches) and normally they can be represented by a complete permutation matrix. The resulting switching properties or ability

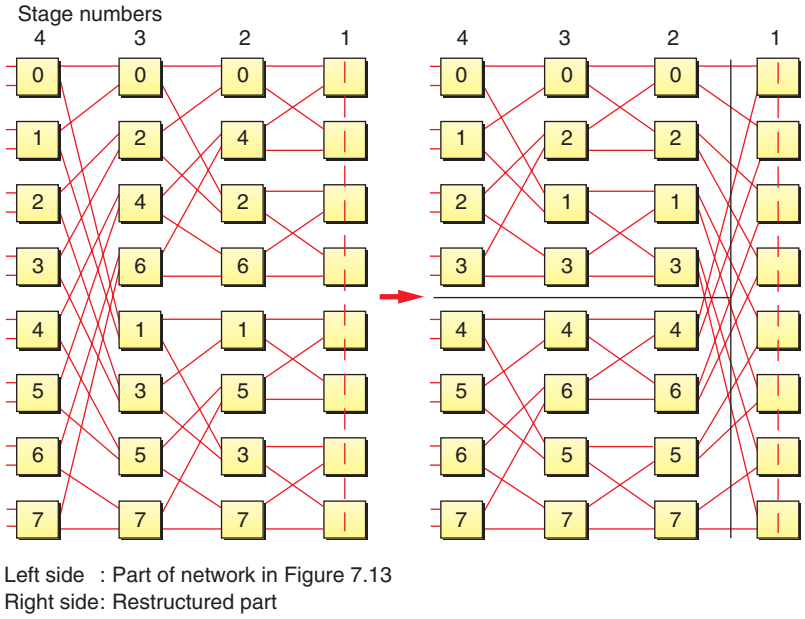


Figure 7.15 Separation of (outer) imprimitive subnetworks with size 8 x 8

to execute permutations can then be given by the relation:

$$\underline{U} = sw_n \cdot (P_{n-1} \cdot (sw_{n-1} \cdot (P_{n-2} \ldots \ldots (sw_1 \cdot (P_0 \cdot (sw_0 \cdot I))) \ldots \ldots))) \quad (7.36)$$

where

I = the inlet vector.

The above indicates that different switching configurations can be described by different permutation matrices. The most interesting of such matrices is perhaps "shuffling matrices".

The "shuffle matrices" is a class of permutation matrices of size N x N. Let $N = n \cdot r$ which means that the inlet vector ($\boldsymbol{b}$) is divided into r groups each containing n elements. Let the output vector ($\boldsymbol{c}$) be divided into n groups with r elements each. The shuffle permutation is given by "filling" the n outlet groups with one element from each of the r inlet groups (see Figure 7.17a).

If $x = 0, 1, 2, \ldots, N - 1$ (i.e. x is an integer), then an n-shuffling is given by the permutation:

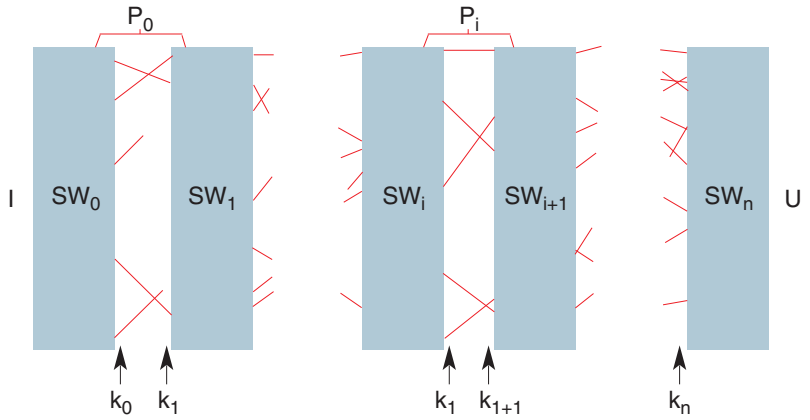


Figure 7.16 Multistage permutation network

$$\pi_n(x) = (nx + \lfloor nx/N \rfloor) \bmod N \quad (7.37)$$

(the outlet $\pi_n(x)$ is the picture of the inlet x).

Associated "shuffle matrix" is given by:

$$S_{n/r} = [s_{ij}] \quad (7.38)$$

where

$$s_{ij} = \begin{cases} 1 & \text{for } i = (nj + \lfloor nj/N \rfloor) \bmod N \\ 0 & \text{otherwise} \end{cases}$$

Shuffling of a vector gives different shuffle matrices. For example, a 16-element vector can give the following shuffle matrices: $S_{16/1}$, $S_{16/4}$, $S_{16/16}$, $S_{1/16}$, $S_{2/16}$, etc. $N = 8$, $n = 2$ and $r = 4$ gives the matrix and topology as illustrated in Figure 7.17b. The shuffling above is achieved by assignment from left to right. Assignment from right to left gives "inverse shuffling" as indicated in the figure. In general, we have:

$$S_{1/N} = S_{N/1} = I_N$$
(= unit matrix of size N x N)

$$S^{-1}{}_{n/r} = S_{r/n} = S^T{}_{n/r}$$
($S^{-1} = S^T$)

$$S^{-1}{}_{n/n} = S_{n/n} = S^T{}_{n/n}$$
(i.e. S is symmetrical) (7.39)

Shuffle-matrices can also be written as a sequence of smaller matrices.

Of special interest is the "perfect shuffle matrix" (PSM). A PSM is defined by:

$$n = 2 \text{ and } r = N / 2 \quad (7.40)$$

Provided that each switching stage is composed of imprimitive bijectors, the capability of the network to implement permutations is determined by the number of stages together with the link assignments.

7.5 Networks with partial permutation

In contradiction to the above, a "partial permutation network" cannot execute all permutations. In many circumstances such networks are of great interest because they give a simple network and also to a certain extent simple control. Such networks are in use in multiprocessor systems. Switching networks based on partial permutation can be characterised by:

- There is only one possible way between a specified pair of inlets and outlets

- The networks have internal congestion since all permutations cannot be executed simultaneously.

Figure 7.18 shows some existing partial permutation networks.

The switching elements in those networks are often of the cell type, i.e. 2 x 2-element with reduced number of states (often called b-element).

Since the number of stages is restricted these networks have internal congestion.

The first property of these networks (only one way between specified inlet/outlet pair) can be used to establish simple control since the way is determined by the output address. For this reason such networks are called “self routing networks”. An address of the input gives the way through the network. Of course, the elements must be provided with sufficient logic to be able to steer the connection through the network. The relatively simple control is one of the reasons for the interest in using such networks in parallel processing.

The number of stages in the partial permutation networks is given by:

$$m = \lfloor \log_2 N \rfloor, \quad N \geq 2$$

Assuming N = number of inlets = number of outlets. This gives:

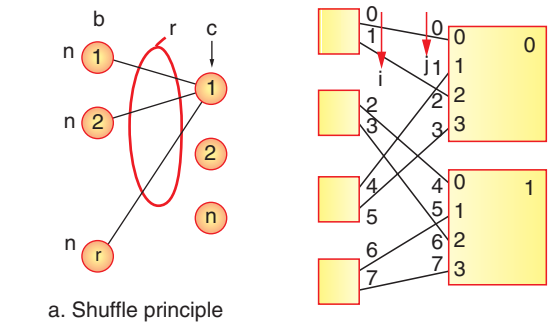
$$N = 2^m \tag{7.41}$$

This is not in general a restriction since it is always possible to set $N = 2^{m+1} > N$ when $2^m < N$, implement the network with this value and then drop unused inlets, outlets and links.

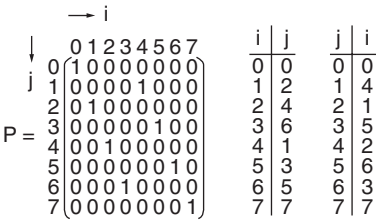
Actual permutations/operations are:

- Perfect shuffle, PS
- Inverse perfect shuffle, IPS
- Cube
- Plus minus 2^i (PM2I)
- β-permutation, β
- Bit reversing, ρ
- Bit switch, δ
- Crossover, γ.

In the following description of the link topology between two succeeding stages for the different permutations, the address of an input/output is represented by a binary sequence:



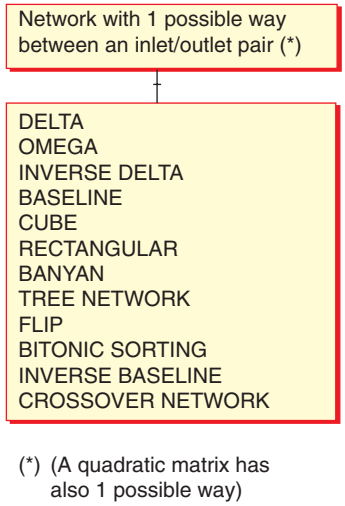
a. Shuffle principle



i = (2j +⌊2j/8⌋) mod 8

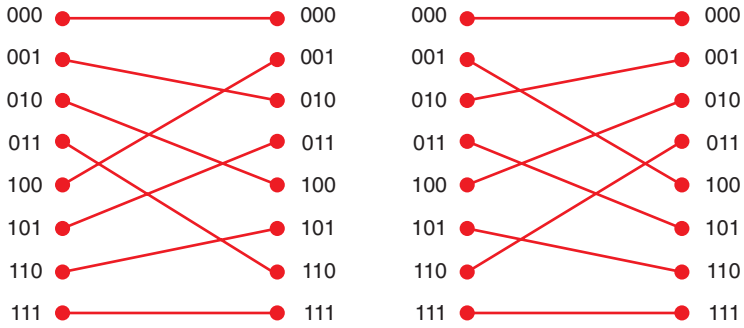
b. Shuffle example

Figure 7.17 The shuffle principle



(*) (A quadratic matrix has also 1 possible way)

Figure 7.18 Partial permutation multistage networks



a) PS permutation with N=8

b) IPS permutation with N=8

Figure 7.19 PS and IPS permutation and connections with N = 8

$$A = (a_{m-1}, a_m, ..., a_1, a_0).$$

The link connection with perfect shuffling (PS) and inverse perfect shuffling (IPS) is given by:

$$PS(a_{m-1}, a_{m-2}, ..., a_1, a_0) = (a_{m-2}, ..., a_1, a_0, a_{m-1}) \tag{7.42}$$

$$IPS(a_{m-1}, a_{m-2}, ..., a_1, a_0) = (a_0, a_{m-1}, a_{m-2}, ..., a_1) \tag{7.43}$$

This corresponds to left rotate and right rotate, respectively.

The permutations and connections are illustrated in Figure 7.19 when $m = 3$.

Figure 7.20 shows a 3-stage (OMEGA) network based on PS-permutation. It has a PS-permutation from left to right and IPS from right to left. To obtain switching functions the switching elements between the stages must have states as indicated in

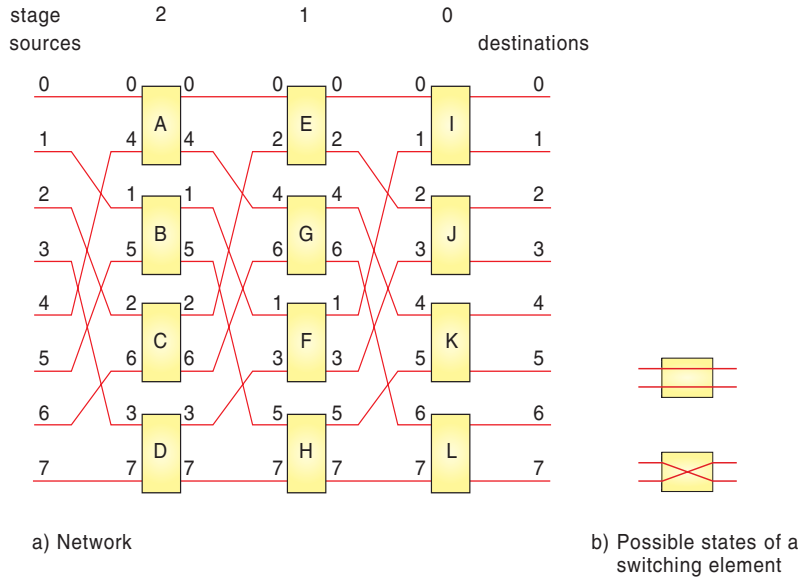


Figure 7.20 3-stage OMEGA network

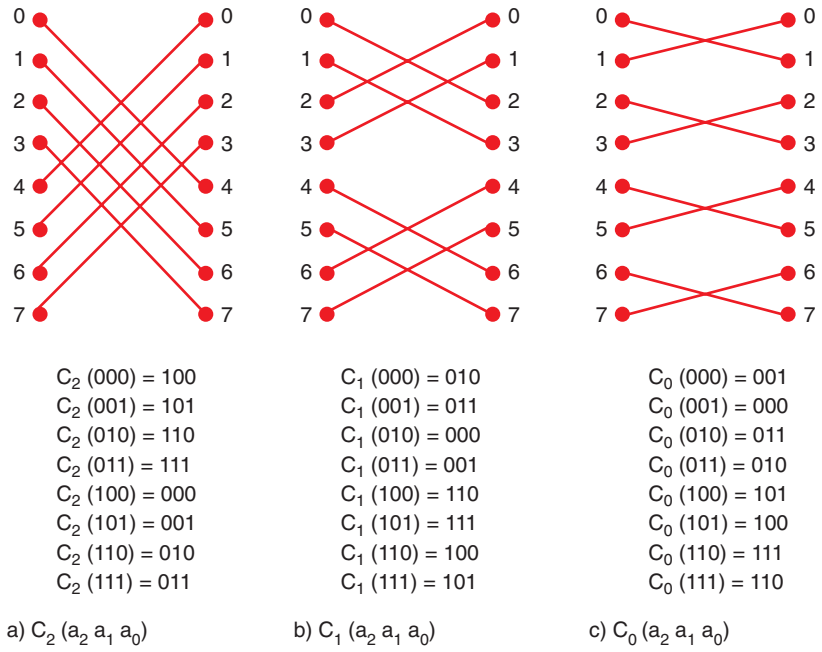


Figure 7.21 Cube permutations and associated connections with N = 8

Figure 7.20b. The exchange/ switching function can be written:

$$S(a_{m-1}, a_{m-2}, ..., a_1, a_0) = (a_{m-2}, a_{m-1}, ..., a_1, \overline{a}_0) \quad (7.44)$$

$\overline{a}_0$ = negation of a_0

Cube permutations are given by:

$$C_i(a_{m-1}, a_{m-2}, a_1, a_0) = (a_{m-1}, a_{m-2}, ..., \overline{a}_i, ..., a_1, a_0) \quad (7.45)$$

$\overline{a}_i$ = negation of a_i

Figure 7.21 illustrates the cube permutation and link pattern when $N = 8$.

Another representation of the cube permutations with $N = 3$ is shown in Figure 7.22.

This representation has been given the name CUBE network.

Multistage network with links (connections) between the stages based on CUBE or/and perfect shuffle permutations gives a network called Generalised CUBE network or only CUBE network.

Set N = number of inlets/outlets = 2^m and $m = \log_2 N$ = number of stages. The number of switching elements (SEs) in each stage is $N / 2$ and every element has two inputs and two outputs. Total number of SEs is $Nm / 2$. The switching element can have stages as shown in Figure 7.20b. The SEs in stage number i are designated SE_i and the stages are numbered 0, 1, ..., m-1.

With two stages the total number of states can be written:

$$n_s = 2^{Nm/2}$$

To get a non-blocking network the number of states must be:

$$N! \leq 2^{Nm/2}$$

The CUBE network is not a non-blocking network according to the statement above.

In stage number i the SE_is can connect the permutations between the inlets and outlets i.e. C_i - and C_{i-1} - permutations if the following rules are followed:

- Inlets/outlets of an SE_i are given the same number.
- Inlets/outlets of an SE_i have to be numbered such that when applying the C_i - permutation the rest can be obtained from the preceding ones.
- The identity of the inlets/outlets give the addresses to the components that have to be connected.
- The links to one stage of SEs, for example SE_{m-1}, are connected to the SEs such that the corresponding C-permutations can be implemented in every SE.
- The C-permutations use binary representations of the links.

Figure 7.23 illustrates the CUBE network with $N = 8$ ($m = 3$ = number of stages).

Connection between an inlet/outlet pair is obtained by using the straight and cross states of the switching elements. When an element C_i is in straight state it implements the C_i-permutation.

Figure 7.24 illustrates how a connection between an inlet and an outlet can be established. It is seen that the operations are quite simple. The following notation is used:

S = source = $S(s_{m-1}, s_{m-2}, ..., s_1, s_0)$

D = destination = $D(d_{m-1}, d_{m-2}, ..., d_1, d_0)$

The rule to be followed when establishing a connection is:

In stage number i:

SE_i is set to straight state if $s_i = d_i$

SE_i is set to cross state if $s_i \neq d_i$.

Figure 7.25 illustrates the isomorphism between OMEGA and CUBE networks. It is seen that by changing switching element F and G in the OMEGA network, the new network has the same link patterns as the CUBE network.

The PM2I operation/permutation can be written as:

$$PM2_{+i}(X) = X + 2^i \bmod N$$

$$PM2_{-i}(X) = X - 2^i \bmod N \qquad (7.46)$$

where

X = inlet address;
N = number of inlets/outlets = 2^m

Figure 7.26 illustrates the permutations and links when $N = 8 = 2^3$.

A multistage network based on PM2I-permutations gives a network called a Data Manipulator Network or only MANIPULATOR network. In addition to the PM2I-connections between the stages, the MANIPULATOR networks have direct connections between non-consecutive stages. An example of an Augmented Data Manipulator network is illustrated in Figure 7.27. The Data Manipulator has more possible routes from any inlet to any outlet than the CUBE or OMEGA network. This on the cost of more complex control since the switching element (SE) has three inlets and three outlets except for the switching elements in the input/output stage which have one inlet and three outlets / three inlets and one outlet each.

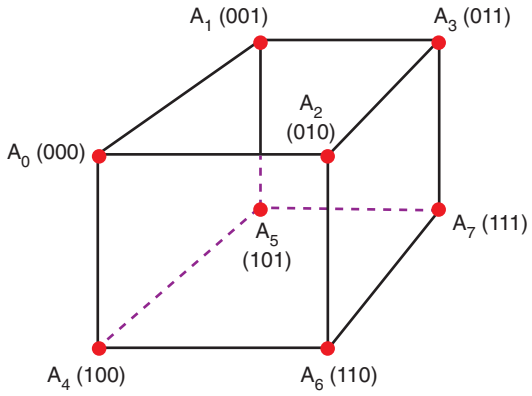


Figure 7.22 Three-dimensional binary cube

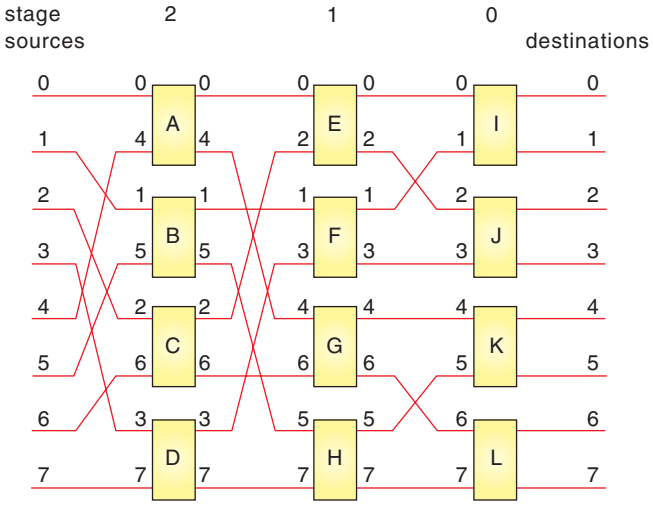


Figure 7.23 Illustration of a CUBE network with N = 8

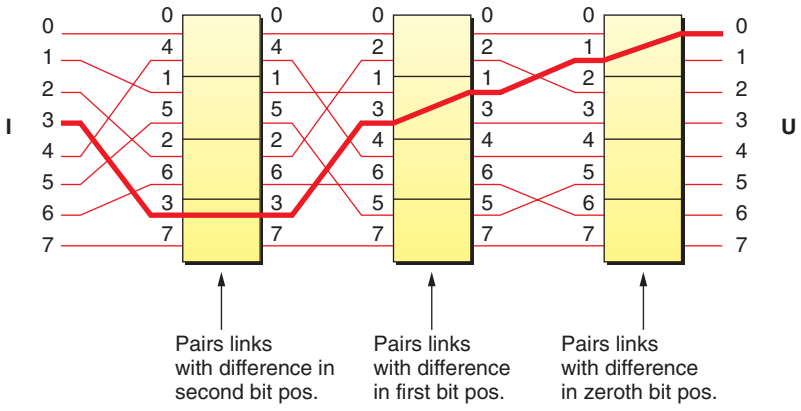


Figure 7.24 Illustration of a connection in a CUBE network with N = 8

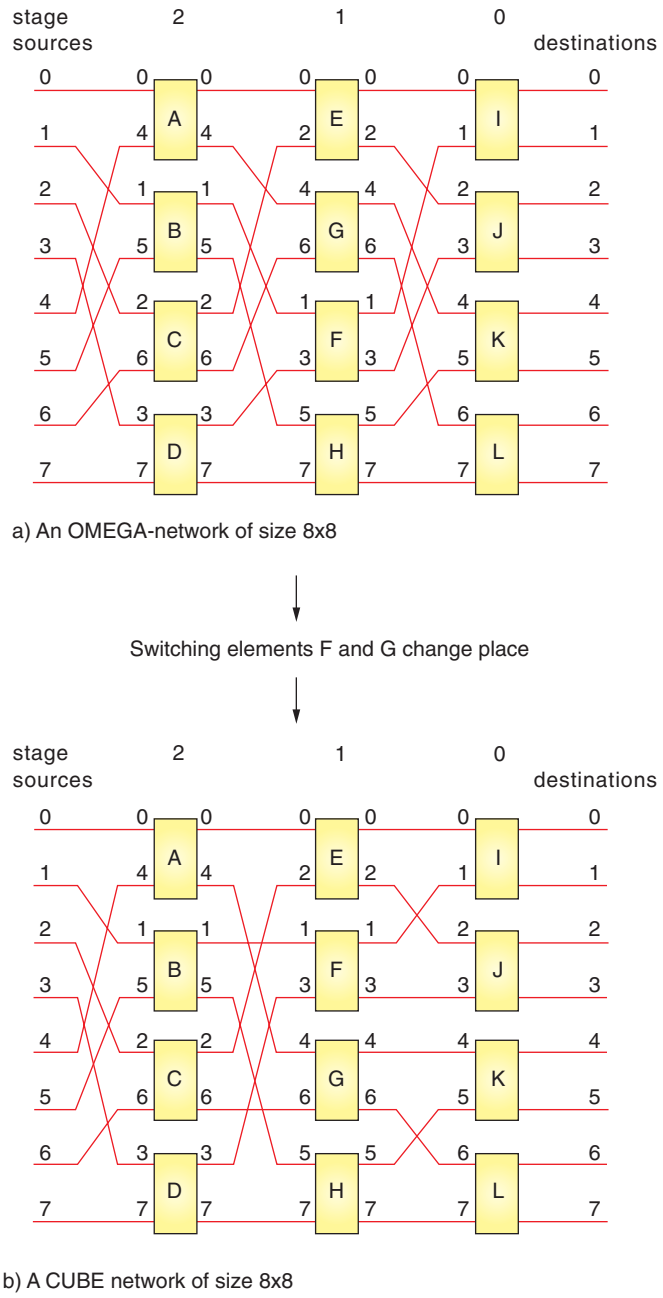


Figure 7.25 Illustration of the isomorphism between CUBE and OMEGA networks

The number of stages is:
$m = \log_2 N$; N = number of inlets/outlets

The number of SEs per stage is N

Total number of SEs is Nm

The class of Data Manipulator networks includes the Data Manipulator, the Augmented Data Manipulator, the Inverse Augmented Data Manipulator and the Gamma network. An interesting alternative to the Banyan- and the Perfect Shuffle (PS) -permutations is the crossover-permutation, as it has potential for a simple optical implementation. It is also of interest in applications where it is important that the lengths of the connections between the switching elements are equal. Networks based on the crossover permutation are topologically equivalent to the PS-network and the Banyan-network. An example of a crossover network is shown in Figure 7.28b. The basics of the realisation is illustrated in Figure 7.28a.

The other permutations mentioned above can be written:

β-permutation:
$\beta_k (a_{m-1}, ..., a_k, ..., a_1, a_0)$
$= (a_{m-1}, ..., a_0, a_1, ..., a_k)$ (7.47)
(the β_k-permutation is stage dependent)

Bit reversing-permutation:
$\rho(a_{m-1}, a_{m-2}, ..., a_1, a_0)$
$= (a_0, a_1, ..., a_{m-1})$ (7.48)

Bit switching-permutation:
$\delta(a_{m-1}, ..., a_k, ..., a_1, a_0)$
$= (a_1, a_2, ..., a_{m-2}, a_{m-1}, a_0)$ (7.49)

Based on these permutations a lot of networks can be defined which differ in the link topology.

In a DELTA network the link topology is defined by:

- A PS-permutation between each switching stage.

An OMEGA network is a DELTA network with:

- A PS-permutation in front of each switching stage.

An INVERSE-OMEGA network (FLIP network) is a DELTA network with:

- An IPS-permutation from left to right behind each stage
- A PS-permutation between the stages when going from right to left
- No permutation before the first stage.

An N-CUBE network is a DELTA network with:

- An IPS in front of the first stage
- A β_{s-k}-permutation in front of each stage where $1 \leq k \leq s$
- No permutation behind the last stage.

An INVERSE N-CUBE network is a DELTA network with:

- A β_{k+1}-permutation behind each stage $0 \leq k \leq s-1$
- A PS-permutation after the last stage.

A (SW) BANYAN or BITONIC SORTING network is a DELTA network with:

- A β_{k+1}-permutation behind each stage $0 \leq k \leq s-2$
- No permutation behind the last stage and before the first stage

- No permutation behind the last stage and before the first stage.

An INVERSE (SW) BANYAN network is a DELTA network with:

- A β_{s+k}-permutation before each stage $1 \leq k \leq s\text{-}1$
- No permutation behind the last stage and before the first stage.

Figure 7.29 shows examples of the networks above with size 8 x 8 and Table 7.1 shows some of the permutations for 4 bit addresses.

As an example of the control of such networks an N-CUBE network of size 8 x 8 is chosen as illustrated in Figure 7.30. The network has $\log_2 N = 3$ stages. The switches are of the β-type and have two states: cross and straight. Inlets and outlets are represented by their binary value (i_2, i_1, i_0) and (u_2, u_1, u_0), respectively.

In switch stage k the operation of a switching element can be written as a binary operation on the inlet and outlet addresses:

$$R_k(i_{m-1}, i_{m-2}, ..., i_1, i_0 \setminus u_{m-1}, u_{m-2}, ..., u_1, u_0) = (i_{m-1}, i_{m-2}, ..., i_1, u_{m-1-k}) \quad (7.50)$$

where

$k = 0, 1, 2, ..., m - 1$
(The stages are numbered from left to right.)

Assume that inlet with binary address (001) and outlet with binary address (101) have to be connected. The above operations give:

Stage 0:
$PS_0(i_2, i_1, i_0 = 001)$
$= (i_1, i_0, i_2) = (010)$

That is;
$R_0(i_2, i_1, i_0 \setminus u_2, u_1, u_0 = 010 \setminus 101)$
$= (i_2 i_1 u_{2\text{-}0}) = (011)$

Stage 1:
$\beta_{3\text{-}1}(i_2, i_1, i_0 = 011)$
$= (i_1, i_0, i_2) = (110)$

That is;
$R_1(i_2, i_1, i_0 \setminus u_2, u_1, u_0 = 110 \setminus 101)$
$= (i_2, i_1, u_{2\text{-}1}) = (110)$

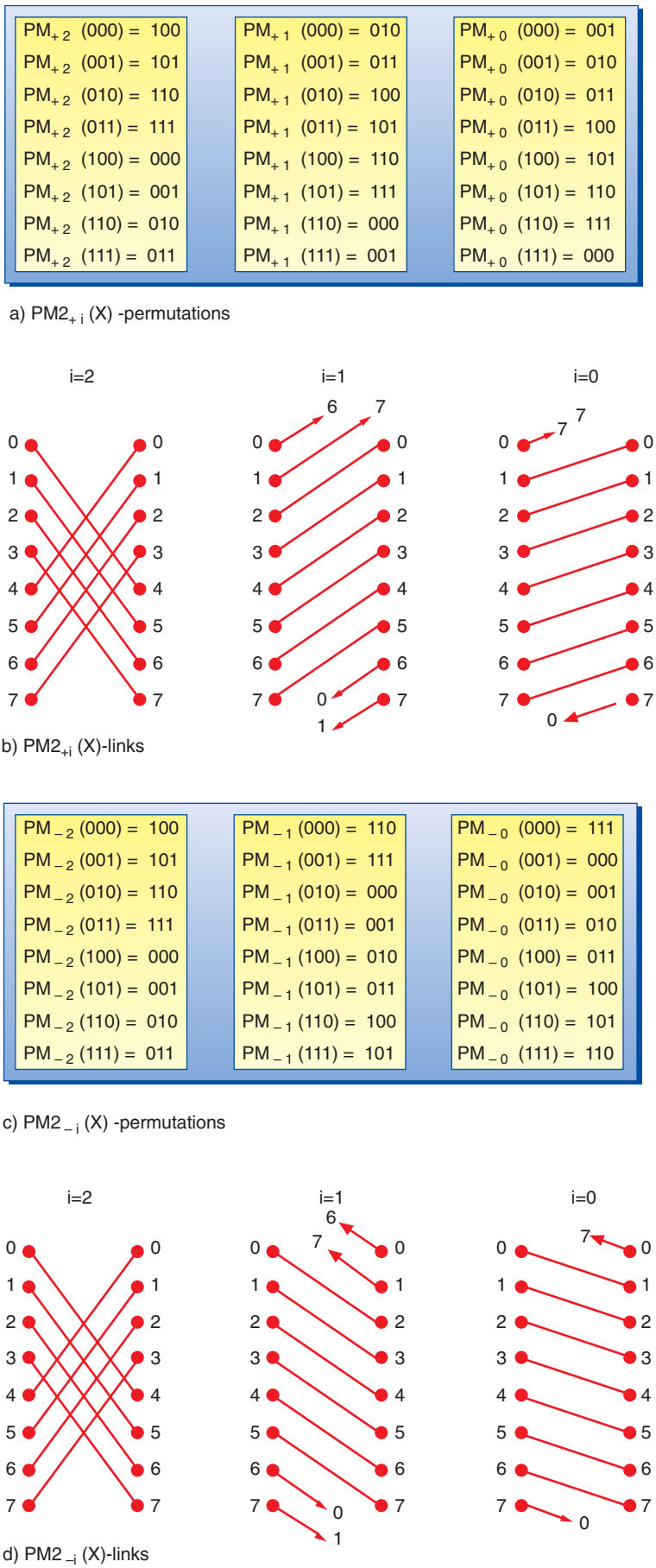


Figure 7.26 PM2I permutations and associated links with N = 8

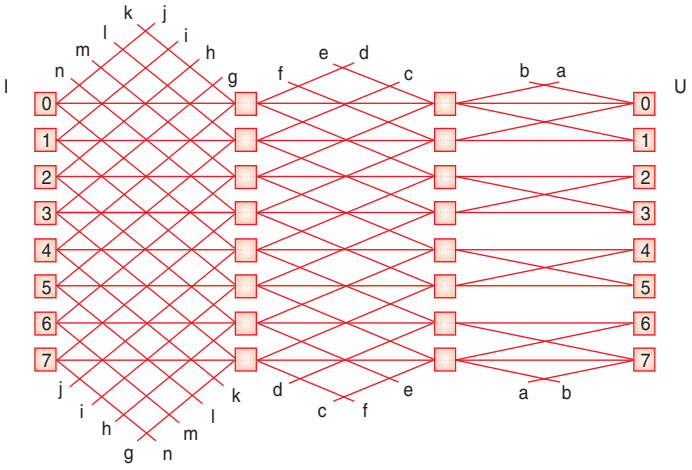


Figure 7.27 Data Manipulator of size 8 x 8

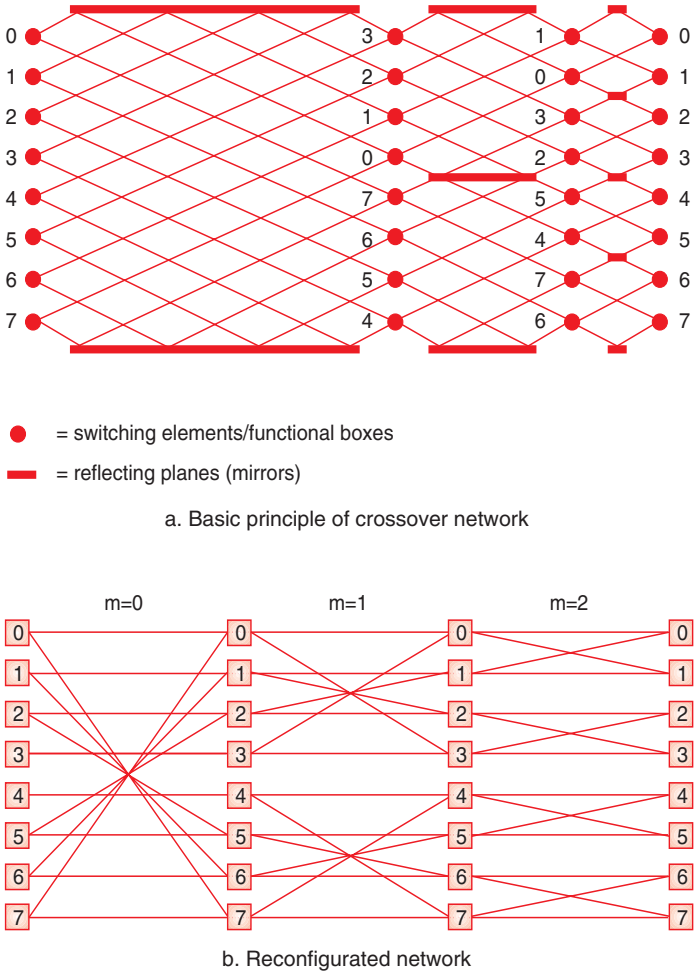


Figure 7.28 A basic crossover network

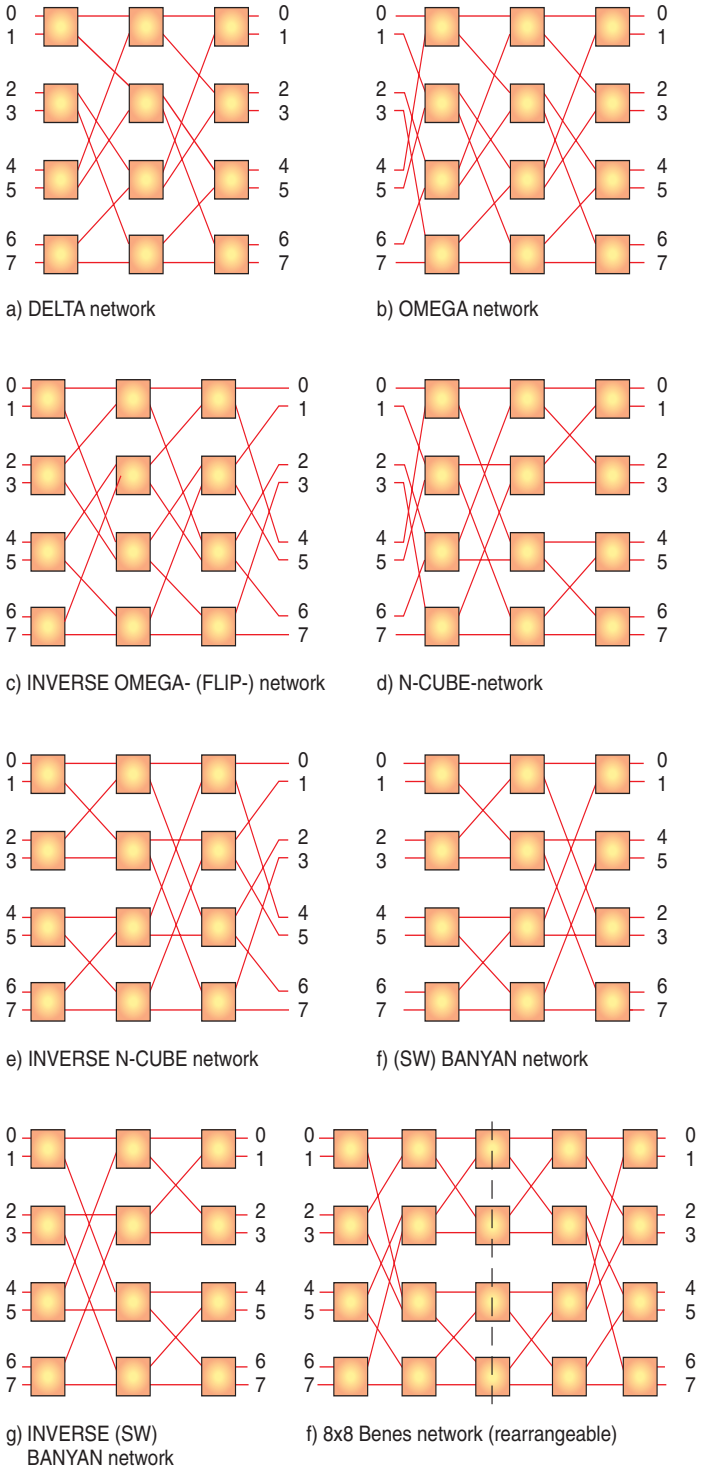


Figure 7.29 Examples of partial permutation of networks of size 8 x 8

Table 7.1 Permutations for 4 bit addresses, $m = \log_2 16 = 4$

I $i_3 i_2 i_1 i_0$	PS	IPS	β_3	β_2	β_1	β_0	ρ	δ	Crossover permutation (m_i = stage number i)		
									$m = 2$	$m = 1$	$m = 0$
0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
0001	0010	1000	1000	0100	0010	0001	1000	0001	1110	0110	0010
0010	0100	0001	0100	0010	0001	0010	0100	1000	0010	0010	0001
0011	0110	1001	1100	0110	0011	0011	1100	1001	1100	0100	0011
0100	1000	0010	0010	0001	0100	0100	0010	0100	0100	0011	0100
0101	1010	1010	1010	0101	0110	0101	1010	0101	1010	0101	0110
0110	1100	0011	0110	0011	0101	0110	0110	1100	0110	0001	0101
0111	1110	1011	1110	0111	0111	0111	1110	1101	1000	0111	0111
1000	0001	0100	0001	1000	1000	1000	0001	0010	0111	1000	1000
1001	0011	1100	1001	1100	1010	1001	1001	0011	1001	1110	1010
1010	0101	0101	0101	1010	1001	1010	0101	1010	0101	1010	1001
1011	0111	1101	1101	1110	1011	1011	1101	1011	1011	1100	1011
1100	1001	0110	0011	1001	1100	1100	0011	0110	0011	1011	1100
1101	1011	1110	1011	1101	1110	1101	1011	0111	1101	1101	1110
1110	1101	0111	0111	1011	1101	1110	0111	1110	0001	1001	1101
1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111	1111

Stage 2:
$\beta_{3\text{-}2}(i_2, i_1, i_0 = 110)$
$= (i_1, i_0, i_2) = (101)$

That is;
$R_1(i_2, i_1, i_0 \setminus u_2, u_1, u_0 = 101 \setminus 101)$
$= (i_2, i_1, u_{2\text{-}2}) = (101)$

In Figure 7.30 links and switching states are marked.

7.5.1 Construction of a partial permutation network

An N x N network can be implemented by a quadratic N x N matrix. This matrix is able to implement all permutations which implies that it is a complete permutations network, i.e. a non-blocking network. A multi-stage network can be implemented from an N x N network by an iterative method illustrated in Figure 7.31. The result is called a BASELINE network and can be used to compare different network configurations.

For simplicity we start with a network of size of N x $N = 2^m$ x 2^m. This network is divided into two equal parts, each with 2^{m-1} inputs. The original 2^m inputs are then connected to the new (2^{m-1}) inputs by β-elements as shown. This process is repeated until one has only β-elements in the network.

Note that the iteration is used only on one (here left) side of the N x N network and so giving a network with restriction on the possibilities of implementing permutations. If, however, the above iteration is also executed on the right side, the resulting network will be able to implement all permutations (after a possible rearranging).

Mathematically, a BASELINE network has a link topology which can be described by a permutation of bit x_0 in the binary representation of the number of the links:

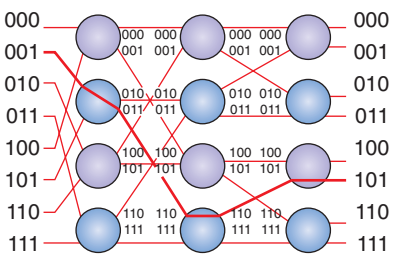


Figure 7.30 Connection through an 8 x 8 N-CUBE network

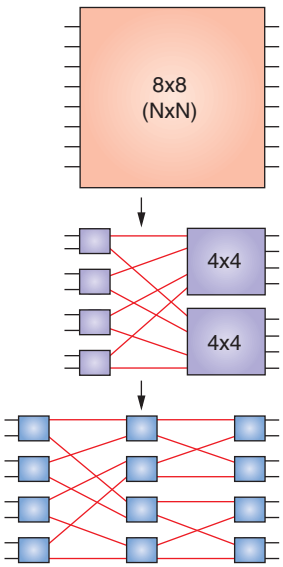


Figure 7.31 Construction of a multistage network from an N x N network

Table 7.2 BASELINE permutation for $N = 16$ ($m = \log_2 16 = 4$)

$x_3x_2x_1x_0$	$\phi_0 = x_0x_3x_2x_1$	$\phi_1 = x_3x_0x_2x_1$	$\phi_2 = x_3x_2x_0x_1$	$\phi_3 = x_3x_2x_1x_0$
0000	0000	0000	0000	0000
0001	1000	0100	0010	0001
0010	0001	0001	0001	0010
0011	1001	0101	0011	0011
0100	0010	0010	0100	0100
0101	1010	0110	0110	0101
0110	0011	0011	0101	0110
0111	1011	0111	0111	0111
1000	0100	1000	1000	1000
1001	1100	1100	1010	1001
1010	0101	1001	1001	1010
1011	1101	1101	1011	1011
1100	0110	1010	1100	1100
1101	1110	1110	1110	1101
1110	0111	1011	1101	1110
1111	1111	1111	1111	1111

$$\phi_0 = x_0, x_{m-1}, x_{m-2}, \ldots, x_1$$

$$\phi_1 = x_{m-1}, x_0, x_{m-2}, \ldots, x_1$$

$$\phi_k = x_{m-1}, \ldots, x_{m-1-k+1}, x_0, x_{m-1-k}, x_{m-2}, \ldots, x_1 \quad (7.51)$$

$m = \log_2 N$ = number of stages

$l = (x_{m-1}, x_{m-2}, \ldots, x_1, x_0)$ = binary representation of the link number

Example:

$N = 16$ gives $m = 4$, gives Table 7.2.

An illustration of a BASELINE network is shown in Figure 7.32.

All networks in Figure 7.29 can be derived from the BASELINE network by reconfiguration of the switching elements. This means that they all have the same properties with respect to connections as the BASELINE network.

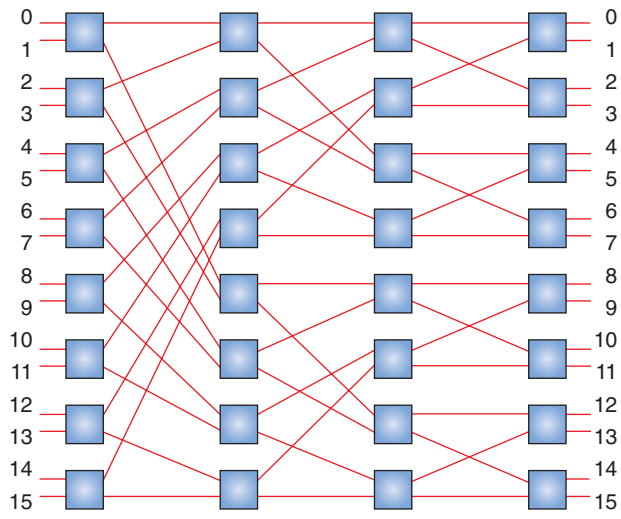


Figure 7.32 Illustration of a 16 x 16 BASELINE network

Since, as mentioned above, all these topologies have congestion, it will be of interest for some applications to reduce this drawback. Several methods are given in Figure 7.33. The topologies are named Parallel, Dilated and Repeated networks. All methods increase the number of components and also means a more complicated control.

References

1 Agrawal, D. Graph theoretical analysis and design of multistage interconnections networks. *IEEE Transactions on Computers,* 32(7), 1983.

2 Ahmadi, H, Denzel, W. A survey of modern high-performance switching techniques. *IEEE Journal on Selected Areas in Communications,* 7(7), 1989.

3 Andresen, S. *Clos permutation networks.* Trondheim, Norwegian Institute of Technology. (Dr. techn. thesis.)

4 Bassalugo, L A et al. The structure of one sided connecting networks. *6th International Teletraffic Conference,* Munich, paper number 245, equation (10), 1970.

5 Benes, V E. *Mathematical theory of connecting networks and telephone traffic.* New York, Academic Press, 1965.

6 Benes, V E. Semilattice characteristics of networks nonblocking in the wide sense. *Proceedings of the Symposium on Computer Communications Networks and Teletraffic.*

7 Benes, V E. Some proposed models for traffic in connecting networks. *Bell System Technical Journal,* 46, p. 115, Jan. 1967.

8 Berge, C. *Principles of combinatorics,* ch. 2 : Partition problems. Academic Press, 1971.

9 Chen, P et al. Interconnection networks using shuffles. *IEEE Computer,* 14(12), 1981.

10 Clos, C. A study of non-blocking switching networks. *Bell Systems Technical Journal,* 32(2), 406–424, 1953.

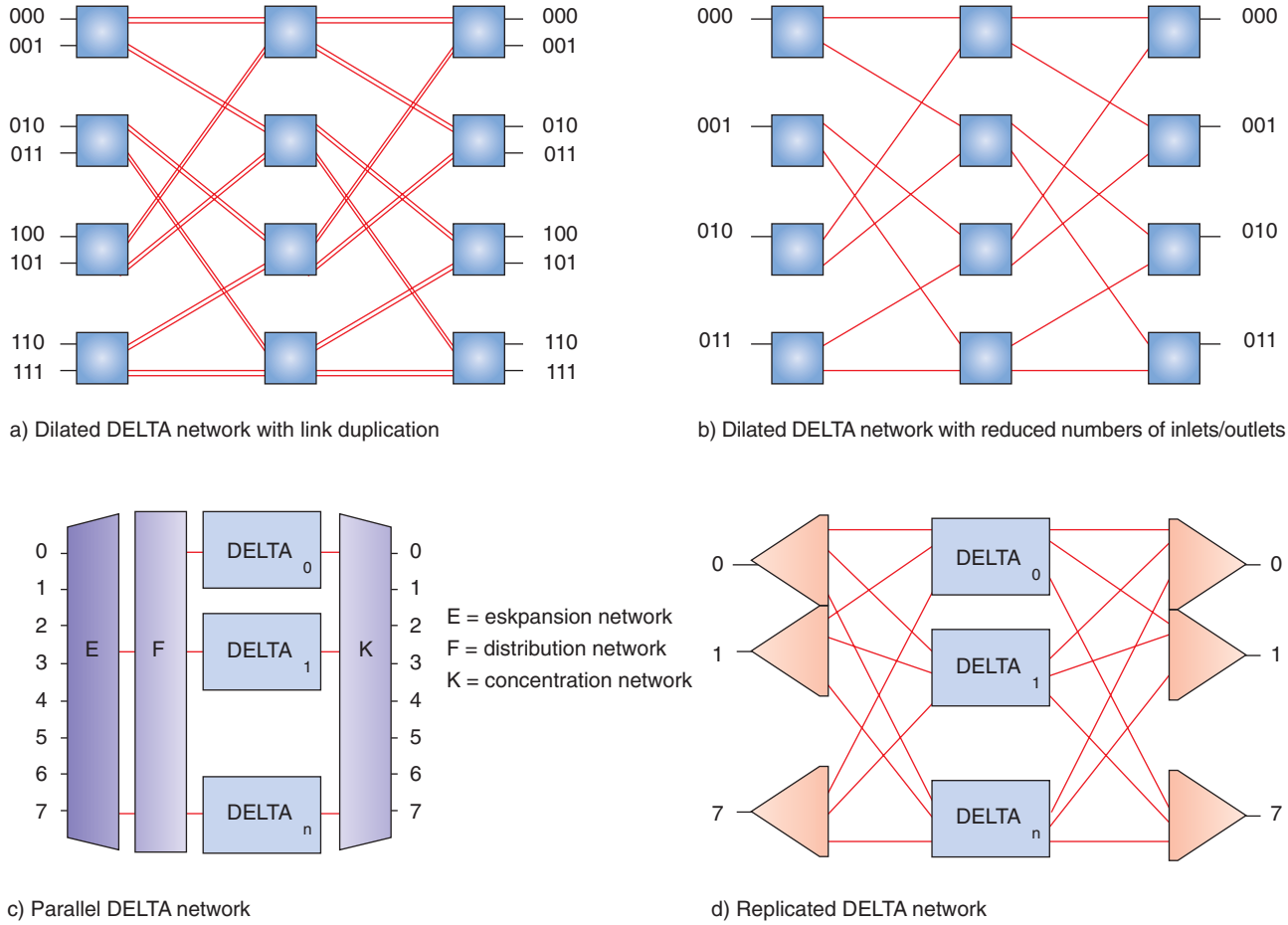


Figure 7.33 Topologies to get more connection paths

11 de Somer, M. Building a mathematical model describing the behaviour of a variable bitrate switching network. *ITC Seminar,* Brussels, May 1986.

12 Duguid, A M. *Structural properties of switching networks.* Brown University, 1954. (Brown University Progr. Rept. BTL-7.)

13 Goldstein, L J, Leibholz, S W. On the synthesis of signal switching networks with transient blocking. *IEEE Transactions on Computers,* EC-16(6), 637–641, 1967.

14 Hluchyj, M G, Karol, M J. Shufflenet : an application of generalized perfect shuffles to multihop lightwave networks. *INFOCOM '88,* 379–390, 1988.

15 Joel, J R. On permutation switching networks. *Bell System Technical Journal,* 813–822, May/June 1968.

16 Kruskal, C P, Snir, M. The performance of multistage interconnection networks for multiprocessors. *IEEE Transactions on Computers,* 32(12), 1983.

17 Lang, T, Stone, H S. A shuffle-exchange network with simplified control. *IEEE Transactions on Computers,* C-25(1), 55, 1976.

18 Lee, C Y. Analysis of switching networks. *Bell Systems Technical Journal,* 34(6), 1287–1315, 1955.

19 Lee, T T. Non-blocking copy networks for multicast packet switching. *IZS '88,* Zurich, March 1988.

20 Mirsky, L. *Transversal theory.* New York, Academic Press, 1971.

21 Myskja, A. *Multistage networks with link connections : introduction to Clos networks.* Trondheim, Norwegian Institute of Technology, 1974. (Readings for course in Teletraffic 14–19 Oct. and 4–9 Nov.) (In Norwegian.)

22 Neiman, V I. Structure et commande optimales de reseaux de connexion sans blocage. *Annal des Télécommunications,* 232–238, July/Aug. 1969.

23 Opferman, D C, Tsao-Wu, N T. On a class of rearrangeable switching networks. *Bell Systems Technical Journal,* 50(5), 1971.

24 Osatake, T, Ogawa, T, Hayashida, T. Optimum structure of one-sided rearrangeable switching networks. *Electronics and Communications in Japan,* 56-A(1), 1973.

25 Osatake, T et al. Optimization of rearrangeable switching networks. *Electronics and Communications in Japan,* 56-A(7), 1–8, 1973.

26 Padmanabhan, K, Lawrie, D H. A class of redundant path multistage interconnection networks. *IEEE Transactions on Computers,* 32(12), 1983.

27 Parker, D S, Raghavendra, C S. The gamma network : a multiprocessor interconnection network with redundant paths. *9th Annual Symposium on Computer Architecture,* Austin, April 1982.

28 Patel, J H. Performance of processor-memory interconnections for multiprocessors. *IEEE Transactions on Computers,* 30, (10), 1981.

29 Paull, M C. Reswitching of connection networks. *Bell Systems Technical Journal,* 41(3), 833, 1962.

30 Pollen, L K. Design rules for optimum non-blocking crosspoint networks. *5th International Teletraffic Conference,* New York, paper number 252, 1967.

31 Ramanujam, H R. Decomposition of permutation networks. *IEEE Transactions on Computers,* C-22(7), 639–, 1973.

32 Richards, G W, Hwang, F. A two-stage rearrangeable broadcast switching network. *IEEE Transactions on Communications,* 33(10), 1985.

33 Sakita, Y et al. A study on exchange functions and cross-connect functions on communication network. *ICC '89,* Boston, 1989.

34 Siegel, H J. Interconnection networks for SIMD machines. *IEEE Computer,* 12(6), 1979.

35 Siegel, H J, McMillen, R J. The multistage cube : a versatile interconnection network. *IEEE Computer,* 14(12), 1981.

36 Tsao-Wu, N T. On Neiman's algorithm for the control of rearrangeable switching networks. *IEEE Transactions on Communications,* COM-22, 737–, 1974.

8 Construction of space switching networks

8.1 Introduction

In this section we concentrate on the implementation of space switching networks. When implementing such networks one has to take into account a lot of parameters which must satisfy the specified limits. Some of these parameters depend on the technology and some do not. One main parameter which is normally independent of technology is the number of switching elements measured in number of inlets. Others, like transmission parameters, however, also depend on existing technology. With increasing bitrate the technology will be more and more important and this will be further discussed in a later section.

In Section 3, a space switching network is defined as a network with a fixed connection (physical or logic) between an inlet/outlet pair during the whole transfer of data. Note that this definition is also valid for binary representation of data since the "whole transfer" of data also incorporates transfer of a single bit, so that connection is established during the transfer of the bit and disconnected in the time intervals between the bits.

The most common example of a space switching network is the matrix. A switching matrix can be implemented by a switching element in every crosspoint in the matrix or by using selectors and multiply inlets and outlets. The name "matrix", however, will be reserved to a matrix of the first kind.

The operation speed can be defined as composed of two parts:

- Time to operate a single crosspoint/ selector
- Time to operate all involved crosspoints/selectors in the connection.

It is necessary to take care of both parts to get the desired specification.

8.2 Clos-(Benes-) network

The Clos-(Benes-)network has a structure as illustrated in Figure 8.1.

The structure is characterised by:

- An odd number of stages
- All switches in a stage are of the same size (and type) and can make the same assignments
- Each switch in the input stage is connected by one link to each switch in the middle stage and each switch in the output stage is connected by one link to each switch in the middle stage.

In a structure as in Figure 8.1 it is also possible to find vertical or horizontal symmetry lines.

In a previous section it was stated that with the number of inlets, *N*, greater than a certain limit, it is more favourable with respect to the number of switching elements per inlet, to divide the network into more stages.

Assume an *N* x *N* matrix network and take the following steps:

- Replace the *N* x *N* network with a 3-stage network (with *N* inlets and outlets) giving a network like the one in Figure 8.1 having $N_1 = N_2 = N$
- Replace each of the middle (central) matrices with a 3-stage network having the same number of inlets and outlets as the matrix
- Compute the total number of switching elements.

This recurrence can be repeated until the (total) number of switching elements do not decrease, or to another convenient limit. The process is indicated in Figure 8.2 for dividing an *N* x *N* network into a 5-stage network.

Another alternative is to start with a number of middle matrices and connect by a link system a set of outer matrices on both sides such that a 5-stage network is the result. Figure 8.2 can illustrate this process by starting from step 2 (i.e. after first iteration above).

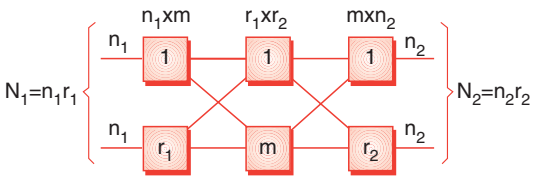


Figure 8.1 3-stage Clos network

Let s be the number of stages this process gives. It is clear that the process gives $s = 2i + 1$, $i = 1, 2, 3, 4, ...$ (i.e. an odd integer). The order of the network is defined by:

$$k = (s - 1) / 2$$

The order can be numbered from the middle stage and out to the inlet/outlet stage with middle stage as number 0.

The number of inlets N of a network constructed in accordance with the above can be expressed by:

$$N = \prod_{j=0}^{j=k} n_j \qquad (8.1)$$

where

n_j = number of inlets of each matrix in stage number.

The size of a middle matrix is given by:

$$n_0 = \frac{N}{\prod_{j=1}^{j=k} n_j} \qquad (8.2)$$

It is assumed that all the matrices in a stage have the same size.

The total number of switching elements of an s-stage network can be written:

$$K_s = 2\left[\prod_{j=0}^{j=k} n_j \cdot \prod_{i=k}^{i=k} m_i\right]$$

$$+2\left[\prod_{j=0}^{j=k-1} n_j \cdot \prod_{i=k-1}^{i=k} m_i\right] \qquad (8.3)$$

$$+2\left[\prod_{j=0}^{j=k-2} n_j \cdot \prod_{i=k-2}^{i=k} m_i\right] + ...$$

$$+2\left[\prod_{j=0}^{j=1} n_j \cdot \prod_{i=1}^{i=k} m_i\right] + n_0^2 \prod_{j=1}^{j=k} m_i$$

or short:

$$K_s = \sum_{i=1}^{i=k}\left[2\prod_{j=0}^{j=i} n_j \cdot \prod_{j=i}^{j=k} m_j\right] + n_0^2 \prod_{j=1}^{j=k} m_j$$

for $k \geq 1$ or $s \geq 3$ (8.4)

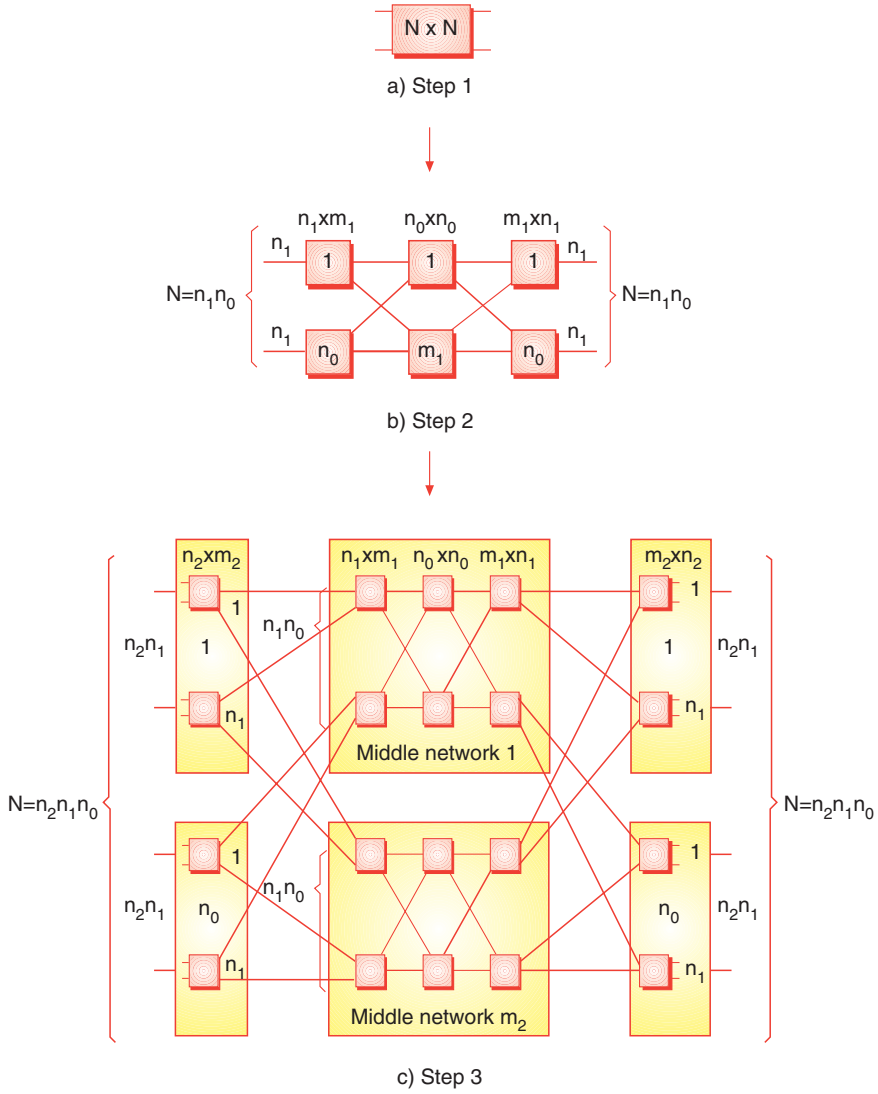


Figure 8.2 Iterative construction of a multistage network

$K_1 = N^2$ for $k = 0$

$k = (s - 1) / 2$

The number of elements per inlet = outlet is given by:

$$k_s = K_s / N \qquad (8.5)$$

No analytic method is known for computing the values of m_i's and n_j's that give fewest number of elements for a given N.

Since m_i and n_j must be integers the situation illustrated in Figure 8.3 can happen with a size of the matrices in the outer stages given by the relation:

$N = \alpha n_k + R$;
$\alpha = 1, 2, 3, ...$ and $0 \leq R \leq n_k$

This gives α matrices of different sizes. With today's technology, however, it is common that matrices have the same size and do not use empty inlets and outlets.

8.2.1 Strictly non-blocking network

In a 1-stage network of size N x N the total number of switching elements is given by setting $k = 0$:

Total: $K_1 = N^2$

Each inlet: $k_1 = N$

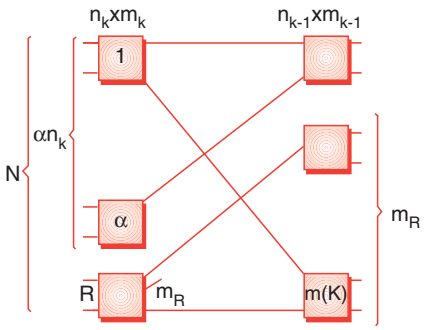


Figure 8.3 Network with matrices of different sizes

Since the total size of the switching networks for most telecommunication applications consists of several thousand inlets/outlets, it is obvious that a 1-stage network is very uneconomical.

A 3-stage strictly non-blocking network like that in Figure 8.2b has $m_1 = 2n_1 - 1$; $k = (3 - 1) / 2 = 1$ and putting this into Eq. (8.4) gives:

$$K_3 = \sum_{i=1}^{i=1}\left[2\prod_{j=0}^{j=i} n_j \cdot \prod_{j=1}^{j=i} m_j\right] + n_0^2 \prod_{j=1}^{j=1} m_j$$

$$= 2n_0 \cdot n_1 \cdot m_1 + n_0^2 \cdot m_1$$

That is: $K_3 = 2\, n_0 \cdot n_1 \cdot (2n_1 - 1) + n_0^2 \cdot (2n_1 - 1) = n_0 \cdot (2n_1 - 1) \cdot (2n_1 + n_0)$ (8.6)

Per inlet: $k_3 = (2n_1 - 1) \cdot (2n_1 + n_0) / n_1$

$n_1 = n_0 = n$ gives $n = N^{1/2}$ and:

$$\left.\begin{aligned} K_3 &= 3\left(2N^{3/2} - N\right) \\ k_3 &= 3\left(2N^{1/2} - 1\right) \end{aligned}\right\} \quad (8.7)$$

By setting $k = 2$ in Eq. (8.4), the total number for a 5-stage network is:

Total:

$$K_5 = \sum_{i=1}^{i=2}\left[2\prod_{j=0}^{j=i} n_j \cdot \prod_{j=i}^{j=2} m_j\right] + n_0^2 \prod_{j=1}^{j=2} m_j$$

or

$$K_5 = 2n_0 \cdot n_1 \cdot m_2 \cdot m_1 + 2n_0 \cdot n_1 \cdot n_2 \cdot m_2 + n_0^2 \cdot m_2 \cdot m_1 \quad (8.8)$$

To get a strictly non-blocking network we must have $m_1 = 2n_1 - 1$ and $m_2 = 2n_2 - 1$ which, by setting $n_0 = n_1 = n_2 = n$, gives:

$N = n_0 \cdot n_1 \cdot n_2 = n^3$

Total: $K_5 = 16 \cdot N^{4/3} - 14 \cdot N + 3 \cdot N^{2/3}$

Each inlet: $k_5 = 16 \cdot N^{1/3} + 3 \cdot N^{-1/3} - 14$ (8.9)

In the same way the number of elements can be found for other *s*-values. In Figure 8.4 this is summarised for some *s*- and *N*-values.

Figure 8.5 shows some simulation results for the number of switching elements as a function of *N* for different networks.

It is seen that the number of crosspoints is much lower in an optimised network. $N = 10^4$ gives 1000 crosspoints per inlet while the corresponding optimised network needs only about 300 crosspoints per inlet.

It can be shown that the upper limit on the number elements can be written:

$K = N \cdot \exp\,[2(\log N)^{1/2}]$

For a 3-stage network the lower limit of number of crosspoints can be found by searching for the extreme values of K_3. Eq. (8.6) gives:

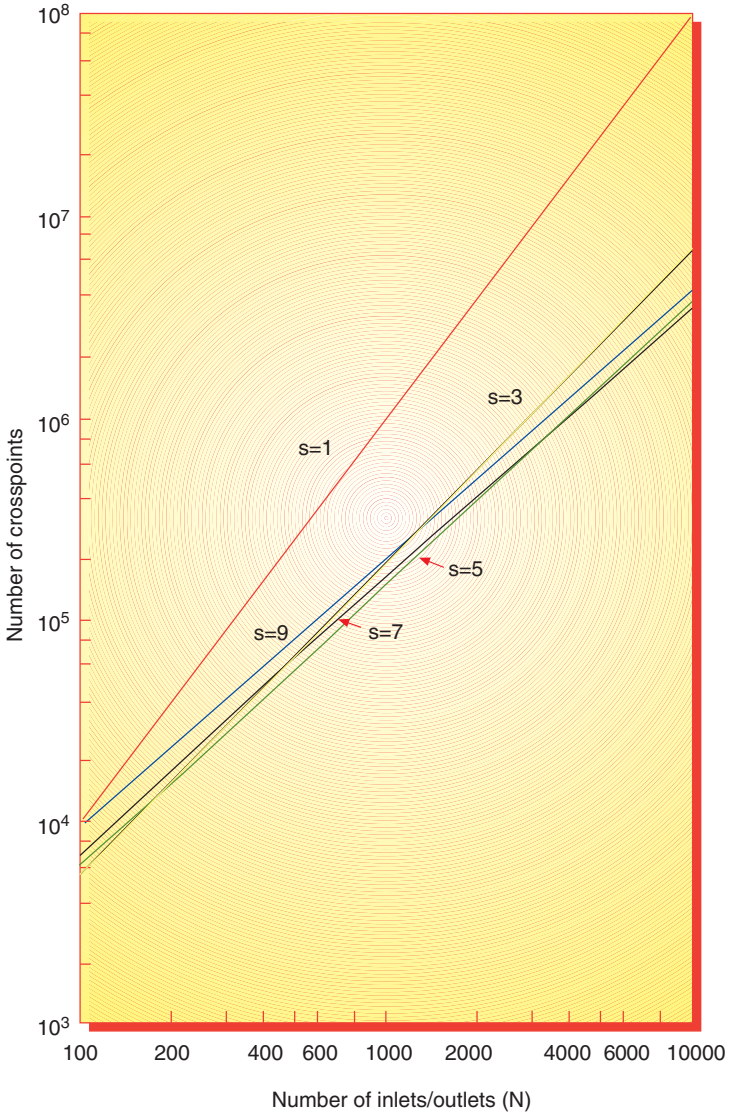


Figure 8.4 Total number of switching elements for some s-and N-values

$$\partial K_3 / \partial n_1$$
$$= \partial[n_0(2n_1 - 1)\,(2\,n_1 + n_0)] / \partial n_1 = 0$$
$$\Rightarrow \delta n_1 = 2\,n_1^3 - n_1 N + N = 0$$

By setting $N >> 1$ this approximately gives:

$$n_1 \approx (N / 2)^{1/2} \qquad (8.10)$$

From the expression of K_3 we get:

$$K_3' = 4 \cdot 2^{1/2} \cdot N^{3/2} - 4 \cdot N$$

$N = 9$ ($n_1 = 3$) gives $K_3' = 117$; ($K_3 = 135$ Eq. (8.6))

$N = 25$ ($n_1 = 5$) gives $K_3' = 608$; ($K_3 = 675$ Eq. (8.6))

In the same way extreme values can be found for higher s-values.

With given sizes of the matrices, which N-value that gives $K_1 < K_s$ can be computed. Applied on K_3 we find:

$$K_1 < K_3 \text{ for } N < [2n_1^2\,(2n_1 - 1)] / (n_1 - 1)^2$$

8.2.2 Rectangular networks

In addition to the quadratic network discussed above, we have an important group of rectangular networks with or without expansion/concentration. These networks have $N \neq M$ where N is the number of inlets and M is the number of outlets.

Figure 8.6 shows an example of a 3-stage network with $N \neq M$. The number of elements is given by:

$$K_{3NM} = n_1 \cdot m_1 \cdot r_1 + r_1 \cdot r_2 \cdot m_1 + m_1 \cdot n_2 \cdot r_2 \qquad (8.11)$$

With $N = n_1 \cdot r_1$ and $M = n_2 \cdot r_2$ this gives:

$$K_{3NM} = (n_1 + n_2 - 1) \cdot (N + M + N \cdot M / (n_1 \cdot n_2) \qquad (8.12)$$

The network in Figure 8.6 can be optimised with respect to n_1 and n_2 which gives:

$$\partial[K_{3NM}] / \partial n_1 = 0 \text{ and}$$
$$\partial[K_{3NM} / \partial n_2 = 0.$$

$n_1 = n_2 = n$ gives minimum:

$$K_{3NM} = (2n - 1) \cdot [N + M + N \cdot M / n^2]$$

$\partial[K_{3NM}] / \partial n = 0$ gives:

$$n^3 - n \cdot (N \cdot M / (N + M)) + N \cdot M / N + M = 0 \qquad (8.13)$$

$N = M$ gives the same equation as above.

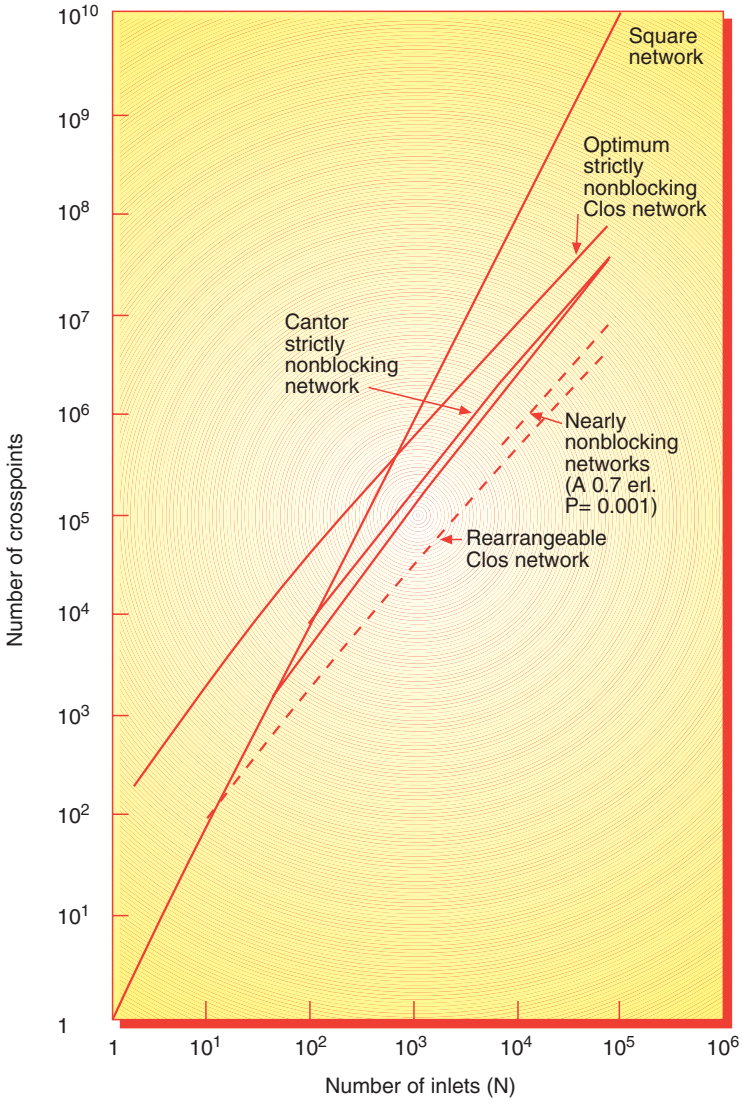


Figure 8.5 Total number of switching elements in different networks

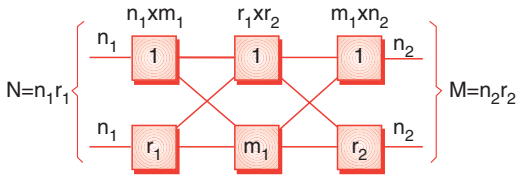


Figure 8.6 3-stage rectangular network

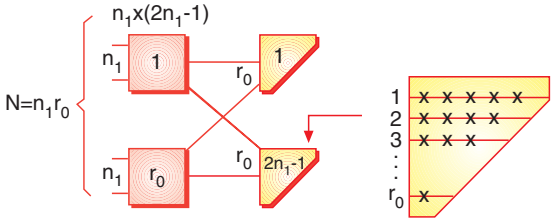


Figure 8.7 2-stage triangular network

8.2.3 Triangular networks

Triangular networks are of interest when one-sided networks have to be implemented. An example of such a network is shown in Figure 8.7.

Total number of elements for this network can be written:

$$K_{3E} = (2n_1 - 1) \cdot (N + (N^2 / (2n_1^2)) - N / 2n_1) \quad (8.14)$$

8.3 Network being non-blocking in the wide sense

A network being non-blocking in the wide sense is given from Eq. (8.4) by setting:

$$m_i = \lfloor 3n_i/2 \rfloor$$

A 3-stage network gives $m_1 = \lfloor 3n_i/2 \rfloor$

and a 5-stage network gives

$m_1 = \lfloor 3n_1/2 \rfloor$ and $m_2 = \lfloor 3n_2/2 \rfloor$

The same equation gives the total number of switching elements.

$N = 25$ (inlets/outlets) in a 3-stage network gives:

$$n_0 = n_1 = 5;$$

$$m_1 = \lfloor 3 \cdot 5/2 \rfloor = 7$$

$$K_{3W} = 2n_0 \cdot n_1 \cdot m_1 + n_0^2 \cdot m_1 =$$

$$\lfloor 15/2 \rfloor (2 \cdot 5 \cdot 5 + 25) = \underline{525}$$

A 1-stage strictly non-blocking network gives:

$$K_{3S} = N^2 = 25^2 = \underline{625}$$

The discrepancy between these will increase with increasing N.

8.4 Rearrangeable networks

A rearrangeable network is given from Eq. (8.4) by setting:

$$m_i = n_i$$

The same equation as above gives the total number of switching elements. $N = 25$ and $s = 3$ gives with $n_0 = n_1 = m_1 = 5$:.

$$K_{3R} = 2n_0 \cdot n_1 \cdot m_1 + n_0^2 \cdot m_1$$
$$= 2 \cdot 5 \cdot 5 \cdot 5 + 5^2 \cdot 5 = \underline{375}$$

It is seen that the number of switching elements is decreasing from strictly non-blocking networks to rearrangeable networks.

Bibliography

1 Braugenhardt, S, Jönsson, I, Nilsson, B. Introduction of the AXE 10 switching system in the telephone network. *Ericsson Review,* 53(2), 60–69, 1976.

2 Buchner, R B, Milort, W. The switching network of Philips' PDX system. *Philips Telecommunication Review,* 31(3), 1973.

3 McMillen, R J. A survey of interconnection networks. *Proceedings ISS '84,* Florence, 7–11 May, 5.1.1–5.1.9, 1984.

4 Myskja, A. *Minimization of the number of crosspoints in recursive connection networks.* Trondheim, Norwegian Institute of Technology, March 1971. (Internal report.)

5 Myskja, A. *Multistage networks with link connections : introduction to Clos networks.* Readings from course in Teletraffic 14–19 Oct. and 4–9 Nov. 1974 at the Norwegian Institute of Technology. (In Norwegian.)

6 Pollen, L K. Design rules for optimum non-blocking crosspoint networks. *5th International Teletraffic Conference,* New York, paper number 252, 1967.

7 Ramanujam, H R. Decomposition of permutation networks. *IEEE Transactions on Computers,* C-22(7), 639–, 1973.

9 Construction of time switching networks

9.1 Introduction

In Section 3 a time switching network was defined as a network which connects an inlet/outlet pair in a fixed time interval and disconnects the same pair in all other time intervals. To implement a time switching network three components are defined.

These three components are (see Figure 9.1):

- The static multi- and demultiplexer here denoted MUX and DMUX
- The space-time and time-space switches here denoted S-switches, i.e. a space switch that switches in defined time slots
- The time switch is here denoted T-switch, i.e. a switch that permutes the contents in the time slots. This is normally the component that executes the switching and is synonymous to time switching.

MUX and DMUX can be considered as serial to parallel and parallel to serial converters. In addition, it is necessary to have units that are able to store data, i.e. to store digital data in a digital system and analogue data in an analogue system.

The S-switches can be viewed as made of one or more programmable MUX or DMUX used to convert from time to space or from space to time. An S-switch in the form of MUX/DMUX is necessary since the inlets and outlets are distributed in space.

9.2 Restrictions on the parameters

Since the time is introduced as a new parameter in time switching, it is necessary to examine if this also introduces

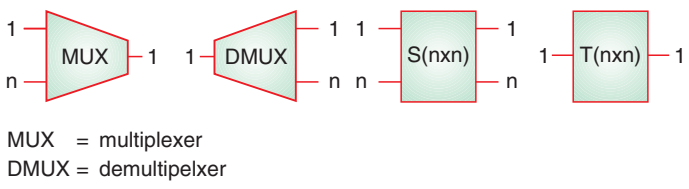


Figure 9.1 Symbols of functional units in a time switching network

any restrictions and try to find solutions of the problems that may arise. With data input as indicated in Figure 9.2, the possible connections this system is able to give can be written as:

$$K \leq \lfloor T_0/(b \cdot \tau_0) \rfloor \qquad (9.1)$$

where

T_0 = time between two succeeding data sets from the same terminal

b = number of bits in a data set

τ_0 = time for transferring one bit.

Example:

T_0 = 488 ns = duration of one bit in a first order PCM system (b = 1)

τ_0 = 100 ns gives $K \leq 4$

τ_0 = 10 ns gives $K \leq 48$

τ_0 = 1 ns gives $K \leq 488$

τ_0 = 100 ps gives $K \leq 4880$

K in the expression above is a measure of number of terminals N the network is able to serve. The relationship between K and N is generally traffic dependent and an expression for the relationship between the two variables can be written as:

$$N = \alpha \cdot K$$

where

α is a traffic dependent concentration factor. Practical values of α can be in the range 4 – 8.

To solve the problem, one can use a more suitable technology or combine space and time switching (see later sections).

The parameter T_0 is related to the behaviour of the data the terminals are generating. To satisfy Shannon's theorem, the transmission of voice with digital representation gives T_0 = 125 µs, transmission of TV-signals with bandwidth 5 MHz gives T_0 = 100 ns, and so forth. In modern telecommunications systems the base time unit is T_0 = 125 µs.

The number of bits in each transfer b, may vary widely. For example, b = 8 bit in the new digital system and b = 424 bit in ATM. A standard base is 8 bit = 1 byte = 1 octet.

The parameter τ_0 is related to the technology used. This is the basic parameter mostly influenced by the development of the technology. Today's components give τ_0 in the nanosecond (ns) region, while it is expected that future components will reduce this figure to the picosecond (ps) region.

Following the above the two parameters T_0 and τ_0 will place an upper limit on the "pure" time switching, see Figure 9.2.

For the first order PCM system we have the basic parameters T_0 = 125 µs, b = 8 bit. With τ_0 = 12.5 ns this gives:

$K \leq 1250$

Some higher order PCM systems give, when $b\tau_0$ = 100 ns:

- Second order system:
  T_0 = 31.25 µs giving $K \leq 312$
- Third order system:
  T_0 = 7.81 µs giving $K \leq 78$
- Fourth order system:
  T_0 = 1.95 µs giving $K \leq 19$.

Total bitrate is: $B = b / T_0$ bit/sec

User bitrate is: $B_b = b / T_0$ bit/sec

The user bitrate (with serial transmission) of first order PCM system is:

$B_b = b = 8 / 125 \cdot 10^{-6} = 64$ kb/s.

The necessary operation speed for the elements involved in the transfer can be written:

$\tau_0 \leq 1 / B$

9.3 Increasing the network capacity

To increase the capacity (measured in number of terminals or in number of "simultaneous" connections) beyond the limit given by eq. (9.1) one can use one of the following methods or their combinations:

- Parallel transfer
- Broadcasting
- More time switches connected by
  - bus system
  - space switches.

9.3.1 Parallel transfer

Transfer of p bit in parallel increases the number of terminals p times. This can be used to increase the capacity of the switch. The capacity can be written as:

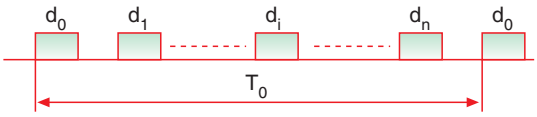


Figure 9.2 Time parameters in time switching

$$K \leq \lfloor p(T_0/b\tau_0) \rfloor$$

where, in addition to eq. (9.1):

p = number of bits transmitted in parallel

Examples:

T_0 = 125 µs, $b\tau_0$ = 100 ns and
p = 32 bits in parallel, gives:
$K \leq 40000$

T_0 = 125 µs, $b\tau_0$ = 100 ns and
p = 8 bits in parallel, gives:
$K \leq 10000$

This method demands that p bits have to be received and stored before switching is executed and that storing capacity is increased with the requested value.

Total bitrate, B, is increased p times.

9.3.2 Broadcasting

Another method to increase the capacity of the switching network or to decrease the demand on the operation time of the switching components, is to use broadcasting. This is illustrated in Figure 9.3 and means that data together with the address of the receiver are broadcast to all receivers. The address is recognized by the receiver and the accompanying data are read by the receiver.

The access times for storing data is the same as for this operation without broadcasting. When a frame of data is stored, all data storage locations can be read. This gives the same value for read and write times for all storage operations.

The capacity of the storage is:

D_{data} = $2s^2nb$, bit (dual data storages)

$D_{control}$ = $sn \cdot \log_2(sn)$, bit

where

s = number of time switches
n = number of time slots
b = number of bits in each time slot.

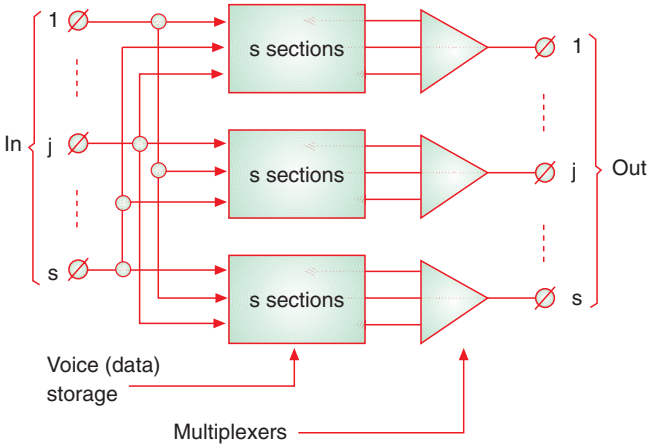


Figure 9.3 Principle of broadcasting to all memory sections

The drawback of this principle is, as the above relation shows, that the necessary capacity of the data storages increases quadratically as a function of time switches.

There are, however, other basic physical conditions that can possibly be a problem. One of the more important ones is the time delay in a physical transmission system.

A simple example is illustrated in Figure 9.4 which shows N terminals connected to a common transmission line with a switch for each terminal. During time slot t_0 inlet 1 has to be connected to inlet 3 and during time slot t_1 inlet i has to be connected to inlet N. To avoid interference, time slots t_0 and t_1 must be separated as the data from inlet 1 must pass before the switch belonging to inlet i is closed.

In a T-switch the switching gives a delay of the data from an inlet to an outlet. The delay will vary from 1 time slot to 1 (or more) frame.

Total allowed delay through an exchange is specified by CCITT to be within the limits of 1500 – 3850 µs. More specifically:

Digital – digital connection:
Δ_{avg} = 900 µs, Δ_{max} = 1500 µs

Digital – analogue connection:
Δ_{avg} = 1500 µs, Δ_{max} = 2100 µs

Analogue – analogue connection:
Δ_{avg} = 3000 µs, Δ_{max} = 3850 µs.

Normally, a fixed delay does not represent any problem, but delay variations can be serious. For this reason it is favourable to keep the transmission lines (within the switching network) of same lengths.

9.3 Synchronization

When incoming data to a time switch are serial binary coded it is necessary to have synchronism between the incoming bitrate and the bitrate in the switch. Figure 9.5 illustrates two terminals where, for one of them, data are displaced compared with the internal time generator CP_0 of the switch. Data in time slot t_0 allocated in terminal j is transferred to the right terminal during time slot τ_0, while data from terminal i in time slot t_0 is transferred during time slot τ_1. This data are "identified" by the demultiplexer as belonging to time slot t_1 and are therefore transferred to the wrong terminal.

In a practical switching network with more exchanges, the inlets represent different TDM systems from other exchanges. Without special effort, the bits in the input signals will not be synchronous, i.e. they have different phases relative to each other and to the internal clock in the switching network. The failures caused by this are lack of bits or bits ending in the wrong time slot (channel). The reason for this can be:

- Different frequencies of the clocks due to frequency variations
- Tolerances of the components
- Different delays due to unequal transmission lines
- Different delays in the switching elements the data have passed
- Jitter due to variation in functional units. (Jitter is very low fluctuation (< 0.01 Hz) due to variation in the clock frequencies caused by temperature variation, different transmission delays, etc.)

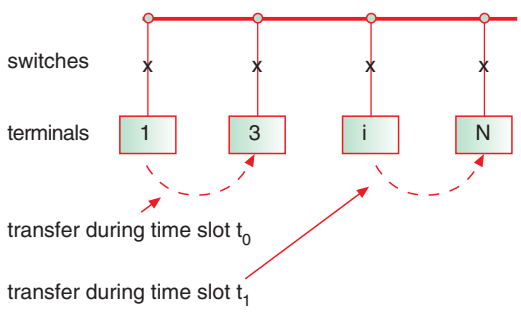


Figure 9.4 Time relations in a 1-stage time switch

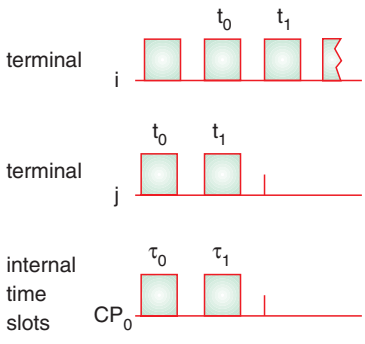


Figure 9.5 Illustration of failure due to unsynchronism

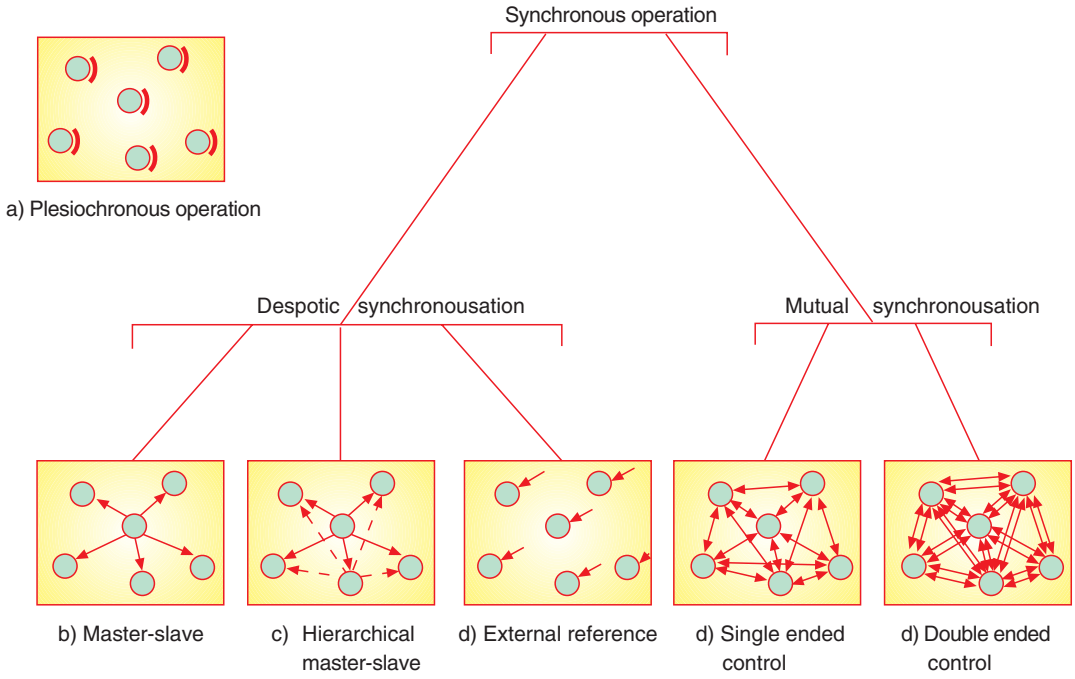


Figure 9.6 Methods of synchronization

To counteract the reasons for failure, three main methods can be used:

- The plesiochronous method which means that the clocks in every exchange are very stable and precise (short and long time stable)
- The synchronous method with all exchanges having the same time reference
- A mixture of the plesiochronous and synchronous method.

The synchronization can be achieved as indicated in Figure 9.6 by:

- The principle of master and slave
- Hierarchical master-slave
- External reference
- Mutual synchronization with single ended control
- Mutual synchronization with double ended control.

Before the synchronization can begin, it is necessary to find the start of a frame so that the succeeding time slots can be identified.

The process starts at “lowest level” by extracting the bitrate from incoming bit streams. The relation between the time slots is determined by counting the bits in every time slot. To indicate the start of a new frame, a special code is placed in the framing slot, see Figure 9.7. When this

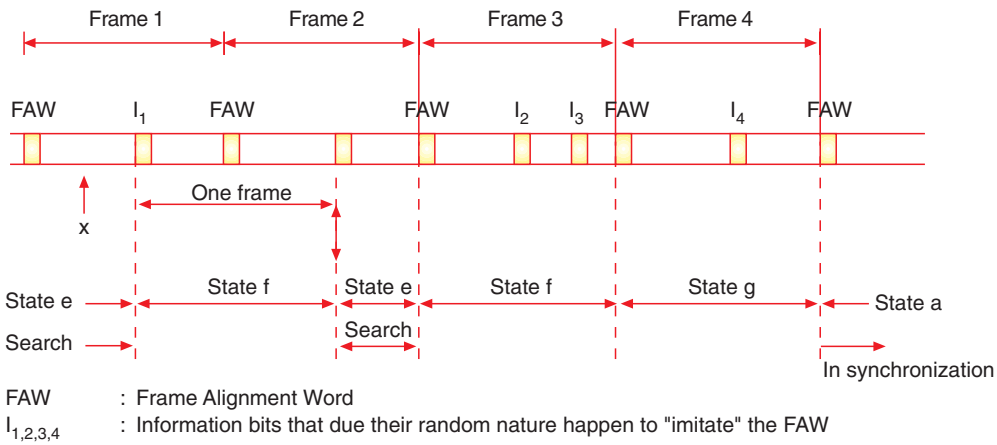


Figure 9.7 Example of frame keying

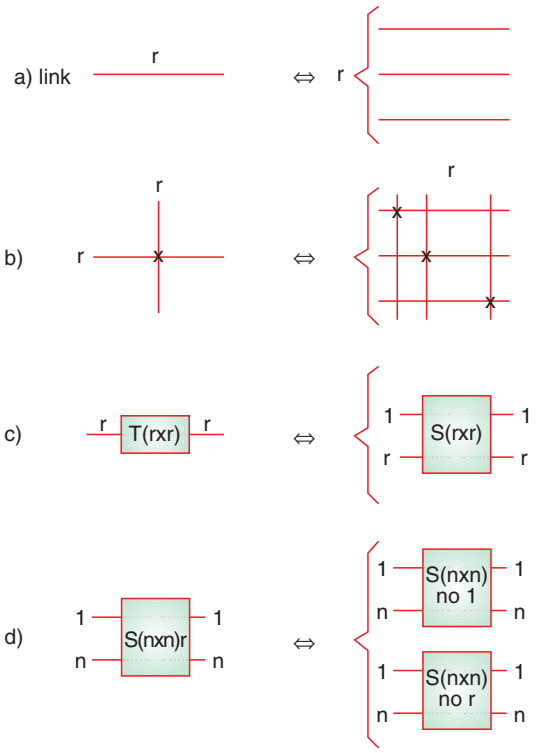


Figure 9.8 Some space – time equivalents

code is detected, we know that the time slots between this and next time slot with the same code belong to the same frame. The code can be called "frame key" or Frame Alignment Word (FAW).

A frame can thus be defined as the number of time slots between two succeeding time slots with the same frame key.

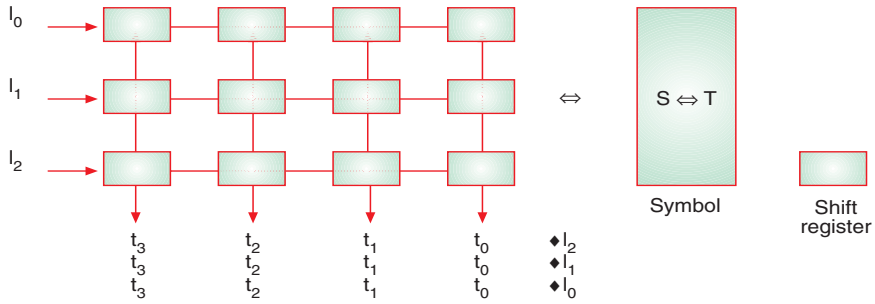
In first order PCM systems time slot 0 is used for the frame key. Higher order PCM systems have more basic frames multiplexed on the same inlet, and in this case a multiframe key is also used.

To find the time slot with the frame key, serial or parallel searching can be used. An example of serial searching is illustrated in Figure 9.7 which shows four frames.

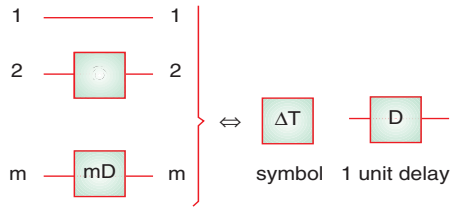
The bitrate is supposed to have been extracted from incoming bit streams. Assume that the frame key word consists of 8 bits. The bit pattern after receiving the first 8 bits is compared with the frame key. If the pattern is not equal to the key, the first detected bit is skipped and a new bit is taken up. The new received word is then compared to the frame key. Assume that the content I_1 in a time slot imitates the frame key word, i.e. a frame key is found. The process continues to count the number of bits and then the number of time slots. After the specified number of time slots it tests whether or not the time slot contains the same key word. Since this is not the case, the search for the right key word starts again. Assume that the right key word is found after two frames. The counting of time slots starts again and if the right frame key is found after 4 – 5 frames, it is concluded that the right frame key was found. Note that when the right frame key word is found, the imitation of frame keys in the other time slots has no influence on the synchronization.

After detecting the start of a frame and identifying the time slots the synchronization remains. The first step is to synchronize the bitrate of the switching network and the incoming bitrate. The next step is to synchronize the frame and time slots.

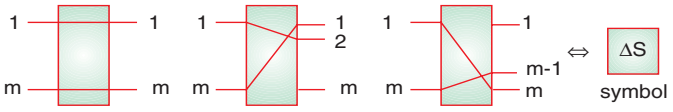
In general, all incoming lines will be unsynchronized, and it is favourable to synchronize all lines so that the switching network "sees" all lines equally, i.e. the start of a frame is equal for all lines. The common method to solve this problem is buffering of the inlet data. It uses two data storages and changes reading and writing from the storages from one frame to the succeeding frame.



a) Time-space transformer



b) Time rotator



c) Space rotator

Figure 9.9 Functional units for space/time – time/space conversion

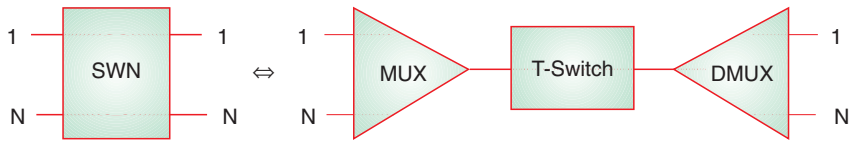
9.4 Space – time equivalents

Figure 9.8 shows the more important space – time equivalents. A link or transmission line multiplexed with r time slots is equal to r parallel single wires in space, i.e. a space bus with r lines (wires).

A time switching element A (1 x 1 matrix) with r time slots in and r time slots out is equivalent to a space matrix with switches in the diagonal of the matrix. A time switch able to permute the content in r time slots is equivalent to a quadratic r x r-matrix.

A space matrix or S-switch working in r time slots is equivalent to a set of r matrices with the same size as the space matrix.

Other important equivalents are units to convert from space to time and vice



a) Space switching network b) Time transformed network

Figure 9.10 TS-network with space / time / space transform

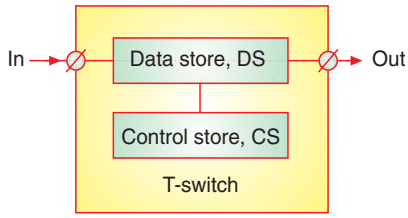


Figure 9.11 Functional units of a T-switch

versa. An example of such a unit is the serial-parallel and parallel-serial converting executed by multiplexers and demultiplexers. Figure 9.9a shows an example of time/space converting of three TDM-lines each with four time slots. One shift register is used per time slot and data are shifted in the register during a frame. After each frame data is shifted vertically out and the result is that the data in the four time slots are distributed on four outlets. Figure 9.9b shows a time rotator constructed by time delays, and Figure 9.9c shows a space rotator which in principle consists of a multipole (m-pole) switch generating one set of connections in each stage between the inlets and the outlets.

The two last units can be used to construct multistage network with only one type of switches (time or space switches).

Other necessary elements are elements for storing data. This need for storing data sometimes makes is difficult to find the exact equivalence between time and space switching networks.

9.5 Switching networks based on time (T-)switches

The pure time switching (T-switching) permutes the content in the time slots without any form of space distribution. In a practical system a space – time transform must be executed before the switching and time – space transform after the switching. This is illustrated in Figure 9.10. The data on the inlets in Figure 9.10 are multiplexed on a common line and appear in sequence. The T-switch then permutes the data, and a demultiplexer distributes the data in space.

The storage units "inside" the T-switch are illustrated in Figure 9.11.

Data on the common input line are stored in DS. The addresses to DS are given from CS and then the permutation can be executed. It is clear that the access time of the stores is very important. The following gives some examples:

- 1 first order PCM system with 32 time slots gives ca. 3906 ns per time slot
- 4 first order PCM systems with 128 time slots give ca. 978 ns per time slot
- 16 first order PCM systems with 512 time slots give ca. 257 ns per time slot
- 64 first order PCM systems with 2048 time slots give ca. 61 ns per time slot
- 256 first order PCM systems with 8192 time slots give ca. 15 ns per time slot

Frame and 8 bit parallel switching have been assumed. These figures show that time is very important and that technological restrictions soon arise.

Important characteristics of a T-switch are then:

- The bitrates on the inlets give the access time for the storages
- The number of single channels give the necessary capacity for the storages.

The demand on access time can be reduced by transferring more bits in parallel. Assume m bytes (1 byte = 8 bits) are stored per channel. By using frame switching and parallel storing this gives a total of $2m$ data storage locations and a demand on access time of:

$$\tau_a = m\tau_0$$

where

τ_0 = access time with one pair of data storages.

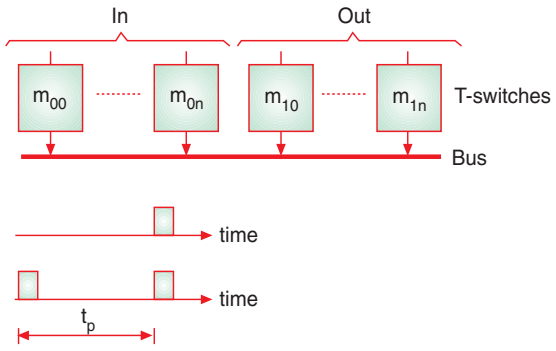
The capacity for T-switching can be extended by using more T-switches and the broadcasting principle mentioned earlier.

9.6 Networks based on T-switching and bus system

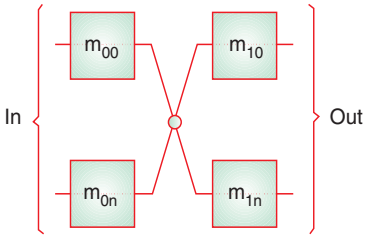
Combining T-switches and bus systems extends the applicability of time switching. At the same time suitable modules of switching networks can be composed.

An example of such a combination was shown in Section 2 and, with another bus layout, repeated in Figure 9.12a. Several T-switches are connected to a common bus and such networks can be characterised by:

- A common transmission media between the modules
- Sequential access to the bus



a) "One-line" bus system



b) "Distributed" bus system

Figure 9.12 T-switched bus system

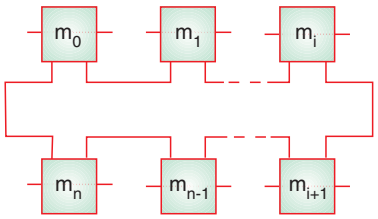


Figure 9.13 Ring bus

- Parallel transfer of data
- Time delay
- High bitrate on the bus.

Internal traffic, i.e. connection of terminals belonging to the same module, can be switched without going out on the bus and so decrease the bitrate on the bus.

The modules can operate in an autonomous way or can be controlled from a central unit. In a practical system there will be a combination of autonomity and central control.

The bus can be divided into separate data, address and control buses, or a time shared bus working in the sequence: control→ address → data, can be used.

In a bus system the number of time slots must be equal to the number of modules connected to the bus giving the relation between the bus bitrate, $1/t_B$, and the number of modules, m:

$t_B = t_0 / m$,
t_0 = duration of a time slot.

Example:

First order PCM system, parallel transfer

τ_k = 3.9 µs = time slot duration

n_k = 32 = number of channels

N = number of terminals

b = 1 Gb/s, bus bitrate

Sequence: control → address → data:

$N \sim 32 \cdot 3900 \cdot 10^{-9} / (3 \cdot 10^{-9})$
= 41600 terminals (subscribers).

By using an optical bus higher bitrates and values of N can be achieved.

Allowed time delay determines the physical distance between the modules. This will also determine the highest bitrate as illustrated in Figure 9.12a. A sending module has to wait until the data from a preceding module has "passed". The highest bitrate is $1 / t_p$ where t_p is the propagation time from module m_{00} to module m_{0n}. In order to handle varying time delays between the modules, a possible solution is illustrated in Figure 9.12b where the bus is divided into smaller buses with the same length and all passing a common point.

Another solution is illustrated in Figure 9.13 where a ring bus is used. Data are received and retransmitted by each module and always "goes" in the same direction.

The data on the bus are identified by the addressed module and read by the module if it is intended for this module. If not, the data are transmitted to the succeeding module. When data arrive at the right module these data (together with the address) can be removed from the bus or circulate on the bus until the sending module is reached, often with a label showing that the data have been received by the right module. The name "slotted ring" is often used for such a usage of the bus. A "token ring" is used in a similar configuration where one module at a time gets a password and is allowed to send data.

An example of the combination of a T-switch and a bus is the basic switching used in ITT 1240 as illustrated in Figure 9.14. The switch is constructed of 16 duplex digital ports communicating over an internal bus. The switching network is intended to switch first order PCM systems and works internally with double speed and with a 39 bits wide bus. The T-switch is called a port.

In Figure 9.15 the basic switching component is shown in more detail. Each T-switch has a sending and a receiving part. Input/output data are serial bit streams with nominal first order PCM bitrate. Data from an input (sending) module are broadcast on the bus and identified by the right receiving module. As shown in Figure 9.16 the bus sequence is divided into four phases:

- Phase *P* where port, channel and command are sent on the bus
- Phase *D* where data are sent on the bus and recognized by the right port

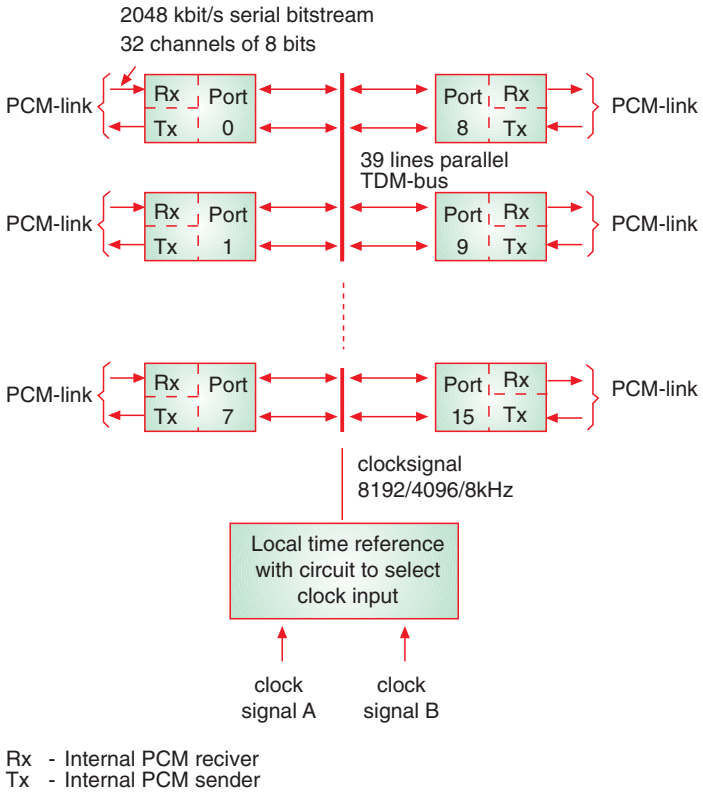


Figure 9.14 Basic switching component in ITT 1240

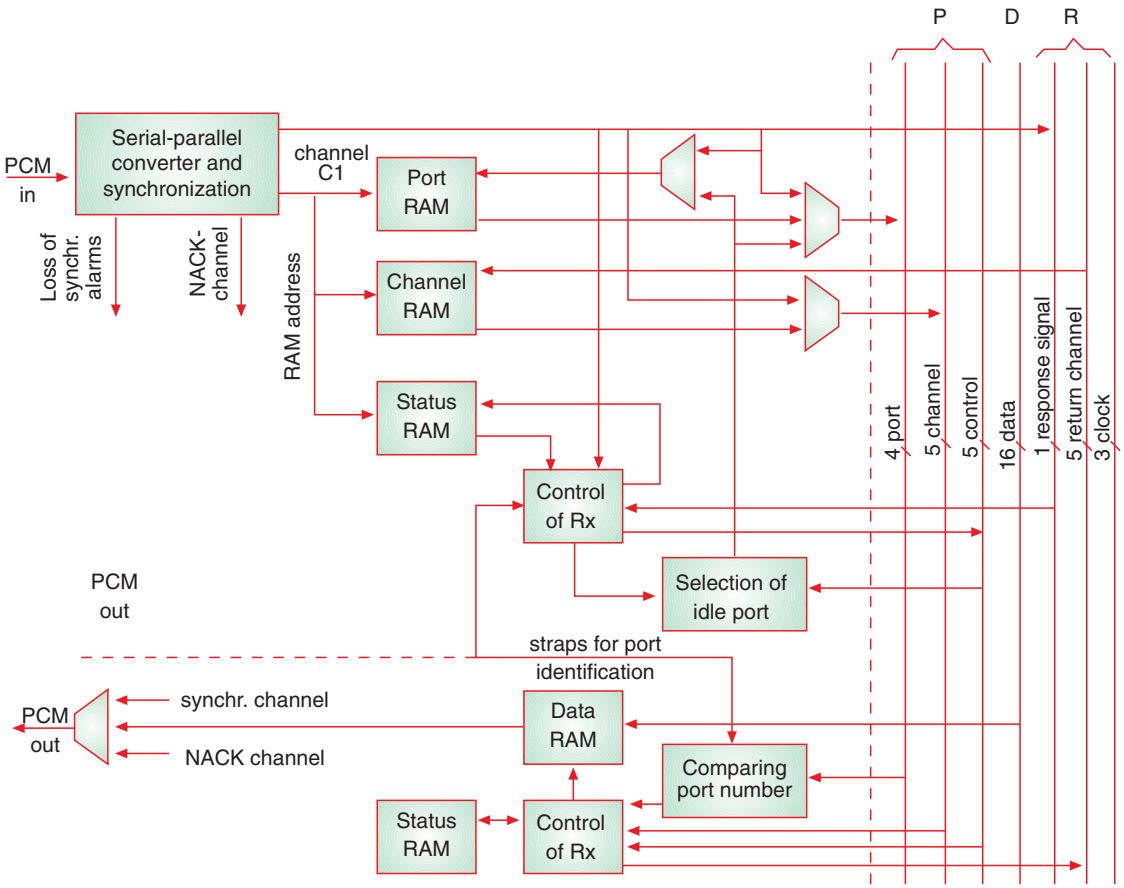


Figure 9.15 Units in the basic switch in ITT 1240

- Phase *W* where data are written into the receiving port
- Phase *R* where response can be returned to the sending port.

Every data transport has this sequence and all operations have to be executed during one time slot (3.9 µs). As indicated in the figure, the phases are used with an overlap to decrease the demand on high operation times of the components.

All data concerning the connection (such as port (T-switch) and time slot) are stored in the control storage of a port during the set-up of the call.

For more information about ITT 1240 see articles in *Telektronikk* 80(2/3), 1984.

The combination of T-switches and a bus system can be extended to a system of buses and T-switches as illustrated in Figure 9.17 for a two stage bus system. The first stage consists of n groups each with A T-switches and one bus. These groups are in stage 2 connected to n T-switches. These are connected with one bus. The process when data have to be transferred from T_{0A} to T_{n1} (both in main group A) is:

- Data from T_{0A} are stored in the associated T-switch T_{11} in group B
- From T_{11}, data are transferred to T_{1n} in main group B
- From T_{1n}, data are transferred to T_{n1}.

The process needs three cycles, but this number can be reduced by applying time sharing.

9.7 Combination of time and space switching

Instead of combining T-switches and bus systems, a time multiplexed space switching matrix can be used so that each input T-switch module gets access to the output T-switch modules. The principle is illustrated in Figure 9.18 where n switches, each with r time slots (channels), can be connected to n output switches by an S-switch of size n x n working with r time slots. This gives a 3-stage network as mentioned in an earlier section. Congestion in the network arises when two T-switches in the input stage simultaneously try to "reach" the same T-switch in the output stage. By changing the number of time slots r in the S-switch the configuration also changes connection properties:

- If $r = n$, the network is rearrangeable
- If $r = \lfloor 3n/2 \rfloor$, the network is blocking in the wide sense
- If $r = 2n - 1$, the network is strictly non-blocking.

A network with T-switches in the input and output stages and an S-switch in the central stage is often denoted a TST network.

Instead of the TST combination one can use an STS combination illustrated in

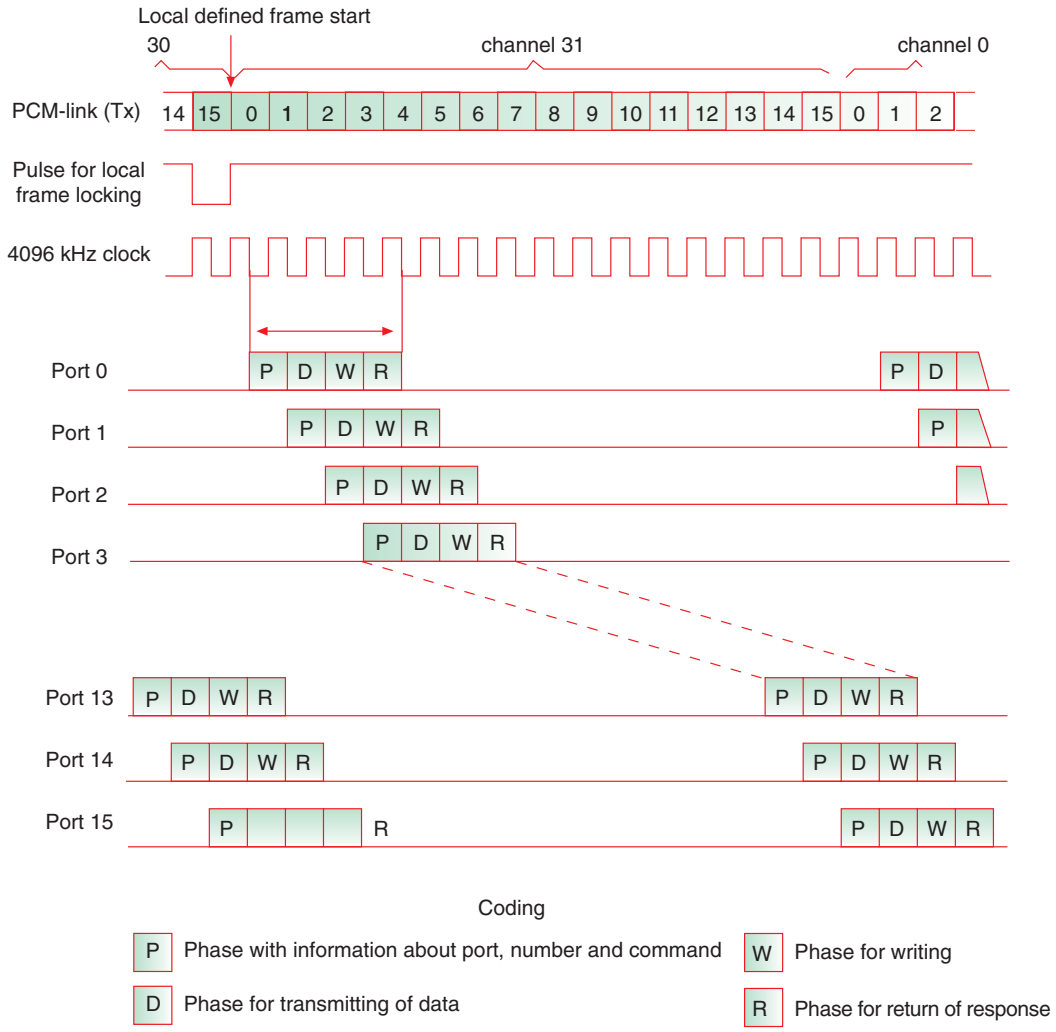


Figure 9.16 Time sequence for operation of the basic switching element in ITT 1240

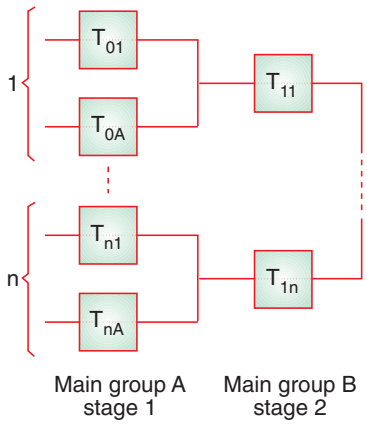


Figure 9.17 T-switches and multibus system

Figure 9.19. An S-switch of size n x n and working with r time slots on each side connects the inlets and outlets to n T-switches, each with r time slots. In this configuration the number of T-switches must be increased to satisfy the condition of strictly non-blocking ($2n$ - 1), widely non-blocking [$3n$ / 2] and rearrangeable networks.

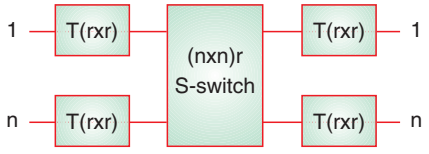


Figure 9.18 Combination of time and space switches, TST-network

A network like that of Figure 9.18 or Figure 9.19 can be seen as an equivalent to a space switching network. In general, space and time switching networks can be deduced from each other by a space/time transformation. The time and space rotator can be used for this purpose. An example is given in Figure 9.20 which illustrates:

- An input stage consisting of r matrices each of size n x m (n inlets and m outlets)
- An output stage consisting of r matrices each of size m x n (m inlets and n outlets)
- A central stage consisting of m matrices of size r x r.

The network has $n \cdot r$ inlets = $r \cdot n$ outlets.

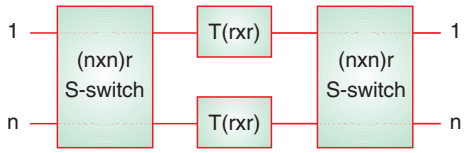


Figure 9.19 Combination of space and time switches, STS-network

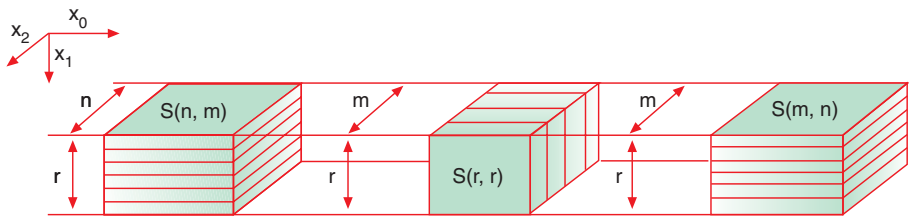


Figure 9.20 A 3-stage network

By multiplexing along different axes this network can be transformed into:

- A TST network by multiplexing (see the figure) along the x_2-axis. This gives r T-switches in the input stage with n time slots in and m time slots out, in the central stage, an S-switch of size r x r working with m time slots, and r T- switches in the output stage with n time slots out.
- An STS network by multiplexing along the x_1-axis. This gives in the input stage, an S-switch of size n x m working with r time slots, in the centre stage, m T-switches with r time slots, and in the output stage, an S-switch working with r time slot.

The result is shown in Figure 9.21.

By using the space – time transforming shown in Figure 9.9, a TST configuration can be transform to a T^3-network with only T-switches and space-time transformers as illustrated in Figure 9.22. Similarly, an STS network can be transformed to an S-network as illustrated in Figure 9.23.

Still higher orders of such networks can be constructed by using more stages. Such networks, however, are not yet in practical use.

Bibliography

1 Ahmadi, H et al. A high-performance switch fabric for integrated circuit and packet switching. *Infocom'88,* New Orleans, March 1988.

2 Andresen, S. *Clos permutation networks.* Trondheim, Norwegian Institute of Technology. (Dr. techn. thesis.)

3 Andresen, S, Harrison, S R. Toward a general class of time-division-multiplexed connecting networks. *IEEE Transactions on Communications,* COM-20(5), 836–846, 1972.

4 Anido, G J, Seeto, A W. Multi-path routing techniques for a class of multistage fast packet switches. *IZS '88,* Zurich, March 1988.

5 dePrycker, M, Bauwens, J. A switching exchange for an asynchronous time division based network. *ICC'87,* Seattle, June 1987.

6 Devalt, M, Quinquis, J P, Rouaud, Y. Asynchronous time-division switching : a new concept for ISDN nodes. *Proceedings ISS '81,* 42b.5.1–42b.5.7, Montreal, 21–25 September, 1981.

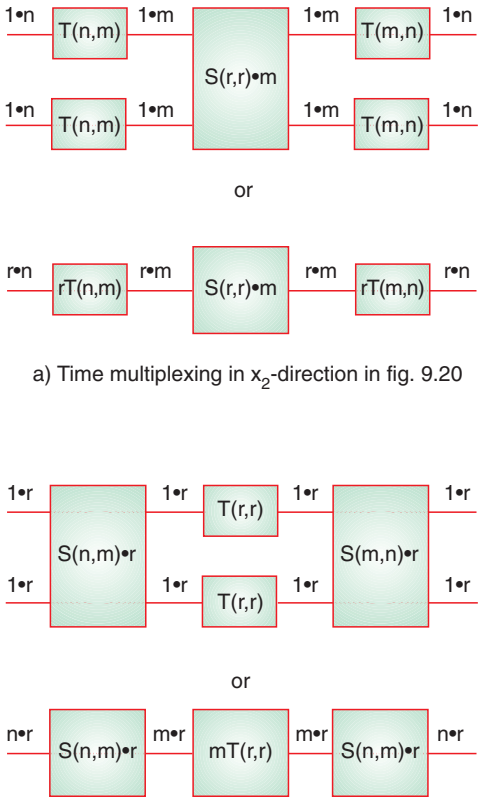


Figure 9.21 Space-time transformation of a 3-stage network without link rotation

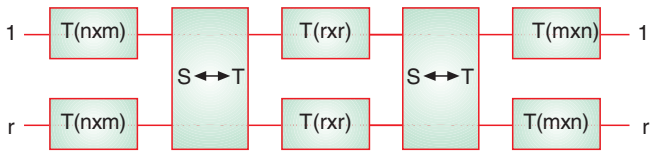


Figure 9.22 Space/time transform of a 3-stage (rearrangeable) Clos network

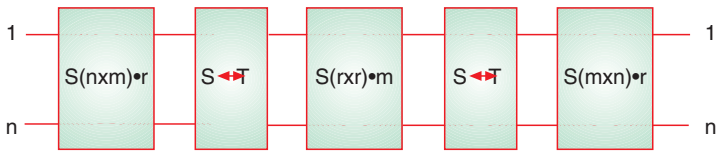


Figure 9.23 STS network transformed to an S-network

7 Duerdoth, W T, Cheeseman, D S. Techniques for an integrated digital network for telephony. *International Switching Symposium,* Boston, 1974.

8 Edström, N H. Synchronization in IST systems. *Telecommunications Symposium,* June 1970.

9 Földes, A. Synchronization of PCM networks. *Period. Polytech. Electr. Eng.,* 19(1), 57–70, 1975.

10 Geissler, H. Planning of a PCM system hierarchy. *NTZ-Report 8,* Berlin, VDE-Verlag, 1971.

11 Gerke, P et al. The time division multiplex switching of digital speech channels. *IEEE Transactions on Communications,* vol. 22, 1721–1275.

12 Granello, G. Switching networks for PCM time-division exchanges. *IEEE Proceedings ISS 1972,* 81, 1972. (72 CH 0 617-1-COM.)

13 Harrison, S R, Andresen, S. TDM strictly nonblocking connecting networks employing fixed delay elements. *Proceedings ICC 72,* Ch. 40-13, 1972.

14 Hartley, G C et al. *Techniques of pulse code modulation in communication networks.* Cambridge University Press, 1967.

15 Huang, N, Lea, C. Architecture of a time multiplexed switch. *Globecom '86,* Houston, Nov. 1986.

16 Karnaugh, M. A model for the organic synchronization of communication systems. *Bell System Technical Journal,* 1705–1735, Dec. 1966.

17 Marcus, M J. Time space equivalents in connecting networks. *Proceedings ICC 70,* vol. 2, ch. 35–25, 1970.

18 McDonald, J C (ed.). *Fundamentals of digital switching.* New York, Plenum Press, 1983.

19 Pan, J W. Synchronizing and multiplexing in a digital communications network. *Proceedings IEEE,* 60(5), 594–601, 1972.

20 Parks, P C, Miller, M R. Results of some studies of PCM network synchronization systems. *Automatica,* 5, 449–463, 1969.

21 Rekiere, B J et al. An introduction to PCM switching. *Automatic Electric Journal,* April 1971.

22 Schindler, H R. Delta modulation. *IEEE Spectrum,* Oct. 1970.

23 Schmidt, R V, Alferness, R C. Directional coupler switches, modulators, and filters using alternating db techniques. *IEEE Transactions on Circuits and Systems,* CAS-26(12), 1099–1108, 1987.

24 Shimazu, Y, Takada, T. High speed time switch using GaAs LSI technology. *IEEE Journal on Selected Areas in Communications,* SAC-4(1), 32–38, 1986.

25 Slabon, R. Introduction of realization of the PCM switching network in the EWS 1 electronic switching system. *NTZ 28,* H.3, 102–107, 1975.

26 Streitlien, B. The switching network in III 1240. *Telektronikk,* 80(2/3), 1984.

27 Vaughan, H E. An introduction to No. 4 ESS. *International Switching Symposium,* Boston, 1972.

10 Broadband switching

10.1 Introduction

The term broadband switching has been used very often over the last years due to the introduction of new services which place greater demand on bandwidth than common services like speech, low speed data, etc. It is then a question of definition: What is broadband switching? In this section broadband switching will be defined as switching of signals either with bandwidth greater than 2 MHz or with bitrate greater than 2 Mb/s. Up till now the upper limit has been 140 – 600 Mb/s, but still higher bitrates can be expected in the future. Switching of analogue signals with bandwidth corresponding to these bitrates must also be expected in the future.

Table 10.1 gives an impression of the variation of bitrates for some broadband services. In this table the measure is bitrate, and approximate equivalent analogue bandwidth can be found by using the sampling theorem of Shannon. The table points out the difficulty of giving an exact definition of broadband switching.

Data with lower bitrates are called low band services and, according to the table, this means services demanding bitrates less or equal to 64 kb/s.

The basic difference between switching of low and broadband services following this is nothing more than switching of data with different bandwidths/bitrates.In principle, it is only necessary to adjust the switch such that it is able to handle greater bandwidths/bitrates. By evaluation of broadband services both the transmission and switching parts of the net-

Table 10.1 Some broadband services and associated bitrate

Service	Bitrate bit/sec	Time per bit
Videophone (coded)	64 k – 384 k	15.625 µs – 2.604 µs
Data communication	64 k – 140 M	15.625 µs – 0.0071 µs
Video library	2 M – 140 M	500 ns – 7.1 ns
TV (common)	34 M – 140 M	29 ns – 7.1 ns
HDTV	140 M →	7.1 ns →
Text-TV	2 M – 140 M	500 ns – 7.1 ns
Pay-TV	34 M – 140 M	29 ns – 7.1 ns
ATM	600 M –	1.67 ns

work must be taken into consideration. With the introduction of the optical fibre, the transmission part is, however, able to offer services with a very high bitrate. This is not the case with today's switching technology.

10.2 Technology for broadband switching

Table 10.1 indicates that:

- Some of the services need technology able to give operation times of nanoseconds
- The existing technology can determine which principle the switch may be based on.

In space switching a connection is held during the whole transfer of the data which means that there, in principle, is no demand on high operation times of the switch. This is true when using relation elements in the switch. If, however, the element is of the logical type, it has to "follow" the input (data) signal and this determines the necessary operation times of the switching element. This is illustrated in Figure 10.1. Relation elements work equally well with both analogue and digital representation of the data and operation times are not in the picture. This is in contrast to using a logical gate as the switching element. Apart from the fact that the gate strongly distorts an analogue signal, the gate must also have operation times such that a bit can be identified at the output of the gate.

Another important figure is the time of transfer data from a user. As an example, the transfer of a file of size 15 Mbyte corresponding to the data on about 10 (3.5") diskettes gives:

64 kb/s: $\approx$ 234 sec
600 Mb/s: $\approx 234 \cdot 10^{-3} = 234$ ms.

It is obvious that this sets restrictions with respect to usable technology. When using a relation element, the following parameters are important:

- Insertion loss
- Crosstalk
- Linearity.

All these parameters can vary with frequency and the significance of this fact has to be evaluated for each technology and also for each switching principle.

As an example, the logical gate in Figure 10.1 regenerates the input signal to a standard output signal as long as the input signal meets certain limits. This means that a stronger signal degeneration can be tolerated than for a corresponding analogue signal (with relation element).

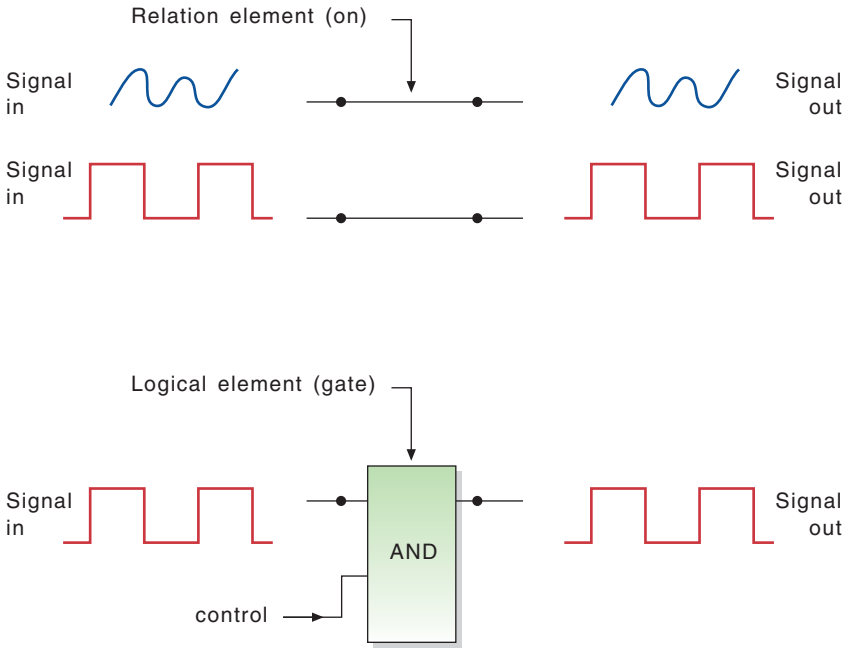


Figure 10.1 Relation between switching elements and input signals

Actual elements for broadband switching will, as for switching of data with low bitrate (small bandwidth), be:

- Metallic elements
- Semiconductor elements
- Optical elements.

The first group of elements can be represented by a relay contact having an equivalent as shown in Figure 10.2. For such an element it is reasonable to depict parameters which are linear up to several GHz but with a crosstalk making this element unusable for frequencies above about 5 MHz (see Section 4).

Another advantage of these metallic elements is that they are relation elements usable for both digital and analogue signals. Because of a small insertion loss many elements can be cascaded without giving too high total insertion loss.

Provided the necessary effort is met with respect to connection of elements, transmission between switches, and with the restriction mentioned above, it seems that metallic elements are well suited as space switches to switch broadband signals. However, in future switching networks, it is assumed that metallic elements pose no realistic alternative.

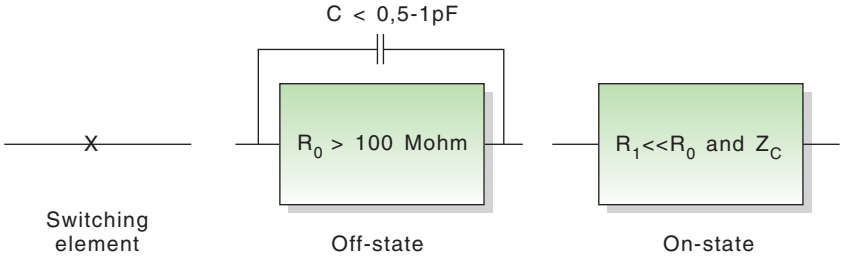


Figure 10.2 Actual parameters for evaluating a switching element

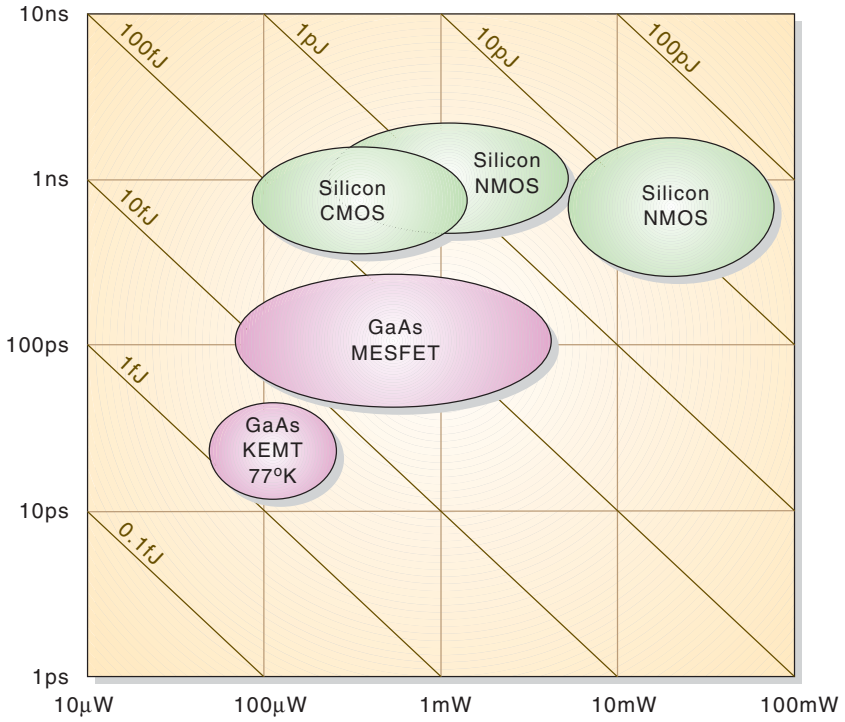


Figure 10.3 Power and switching time for some semiconductor technologies

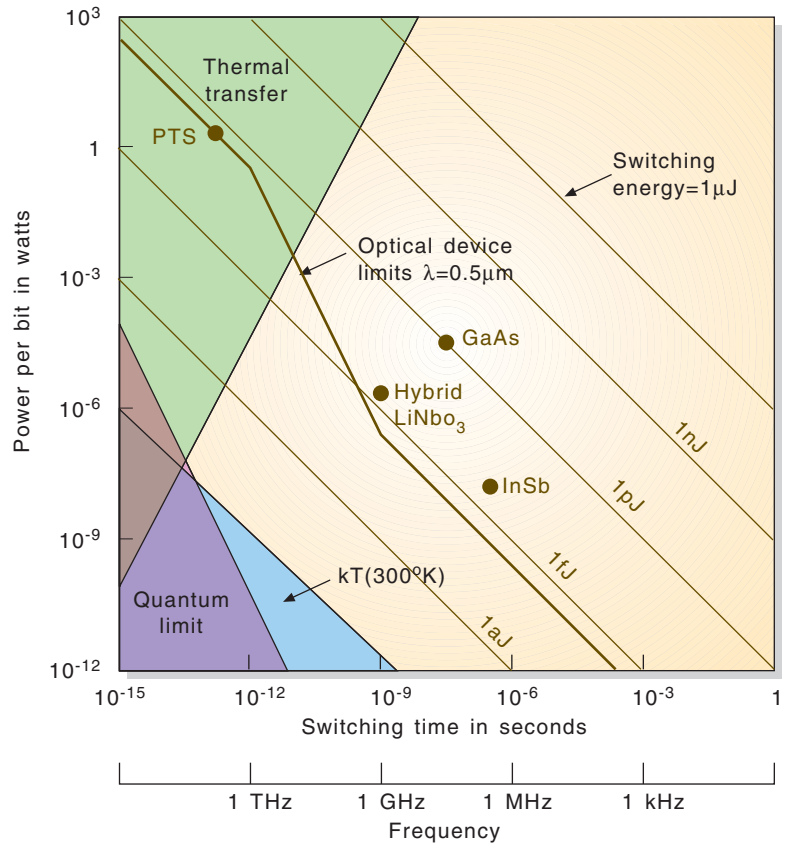


Figure 10.4 Regions of performance for semiconductor and optical components

The semiconductor technology (giving the second element group) now gives elements with satisfactory parameters for space switching. One drawback is that the insertion loss does not allow cascading of many elements without amplification. The power consumption when using logical elements can be high when data are switched with high bitrate (see Section 4).

The optical technology is of special interest because of the optical fibre introduced in the transmission network. Optical switching elements enables us to avoid the electronic bottleneck with respect to speed and also to avoid conversions between optical and electric power. The optical elements can be characterised by:

- Very high bandwidth
- Short switching times
- Low power consumption when used in space switching
- Insertion loss making it unfavourable to connect many elements in serial
- Physical dimension somewhat greater than the electronic elements
- Crosstalk not satisfactory
- Low linearity.

Implemented optical switches have crosstalk < 40 dB and loss < 0.3 – 0.5 dB. Some examples of optical switches are given in Section 5. It is assumed that optical technology will be dominating, along with the semiconductor technology, in the future.

Figure 10.3 shows the relation between switching time and power consumption for some semiconductor technologies. As stated in Section 5, the power consumption sets a practical limit on attainable switching times. Another important parameter is the degree of integration which, in practical terms, is restricted by the power consumption.

Optical technology is very promising in conjunction with time switching. Figure 10.4 shows regions of performance for optical components.

In that figure some semiconductor technologies are also marked. The lower limit of power consumption is given by quantum theoretical consideration. This limit is given by:

$$E = h / \tau$$

where

h = Planck's constant
τ = switching time.

The figure shows that optical technology seems to be an attractive technology for "switching of broadband services".

Figure 10.5 shows a comparison between the performance of three technologies:

- Semiconductor technology
- Supra technology (Josephsons-elements)
- Optical technology.

For the sake of comparison, the figure also shows the performance of a neuron (a biological switch). Components in supra technology are expected not to be relevant until materials with higher supra temperature are found.

It is obvious that in a network with only broadband users there will soon be very strict demands on the technology. Access and switching times down to the ps-region (10^{-12} sec) are necessary when implementing switches with, say, 10,000 users each with a bitrate in excess of 2 Mb/s.

Table 10.2 shows some data for commercial technologies. However, these performances are expected to be much improved within a few years.

In Table 10.3 is shown data for some available data stores which can be used for broadband switching.

10.3 Broadband space switching

As stated previously time is not a parameter in a space switching network so time to establish the connection is not an important parameter. The example above shows, however, that this can be somewhat misleading since a network switching data with very high bitrate also has to establish the connection in a short time to avoid an unfavourable ratio between time of data transfer and time for establishment of the connection. This parameter has an impact on what technology to be used rather than the switching network itself. So in general, a space switching network can be used also for switching broadband signals. It is also important that there is no need to synchronise the incoming bit stream before switching.

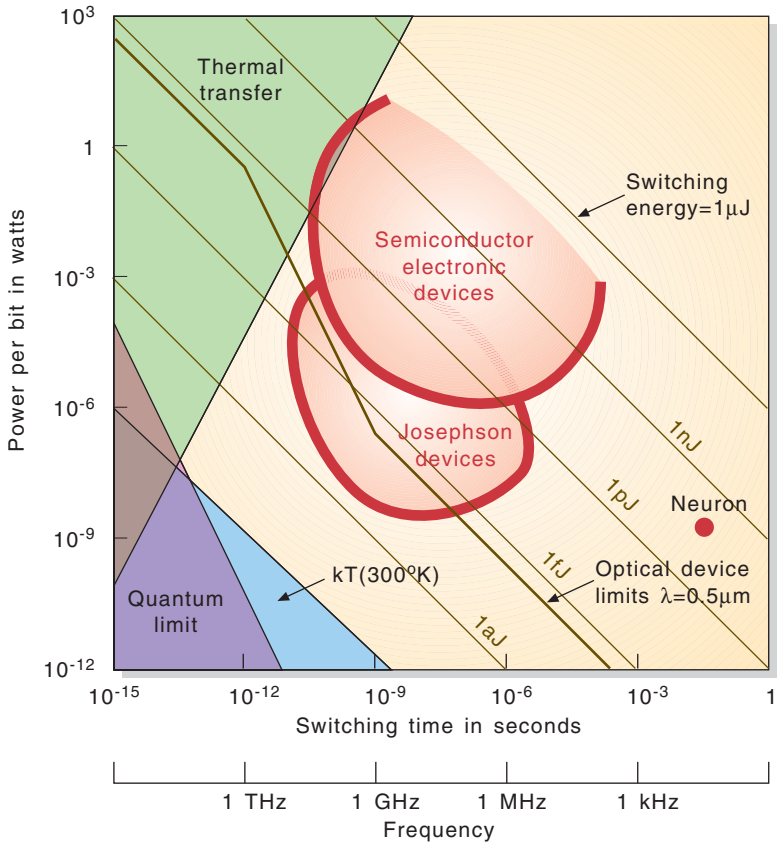


Figure 10.5 Regions of performance for different switching technologies

Table 10.2 Data for some commercial gates

Technology	Delay ns	Loss/gate mW	Number of gates/circuits
CMOS	2.7	1.2	50,000
Bi-CMOS	0.8	1.8	3,000
Bipolar (not ECL)	1.5 – 2.5	0.3 – 1	10,000
ECL	0.15	1	15,000

Table 10.3 Data for some commercial data stores

Technology	Access time ns (1)	Loss/RAM W	RAM-module
F100K ECL	12.8	1.2	256 x 4
F100K ECL	8.8	0.9	16 x 4
Harris BFL	1.85	1.5	64 x 1
GBL SDFL	3.0	1.8	256 x 4

(1): Included delay through address register and in-out registers

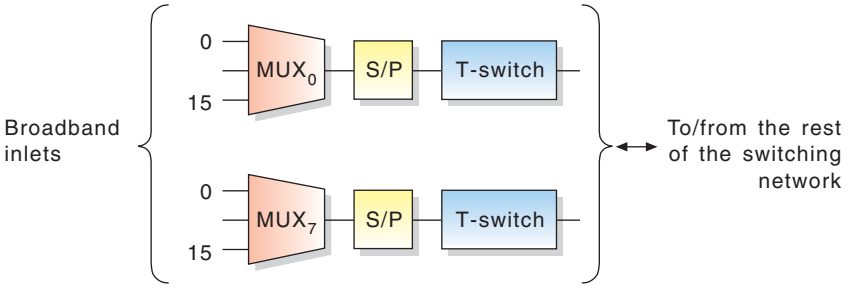


Figure 10.6 Input stage of a time switching network

10.4 Broadband time switching

In contrast to space switching, time is an important parameter in time switching. In practically all (larger) switching networks, data are represented in a digital way. It is common to use bitrate measured in bit per second as performance with respect to the bandwidth of the switch. Equivalent analogue bandwidth can be calculated by using number of bits per transfer and time between each transfer (between the same user). Such a calculation can be useful when comparing a space and a time switching network.

Example 1:

Number of bits per sample d = 8 bit, and sampling frequency f_s = 12 MHz, gives user bitrates:
$b = 12 \cdot 10^6 \cdot 8 = 96$ Mb/s.

This is the bitrate necessary in order to transfer a "common" analogue TV signal in a digital switching network.

With given bitrate B, the number of channels (users), each with bitrate b, is found from:

$$N = \lfloor B/b \rfloor \qquad (10.1)$$

where

$\lfloor X \rfloor$ = the greatest integer less or equal to X.

Example:

A bitrate of 1 Gb/s (10^9 bit/s) gives

- for b = 96 Mb/s:

  $N = \lfloor 10^9/96 \cdot 10^6 \rfloor = 10$ channels

- for b = 64 kb/s:

  $N = \lfloor 1 \cdot 10^9/64 \cdot 10^3 \rfloor = 15624$ channels

These examples illustrate how the bitrate influences the number of channels (or users) in time switching. In a public broadband network one can expect 10,000 – 20,000 "broadband" users in an exchange. This implies that with today's technology one has to use the earlier mentioned combination of time and space switching. Increasing bitrate also means that the operation times of the elements must decrease. Bitrate B = 1 Gb/s gives switching times down to 0.5 ns with the same rise and fall times. The power consumption will also increase with increasing bitrate as stated above.

10.4.1 Synchronous time switching networks

Synchronism in time switching networks means that there is:

- Synchronism with respect to bitrate
- Synchronism with respect to channels
- Synchronism with respect to frames.

In addition, the number of bits per channel and frame is fixed. The problems with broadband time switching (BTSW) are the same as for narrowband time switching (NTSW) at higher level in the network where many channels are multiplexed into a common inlet to the switch. This seems to be reasonable, since one difference between a broadband and a narrowband system is a higher bitrate in the first. The main problems will be operation times of switches, logic and data stores.

In principle, the difference between BTSW and NTSW can be seen as a reduction in number of inlets (and outlets) in a narrowband switch such that existing technology can be used. Figure 10.6 shows the input stage of a switching network with 8 T-switches. If each MUX has 16 inlets, each with a bitrate of 75 Mb/s, this gives the following operation times when assuming parallel switching of 8 bits and frame reading (2 stores each group, see Section 3):

MUX and S/P:
$(16 \cdot 75 \cdot 106\ \text{Hz})^{-1} = 0.83$ ns

Access time of data stores for T-switches: 0.83 ns · 8 = 6.6 ns.

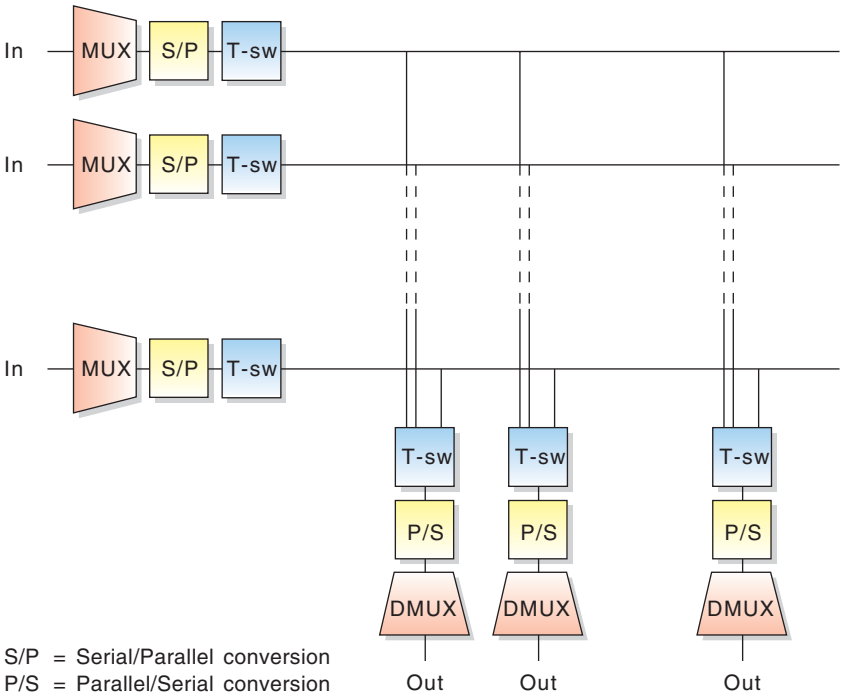


Figure 10.7 T-switches and multibus system

These operation times are within limits of today's technology. To increase the number of inlets, bitrate and/or to decrease the demand on faster components, several methods can be used:

- Switching more bytes in parallel giving a more expensive switch but reducing the technological stress
- Increase number of modules to increase number of inlets at the cost of more expensive technology
- Use a combination of time and space switching to give more inlets without demand for shorter operation times
- Combine T-switches and bus systems.

A combination of T-switches and buses is illustrated in Figure 10.7 and a combination of T-switches and an S-(space-) switch is illustrated in Figure 10.8.

When using T-switches and a bus the time delay on the bus must be taken into account. With high bitrates a method to avoid such delay problems is illustrated in Figure 10.9 where the bus is divided into equal parts each "passing" a common point. The maximum acceptable delay is given by the delay from a T-switch to the common point.

10.4.2 Combination of narrow- and broadband users

A reference in later years with respect to bitrate is the digital coded voice, i.e. a bitrate of 64 kb/s. To increase the bitrate of a user in such a network, more 64 kb/s channels can be combined with the effect that number of users is reduced and that the available bitrate of the system is restricted. In Figure 10.10 this principle is illustrated. A user (terminal) is connected directly to a T-switch which has a capacity of, say, 30 channels each with bitrate 64 kb/s. The user can now be assigned bitrates up to 2 Mb/s in steps of 64 kb/s. The T-switch operates as a "normal" narrowband T-switch. (The receiving user has to be connected in the same way to make advantage of the higher bitrate from the transmitter.)

The number of single channels assigned to a user in this way will vary between 1 and a number *k* depending on:

- The total capacity of the switch
- The number of terminals which simultaneously is allowed to increase the bitrate
- The total number of narrowband channels available.

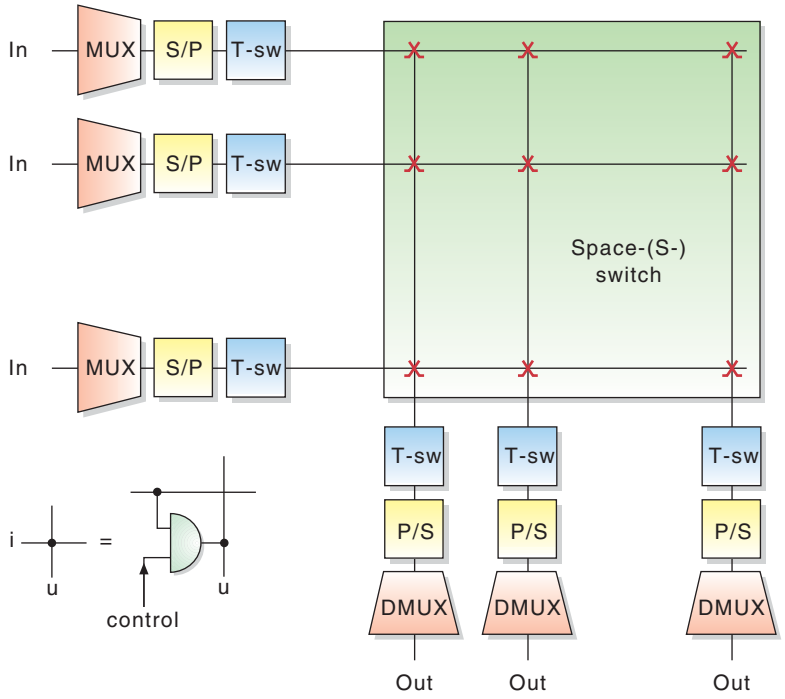


Figure 10.8 T-switches and an S-switch

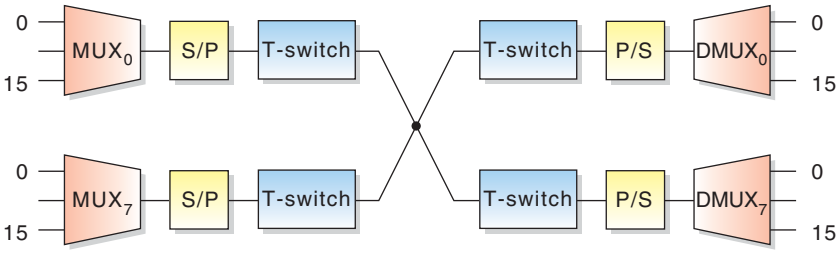


Figure 10.9 T-switches and a modified bus

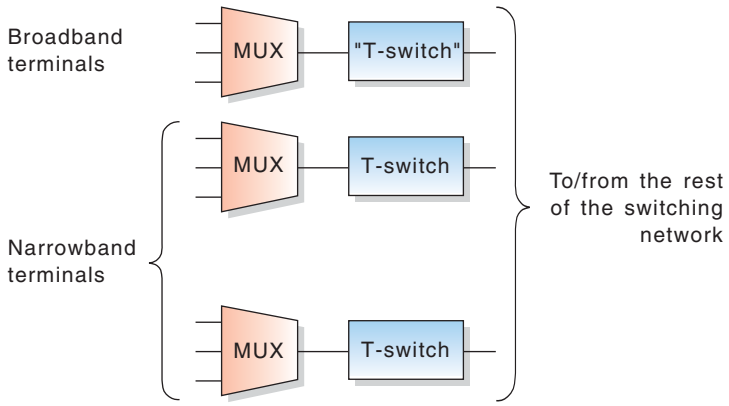


Figure 10.10 Possible combination of narrowband and broadband users

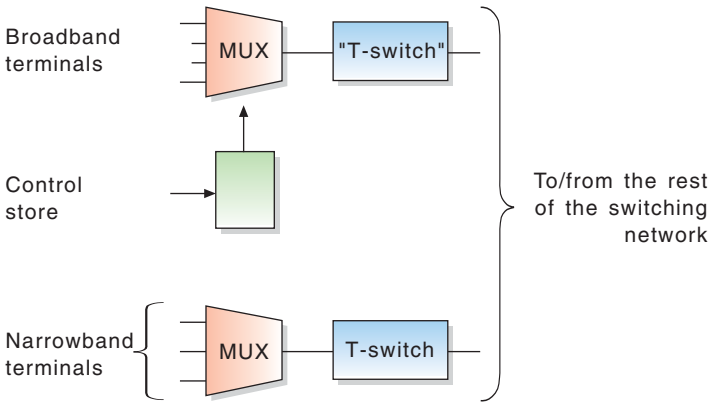


Figure 10.11 Dynamic allocation of bitrate

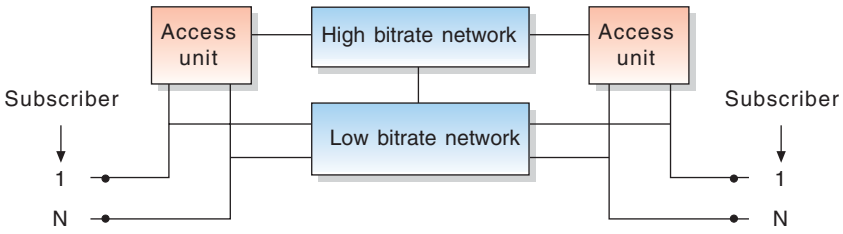


Figure 10.12 Separate broadband network

The control of this "collection of channels" is the same as of a single channel. The same technology can be used for both types of users.

The reduction ratio of number of users can be expressed as:

$$r = [N - \{(k_1 - 1) + (k_2 - 1) + ... + (k_n - 1)\}]/N$$

where

k_i = number of narrowband channels needed by broadband user i

n = number of broadband users

N = total number of narrowband users.

Example:

$N = 16 \cdot 30$, $k_1 = k_2 = ... = k_n = k = 30$ and $n = 3$ gives: $r = 0.818$

$N = 16 \cdot 30$, $k_1 = k_2 = ... = k_n = k = 30$ and $n = 8$ gives: $r = 0.516$.

This principle is perhaps the easiest way of solving the problem in an existing system based on time switching. It is, however, seen that even a relatively low broadband demand, like 2 Mb/s, drastically reduces the number of inlets and outlets compared to a "narrowband switch".

Figure 10.11 illustrates a more dynamic assignment of bitrates. The input multiplexer is made dynamic by addressing it from a control store containing the addresses of those inlets which have to be connected to the T-switch. In this way a switch can be used more efficiently by using idle channels (idle time slots) between busy channels.

The highest bitrate available for a broadband user connected to a single T-switch is set by the number of time slots in the T-switch. It is assumed that this number is restricted to 512 time slots corresponding to a total bitrate of 32 Mb/s.

By merging narrowband and broadband users on more T-switches, a more flexible use of existing channels can be achieved.

Another alternative to introduce broadband services is to establish a separate broadband network as illustrated in Figure 10.12. A user can be given access to the broadband network via an access unit.

This could be an easy way to introduce new technology.

Such a synchronous network can be characterised by:

- Fixed sequence of the channels (time slots)
- No extra equipment to identify number of time slots, etc.
- Minimum total delay through the switching network
- Extra equipment to handle loss of data during loss of synchronism
- Bitrate available in steps of 64 kb/s
- Data arrive in right sequence.

10.5 Time switching of asynchronous data

Asynchronous data will here be defined as data not transmitted in fixed time slots. It could be expected that this will influence the switching process. Depending on the actual situation, this is correct or incorrect. This can be explained as illustrated in Figure 10.13a which shows a time switching network consisting of:

- N terminals in and N terminals out, each with associated terminal equipment (LU)
- A multiplexer, MUX, and a demultiplexer, DMUX
- An address register, ADR
- A control store, CS, controlling DMUX
- A pulse generator, PG.

Assume that data on each of the inlets are making a bit stream with no correlation on data on other inlets with respect to phase or time. Every bit, however, is assumed to have the same level, duration t_b and guard interval $t_g = t_b$. This is illustrated in Figure 10.13b. Assume that the bitrate of the pulse generator PG is given by:

$$b_G = 1 / \tau_0$$

where

τ_0 = time between pulses from PG.

The function is as follows:

- An input is addressed during a time interval τ_0
- The result (a piece of duration τ_0 of the input bit) is transferred through the

MUX and DMUX under control of the address from CS

- At the output of DMUX the result can be transformed to normal bit level and duration by the terminal equipment.

The cycle time for MUX (i.e. time between the addressing of the same terminal) is:

$$\tau_b = \tau_0 \cdot N$$

An expression for maximal value of N can be given as:

$$N_{MAX} = \lfloor \tau_b / 2k\tau_0 \rfloor;\ k = 2, 3, \ldots$$

where

k = number of times a terminal is addressed during a bit.

Example:

τ_b = 488 ns (first order PCM),
τ_0 = 100 ps, k = 2 gives N_{MAX} = 1220

τ_b = 1667 ps (600 Mb/s),
τ_0 = 10 ps, k = 2 gives N_{MAX} = 41.

This simple example shows that time switching can be used to switch asynchronous digital data. The principle can, of course, also be used to switch analogue data.

The concept asynchronous data communication refers to the transfer of data packets each with a construction as illustrated in Figure 10.14:

- A header consisting of a start label and the destination address
- Data with variable length (number of bits or bytes)
- A tail consisting of stop indicator, control bits, etc.
- The time between packets arriving of a receiver can vary.

The switch executing the switching is often called a "packet switch". It must be stressed, however, that the switching itself is as for all types of data.

For services like voice and video, the data set must arrive in right sequence. The distance (measured in time) can vary within certain limits but the sequence has to be correct in order to avoid distortion. To a certain limit this can be obtained by:

- Same path is used during the whole connection time
- Each packet is assigned a number giving the place in the sequence.

The main advantage and disadvantage of "packet switching" can be stated as:

- Flexibility with respect to assigned bitrate to a user
- The effective transfer rate is decreased.

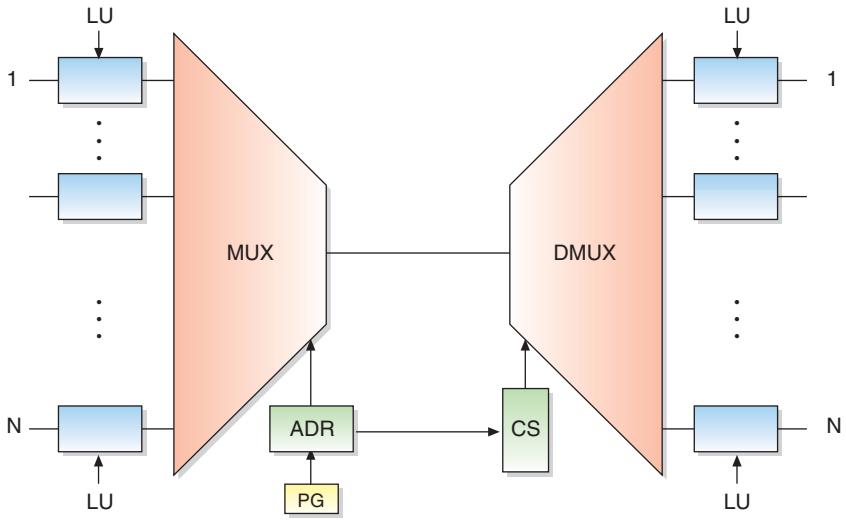


Figure 10.13 Time switching of asynchronous digital data

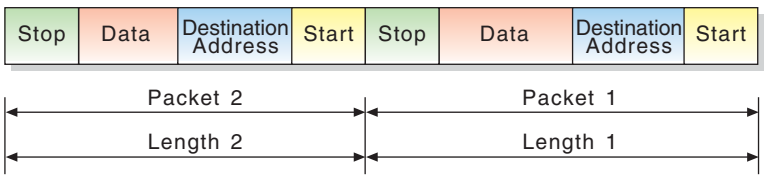
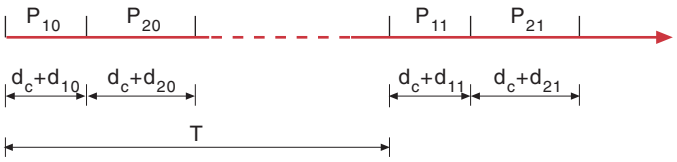


Figure 10.14 Illustration of input data in "packet switching"



P_{ik} = packet No. k from user No. i
d_{ik} = number of bits from user i in packet k
d_{ci} = number of control bits in a packet
T = time interval between two packets from the same user

Effective bitrate for user 1, $b_1 = d_{10} / T$

Figure 10.15 Illustration of effective bitrate

Total length: 424 bit						
Data field: 384 bit	Header 40 bit					
	24 bit	8 bit	2 bit	4 bit	1 bit	1 bit
User data Connection data No data	Control data: VCI+VPI	Error detection HEC	Cell type PT	User control GFC	Spare RES	Cell loss CLP

VCI = Virtual Circuit Identifier
VPI = Virtual Path Identifier
GFC = Generic Flow Control
HEC = Header Error Control
PT = Payload Type
CLP = Cell Loss Priority

Figure 10.16 Data contained in an ATM cell

The effective bitrate of a transfer is determined by the time between transfers to the same user and the number of bits in each packet. This is illustrated in Figure 10.15.

Number of packets transferred in time interval T can be calculated from:

$$\tau_0 \sum_{j=1}^{j=n_A} d_{Aj} = T$$

for asynchronous transfer, and (10.2)

$$\tau_0 \sum_{j=1}^{j=n_c} d_{0j} = T$$

for synchronous transfer (10.3)

where

τ_0 = bit duration
$d_{Aj} = d_{Cj} + d_j$ = packet size by asynchronous transfer
d_{0j} = packet size by synchronous transfer
d_{Cj} = size of control part
d_j = size of user data part.

If $d_{C0} = d_{C1} = ... = d_{Cn} = d_C$

$d_0 = d_1 = ... = d_n = d_0$

then: $n_A = T / (\tau_0(d_C + d_0))$

$n_C = T / (\tau_0 d_0)$

and $n_C / n_A = 1 + d_C / d_0$ (10.4)

The number of bits in d_C is determined by:

- Number of start/stop bits
- Number of parity/control bits
- Number of address bits.

To avoid imitation of control bit codes it is favourable to use a fixed length of the packets.

Normally, 2 byte (16 bit) is necessary to indicate start/stop and parity code.

The size of the address field depends on how the path is established. If the path through the network is established before the packets to a receiver are sent, it is not necessary to include the whole destination address in every packet. This is the case in ATM where only a stepwise path address is given (see later). The other main alternative is to send the whole address together with the packet. With 8 digits (32 bit) in the address, 16 bit for other control information and $d_0 = 8$ bit data, this gives from eq. (10.4):

$$n_C / n_A = 1 + 48 / 8 = 7$$

This gives an impression of how much an asynchronous transfer can reduce effective bitrate to the user relative to a synchronous transfer.

To give a higher bitrate it is common to "collect" more packets with the same destination address. If 32 samples (of 8 bits) are "collected" and sent in the same packet, this gives with the same parameters as above:

$$n_C / n_A = 1 + 48 / (32 \cdot 8) = 1.18$$

This is a great improvement but still asynchronous transfer is more unfavourable than synchronous transfer when measured in this way.

10.6 ATM

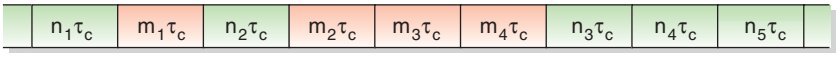
In the last few years the asynchronous transfer concept has been standardised under the name ATM (Asynchronous Transfer Mode). The main properties of ATM are:

- A packet in ATM is called a cell consisting of a header data part
- All cells are of fixed length with 53 octets (bytes) of 8 bits each, i.e. 424 bits distributed as follows (see Figure 10.16):
  - 48 octets bit information (user data)
  - 5 octets header containing:
    - 3 octets to specify:
      - Virtual Circuit Identifier
      - Virtual Path Identifier
    - 1 octet (HEC) for detection of right start of the cell and for failure correction
    - 4 bits (GFC) for user purposes

- 2 bits (PT) to indicate type of cell:
  - cell with path data
  - cell with user data
  - empty cell
- 1 bit (CLP) for indicating priority
- 1 bit spare (RES).

ATM is characterised by:

- All cells have the same length (same number of bits)
- Cells are sent continuously, i.e. there is no empty space between the cells
- Empty cells are sent if there is no information to be transferred
- All cells belonging to the same connection have to follow the same path established prior to the transfer
- The path through the network is established by using data in cells marked as "connection cells"
- Data about a selected path are stored in a table in every switching node and is looked up during the switching process
- Releasing a connection is executed by cells marked as "release cells"
- VCI and VPI in the header of the cell identify the connection for the data cells following the "connection cells", i.e. after the path is established
- A new header (new VCI, VPI and HEC) is added to the cell after passing the switching network
- Empty cells may be used to find the start of a cell, but instead of that the following procedure is introduced in ATM:
  - The header and HEC are received and the HEC value is calculated
  - If calculated and received HEC values are equal the right start of the cell can be assumed
  - If calculated and received HEC values are not equal, the cell is dropped
  - To find the start of a cell when calculated and received HEC are different for 5 – 8 cells, the following procedure is used:
    - The first bit in the received header is dropped and the next bit after the header is added to the rest of the received header
    - A new calculation of HEC is executed, and if the received and the calculated HEC are found to be equal, it is assumed that the right start of the cell is found
    - If the result of the comparison is that they are different, the process is repeated.



n_i = active cells from user *i*
m_j = empty cells or cells from other users
τ_c = duration of a cell

Figure 10.17 Assignment of "user bitrates"

The whole cell is found by counting the number of bits.

In ATM, a user is offered the bitrate that the service requests; in principle, this could mean a user selected bitrate in steps of 1 cell. If the user bitrate is denoted by b, the following expression can be found (see Figure 10.17):

$$b = (\Sigma n_i)n_c / ((\Sigma n_i + \Sigma m_j)\tau_c)$$

where

n_i = active cells from user i

m_j = empty cells or cells from other users

τ_c = duration of a cell

n_c = number of bits per cell.

With $\tau_c = n_c / B_0$, where B_0 is the maximal bitrate, this gives:

$$b = [\Sigma n_i / (\Sigma n_i + \Sigma m_j)]B_0$$

Example:

$m_j = 0$ for all j gives $b = B_0$
i.e. max user bitrate

$m_j \rightarrow \infty$ gives $b = 0$
i.e. no active cells are transferred from this user.

In ATM there are two statistical processes:

- The statistical variation of calls
- The statistical variation of bitrates.

ATM is intended to be used for transfer of all data types. Speech can be transmitted by collecting 48 (8 bit) words in each cell giving a total delay of 6 ms.

Besides the bitrate, ATM has other characteristic parameters like:

- Bandwidth for instance expressed by:
  - distribution function
  - max value R_{max}
  - mean value R_{mean},
    e.g. $R_{mean} = N_c / T$,
    where N_c = number of cells transferred in period T.
- Burstiness defined as:

$$\alpha_{burst} = R_{max} / R_{mean}$$

$$\alpha_{burst} = S / R_{mean}$$

$$\alpha_{burst} = S^2 / (R_{mean})^2$$

where S = standard cell rate.

10.6.1 Switching of ATM data

To switch ATM data the same principles as for synchronous data can be used. An "ATM switch" has to satisfy the same definitions as defined in the preceding sections. Switching networks for ATM can then be divided into space and time networks and their combination.

Important parameters which can influence on the actual implementation of the switch are:

- There are always cells at an inlet of the switch
- Each cell contains the address of the outlet of the network and a new address is added at the outlet.

Two succeeding cells can belong to two different users.

The control unit has to process the address contained in each cell to be able to establish the "path" through the network. To avoid loss of cells, the time available for this processing is given by:

Serial transfer: $T_s = \tau_d / n$

Parallel transfer: $T_p = \tau_C p / n$

where

τ_C = time per cell

τ_d = time between header and data

n = number of simultaneous cells

p = number of cells transferred in parallel.

In case of serial transfer the connection through the network has to be determined and established in the time interval between the last bit in the header and the first bit in the data. If B_0 = 600 Mb/s (max transfer bitrate) this gives with n = 1:

$$T_s \leq \tau_d \approx 1.7 \text{ ns}$$

where τ_d is the duration of one bit.

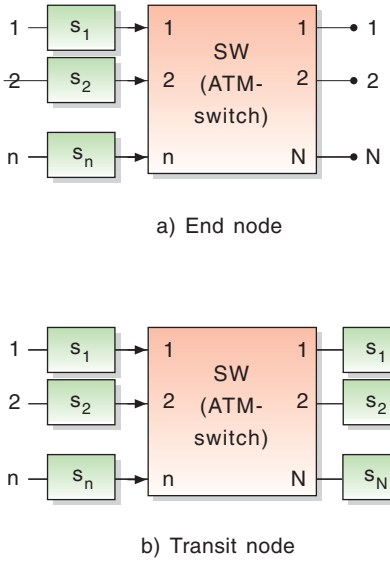


Figure 10.18 Different types of switching nodes

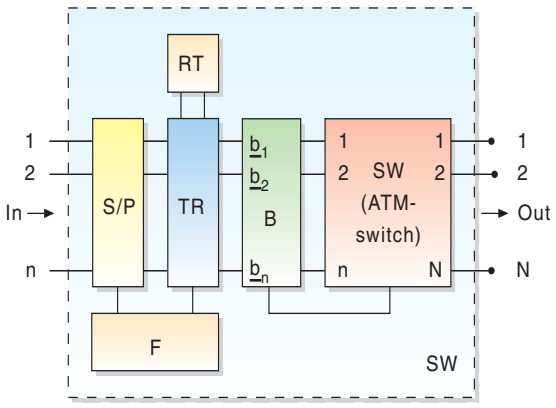


Figure 10.19 Functional units in an ATM node (with input buffer)

In parallel transfer, the time can be extended, but implying that more components are needed. If B_0 = 600 Mb/s and parallel transfer of all bits in the cell, the time available with p = 1 and n = 1 is given by:

$$T_p = \tau_d \cdot 424 \approx 700 \text{ ns}$$

The necessary processing time can be obtained by several means such as:

- By selecting switching networks with a short time to establish a call (transfer)
- By distributing the call processing to smaller groups of the network or even to the single switching element.

The demand for short processing times when switching data with high bitrates indicates a need for more parallel processing.

It may be suitable to distinguish between two circumstances as illustrated in Figure 10.18. In Figure 10.18a the data on the n inlets have to be distributed to N outlets connected to subscribers. SW is a subscriber node. In Figure 10.18b the outlets from SW go to other SWs. SW is then called a transit node.

In case no cells arrive simultaneously to the same outlet which means that there is no need for outlet buffering to avoid loss of cells.

In case it is expected that two succeeding cells belonging to different users have to be transferred through the same outlet of the switch and arrive simultaneously to the outlet. The common methods to solve this contention problem is to use:

- Input store
- Common store
- Output store.

Figure 10.19 shows possible functional units in an ATM. The main units are:

- Serial-parallel converting unit, S/P
- Unit F for detecting errors, testing and finding start of cells
- Unit RT with routing table and unit TR with translator table
- Unit B to store input cells (input buffer)
- Unit SW to execute the switching.

If SW is a non-blocking switching network, its need for storage is only at the inlets with a capacity of 1 cell per inlet.

If SW is not non-blocking, the storage capacity will increase since a cell has to wait for an idle connection. In this case it will be necessary to have storage capability inside the switch.

In Figure 10.20 we find examples of $n \neq N$. n cells arrive simultaneously at the n inlets and represent n different users, assuming that cells belonging to the same user follow the same route in the preceding network. A mean multiplex factor for the n inlets is given by:

$$\alpha = \lfloor N/n \rfloor$$

$n < N$ gives an expansion network
$n = N$ gives a quadratic network
$n > N$ gives a concentration network.

With a non-blocking network (SW), the total bitrate in could be equal to the total bitrate out giving:

$$B_0 \cdot n = \sum_{i=0}^{i=N} b_i$$

where

B_0 = max bitrate per inlet
b_i = bitrate per outlet (user bitrate).

With a constant user bitrate $b_i = b_0$, the number of outlets is given by:

$$N = B_0 \cdot n / b_0$$

Example:

B_0 = 600 Mb/s, n = 16 and b_0 = 64 kb/s gives: N = 150,000

B_0 = 600 Mb/s, n = 16 and b_0 = 2 Mb/s gives N = 4,800

B_0 = 600 Mb/s, n = 16 and b_0 = 140 Mb/s gives N = 64.

This simple example shows the large variation that arises when offering free bitrates to the users. In practical systems this will be a great problem, but it can be solved in different ways for example by using a higher mean bitrate in the system than necessary or by using more inlets than necessary.

A solution with higher bitrates than necessary is possible when using optical technology.

By using more than one wavelength the capacity of an optical fibre can be written as:

$$k_{fibre} = G \cdot n_w / B_0 \text{ (channels)}$$

where

G = bitrate per wavelength
B_0 = bitrate per channel
n_w = number of wavelengths.

The present state of the optical technology indicates the following capacities:

G = 8 Gb/s, n_w = 8 and B_0 = 600 Mb/s gives: $k_{fibre} \approx 104$ channels

G = 8 Gb/s, n_w = 8 and B_0 = 150 Mb/s gives: $k_{fibre} \approx 416$ channels

G = 8 Gb/s, n_w = 8 and B_0 = 64 Mb/s gives: $k_{fibre} \approx 1000$ channels

G = 8 Gb/s, n_w = 8 and B_0 = 34 Mb/s gives: $k_{fibre} \approx 1880$ channels.

In the future, the optical technology is expected to increase the capacity to values indicated below:

G = 4 Tb/s, n_w = 16 and B_0 = 600 Mb/s gives: $k_{fibre} \approx 10 \cdot 10^4$ channels

G = 4 Tb/s, n_w = 16 and B_0 = 150 Mb/s gives: $k_{fibre} \approx 42 \cdot 10^4$ channels

G = 4 Tb/s, n_w = 16 and B_0 = 64 Mb/s gives: $k_{fibre} \approx 100 \cdot 10^4$ channels

G = 4 Tb/s, n_w = 16 and B_0 = 34 Mb/s gives: $k_{fibre} \approx 188 \cdot 10^4$ channels

In particular, it is expected that n_w will be increased to values up to 100.

The architectures relevant for switching ATM data are as earlier mentioned:

- Bus structures
- Ringbus structures
- Multistage permutation network.

The different principles will be briefly discussed below.

10.6.2 Bus structures

Bus structures for ATM switches have minor differences compared with the structures for switching data with low bitrates. Figure 10.21 illustrates a bus structure adjusted to ATM. The principle is the broadcasting principle, where each receiver identifies the address placed on the TDM bus by the sending port. Each receiving port is addressed during one time interval and within this interval address and data have to be placed on the bus, i.e. each sending port has the same functional units as when data are handled with lower bitrates.

The bitrate of the bus is:
$B_{bus} = N \cdot B_0$; B_0 = bitrate per inlet.

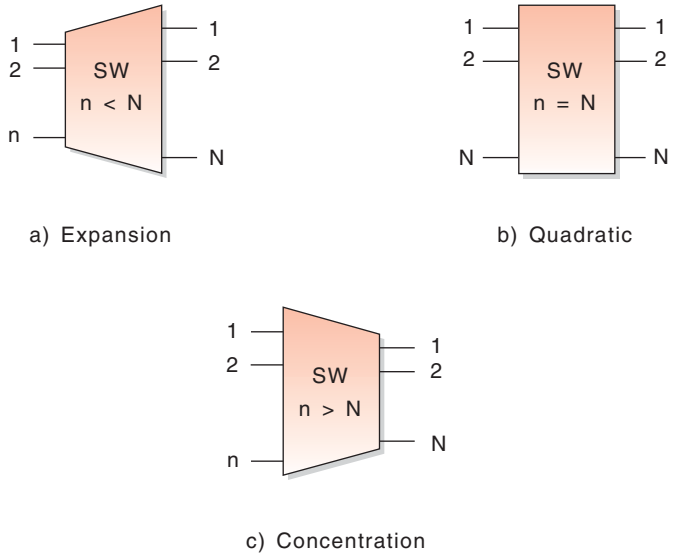


Figure 10.20 Topologies of SW

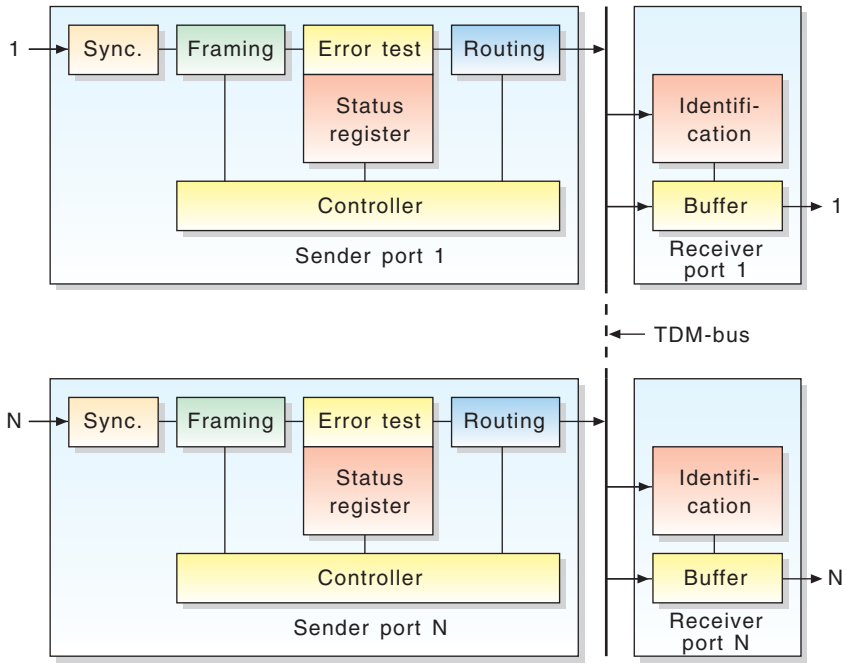


Figure 10.21 Bus system with output buffering for switching of ATM cells

Transferring bits in parallel reduces the demand on the bitrate of the bus given by:

$B_{bus} = (N \cdot B_0) / p_p$;
where p_p = number of bits in parallel.

Example:

B_0 = 600 Mb/s and N = 16 gives:

p_p = 8: B_{bus} = 1200 Mb/s

p_p = 16: B_{bus} = 600 Mb/s

p_p = 32: B_{bus} = 300 Mb/s

p_p = 128: B_{bus} = 75 Mb/s.

Note that problems with time delays increase when the bitrate increases (see Section 3).

The architecture in Figure 10.21 can be characterised by:

- No internal congestion
- Only outlet buffering
- No priority
- Time delay proportional to number of cells arriving simultaneously to the same outlet
- Increasing N changes the bitrate on the bus
- Possible different bitrates on the inlets.

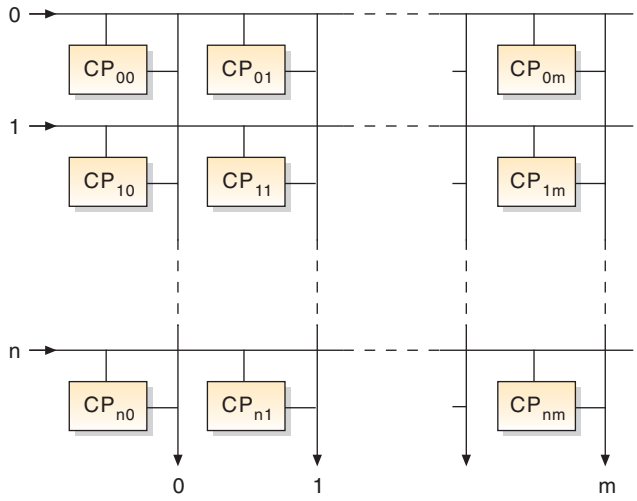


Figure 10.22 Matrix bus system

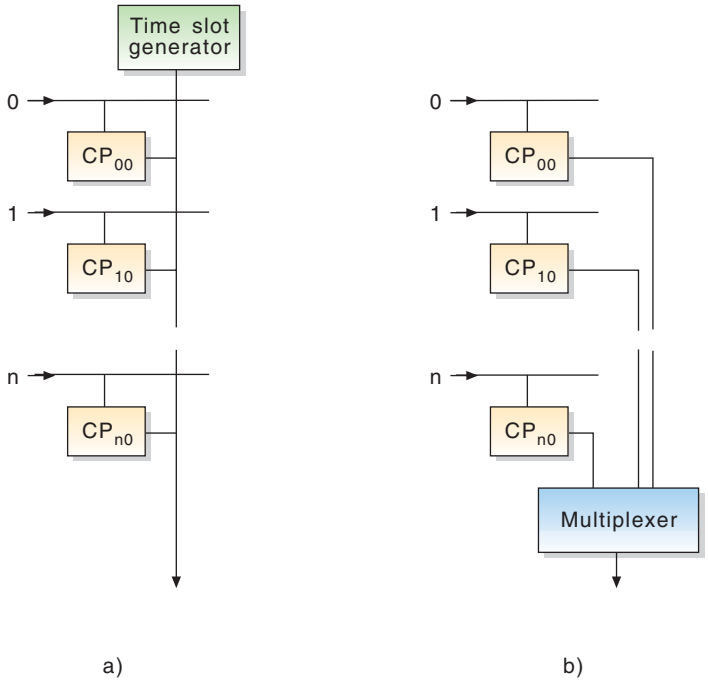


Figure 10.23 Different methods for accessing the output

The capacity of this structure is very dependent on the technology.

Figure 10.22 illustrates a matrix-bus-system (slotted buses) consisting of:

- $n + 1$ inlet buses
- $m + 1$ outlet buses
- An "intelligent" switch CP_{ij} in every crosspoint.

This structure corresponds to a matrix of size $(n + 1)$ x $(m + 1)$.

Data (cells) on an inlet are broadcast to all CPs on the same inlet (horizontal) bus. The header of the cell is examined and the information used when looking up in the tables in each of these CPs. The cell is stored in the CP that corresponds to the outlet given by the information in the header. The other CPs along the bus do not store the cell. A cell stored in a CP is switched out on the vertical bus in an idle time slot. A time slot generator (Figure 10.23a) produces time slots. Instead of a bus, a multiplexing system can be used as indicated in Figure 10.23b.

This principle can be characterised by:

- A relatively complex crosspoint with storing of cells (internal or distributed buffering), routing table, etc.
- No synchronism between the different inlet data
- Without multiplexer, the CP "nearest" to the time slot generator has the highest priority
- Flexible with regard to increasing number of inlets and outlets
- Varying time delay
- Time delay in each crosspoint given by:

$$\Delta = (n + 1) \cdot b_c / B_{bus}$$

where

b_c = storage capacity of the crosspoint CP

B_{bus} = bus bitrate.

10.6.3 Ringbus structures

A ringbus structure consists of a ring bus transferring data and to which the users can be connected. An actual structure is shown in Figure 10.24. It is known by the name of Orwell ring and was primarily intended for MANs (Metropolitan Area Networks).

The Orwell ring is a parallel bus with a fixed time delay divided into fixed time slots. The number of time slots depends on the total delay of the ring. Data can be sent in a time slot, indicating whether the time slot is idle or not. This indication is a packet marked TRIAL and has the address of the node sending the TRIAL packet as its destination address. The nodes (or exchanges) are connected to the ring. All data on the ring go through the nodes passed, and data always go in the same direction. The destination address of the packet also circulates together with the user data. A node can be in the following states:

- Active state, sending packets stored in its internal queue (buffer)
- Pause state, not allowed to send packets from the queue
- Idle state with no packets in the queue.

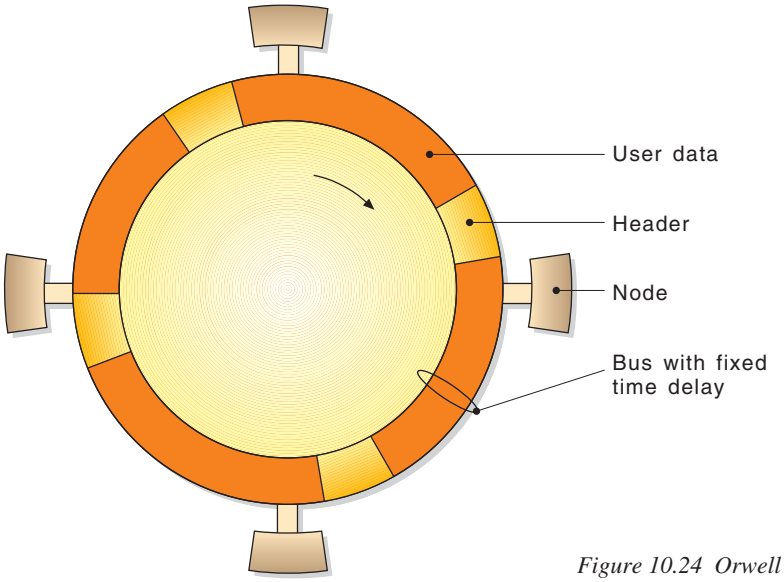


Figure 10.24 Orwell ring

The Orwell protocol is characterised by:

- A node identifies its address and removes data from the ring so that the succeeding part of the ring can be used by other nodes
- A node removing a packet from the ring also puts on the ring a TRIAL packet in the emptied time slot
- Each node has a counter D_i which is incremented by one when a packet is sent
- When the counter reaches a certain value, the node stops the sending of packets and goes to the pause state where no packets can be sent until the counter is reset.

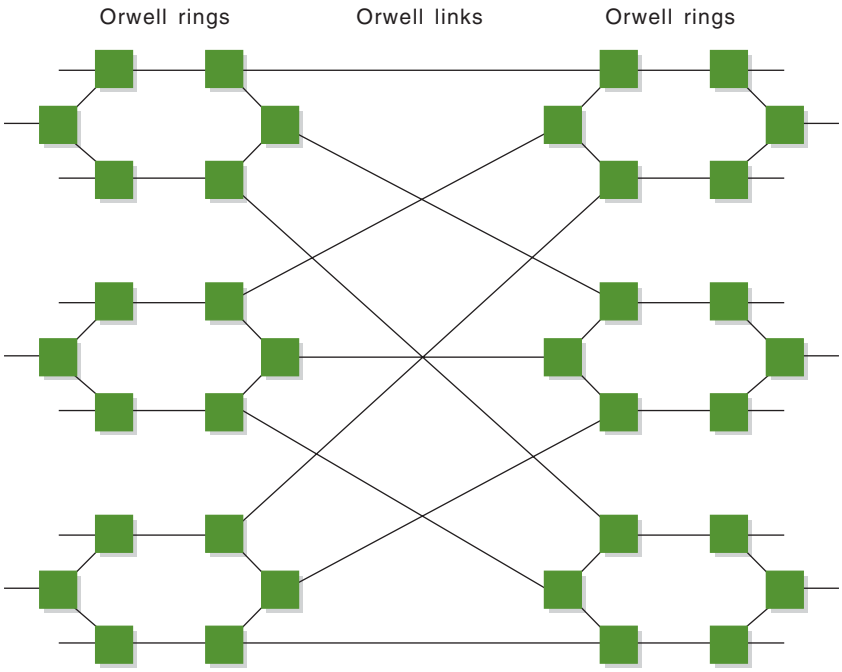


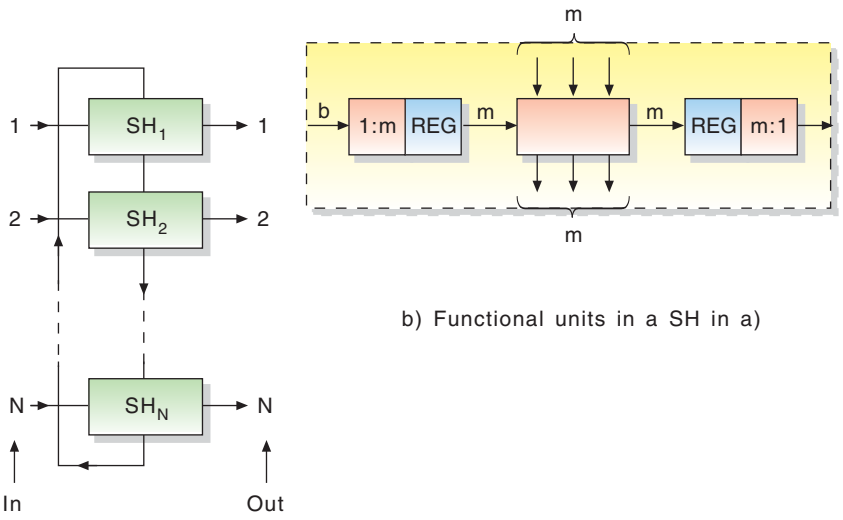
Figure 10.25 Switching network based on Orwell rings

The control of the reset of the counters is distributed and is executed in the following way:

- A node in pause or idle state marks succeeding unmarked time slots with the TRIAL mark together with its own address as destination address
- Time slots marked TRIAL can be taken by succeeding nodes which are not in pause state (the TRIAL mark is removed when a node uses the time slot)
- Each node going to pause state marks succeeding idle time slots (not having the TRIAL mark) with TRIAL mark
- When a node detects its unchanged TRIAL packet, this packet is changed to a RESET packet
- The RESET packet goes to all nodes and resets the counters of the nodes as it passes
- The content of a RESET packet can be changed only by the node sending the packet.

When all counters are reset, the sending process can start again and may continue until the queue is empty and then the node goes to idle state.

The RESET process ensures that all nodes get access to the ring within a pre-



a) Configuration

b) Functional units in a SH in a)

Figure 10.26 Shift register bus for switching of ATM data

set time. The protocol also ensures that a node by counting the number of idle time slots can reserve the necessary number of time slots to get the desired bitrate. This bitrate is obtained by setting the counter to a certain value.

The Orwell ring can be used to construct large networks as illustrated in Figure 10.25.

Other ring based structures can also be relevant for switching ATM data.

Figure 10.26a illustrates a configuration consisting of shift registers connected as a parallel ringbus. Each shift register consists of an input part, a shifting part and an output part. In the input part a serial/parallel conversion is executed.

The process is as follows (see Figure 10.26b):

- The input part executes a serial/parallel conversion and places the result in REG
- From REG the data is transferred to the associated shifting part if it is empty. This process is executed simultaneously for all REGs
- The shifting part shifts the data around the bus from one shifting part to the next
- When data reach the right shifting part, they are shifted out to the output REG which converts the data from parallel to serial form.

The purpose of REG is to ensure that data arrives correctly to the shifting part. The bus consists of an address part and a data part. The switching is executed by sending the destination address together with the data. Address and data circulate on the bus until they reach the right SH where the address is identified by the shifting part of SH. The data are then taken away from the bus and sent to the associated REG. To solve the contention problem, input or central buffer can be used.

Time between two packets is given by:

$$t_0 = n / b$$

where

b = bitrate on the input line
n = number of bits in the packet.

Time per packet shift is:

$$t_1 = t_0 / N,\ N = \text{number of inputs}$$

Example:

b = 600 Mb/s, N = 16, n= 424 bits gives:

$$t_0 = 424 / 600 \cdot 10^6 \approx 706 \text{ ns}$$

$$t_1 = (424 / 600 \cdot 10^6) / 16 \approx 44 \text{ ns}.$$

The bitrate on the bus is:

$$B = 1 / t_1 = N/(n / b) = bN / n$$

The same values as above give:

$$B = 600 \cdot 10^6 \cdot 16 / 424 \approx 23 \text{ Mb/s}$$

This is a moderate bitrate and the bitrate (on the bus) can be further reduced by extending the bus to double capacity and also extend the capacity of the shift registers.

10.6.4 Examples of switching elements for switching ATM data

An example of an "ATM" switch is described in a special contribution to this issue.

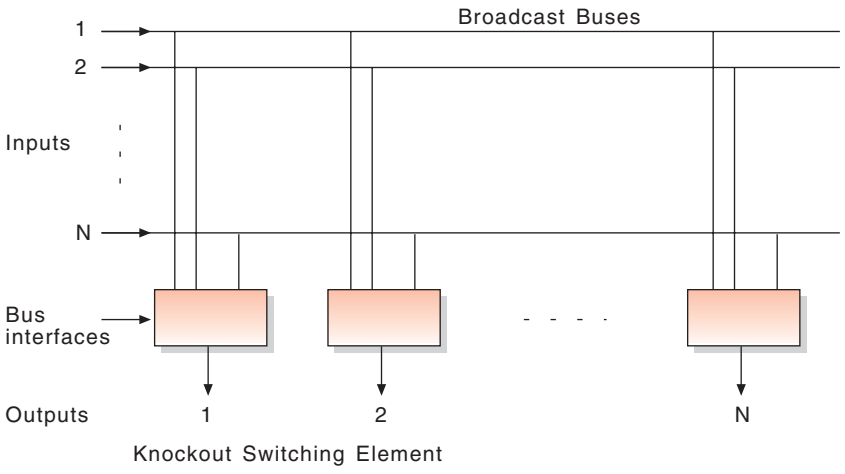


Figure 10.27 Knockout principle

Another switch is the Knockout switching element from ATT Bell Laboratories (Yeh, 1987). The principle is shown in Figure 10.27. With *N* inlets it consists of:

- *N* (broadcast) buses, one for each inlet
- *N* bus interfaces.

A cell arriving on an inlet is broadcast to all interfaces and received at the output interface identified by the header of the cell.

The bus interface executes the following functions:

- Identifying the cells addressed to the output (or filter out the cells not intended for the output)
- Concentrating and reducing the operating speed of the buffers
- Distributing cells to the different buffers
- Buffering cells.

A block diagram of an interface is shown in Figure 10.28.

It is clear that the Knockout switch may result in some cell losses. However, it can be shown that the probability of cell loss is small even with large *N* and a relatively small *L*, where *L* is the number of outputs from the concentrator.

The concentrator is based on a so-called Contention Switch illustrated in Figure 10.29a. Its functions are as follows:

- If only one (ATM) cell is present at an inlet, the cell will be selected as a winner (and routed to the winner output)
- If two cells are present, the left cell will be selected as a winner.

Figure 10.29b illustrates the use of the contention switch in a concentrator with 8 inputs and 4 outputs. In addition to the switch single-input / single-output delay elements marked with a "D" have also been used. The $N = 8$ inputs to the Concentrator (from the cell filters) enter four contention switches of the first stage. The four winners from this first round advance to the second round where they compete with each other using two contention switches. The winners from this second round then advance to the third round, and so on. The result is that four out of the eight originating cells from the cell filters will arrive at the output of the concentrator.

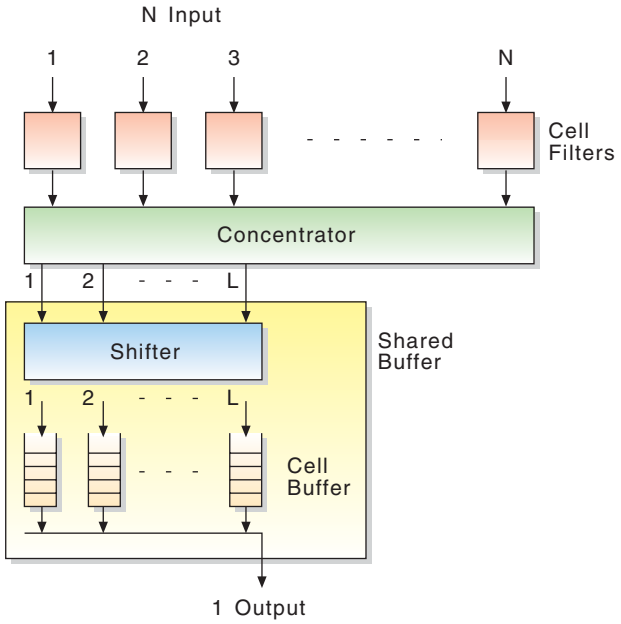
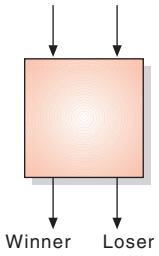
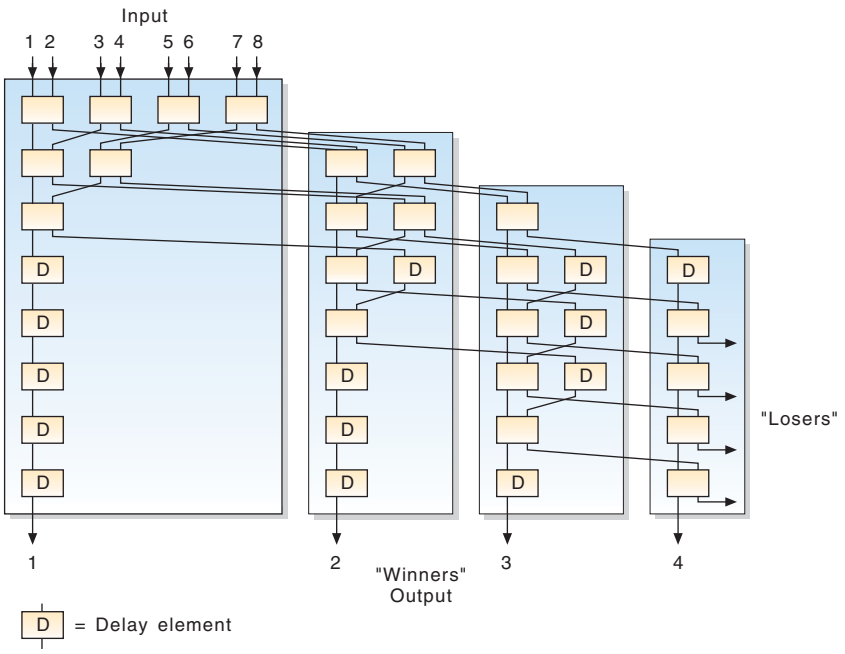


Figure 10.28 Knockout bus interface



a) Contention switch



b) Contention switch (used in a concentrator)

Figure 10.29 Concentrator

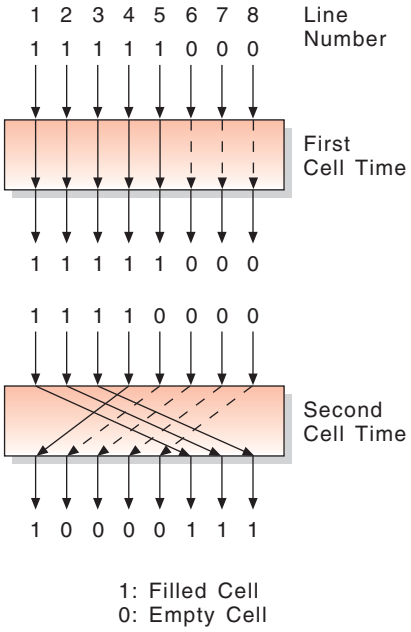


Figure 10.30 Shifting function for 8 inputs

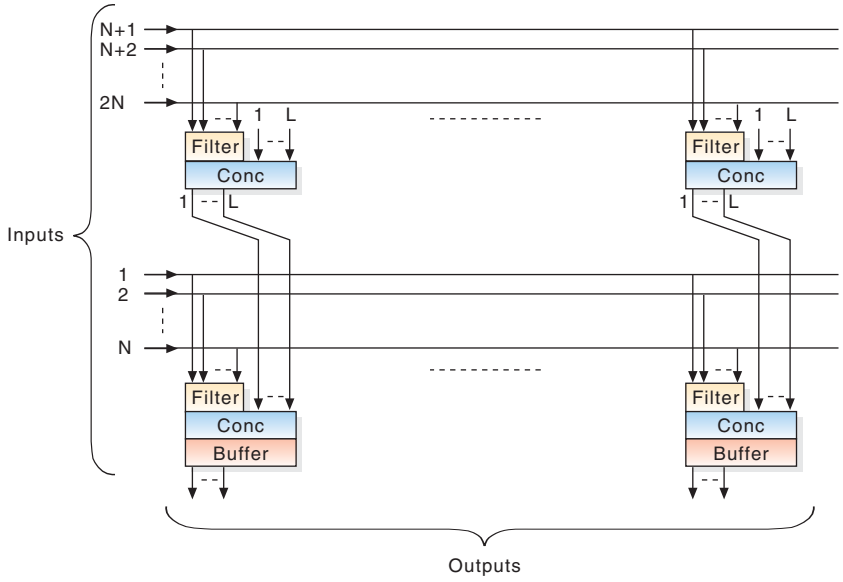


Figure 10.31 Expansion of the Knockout switch

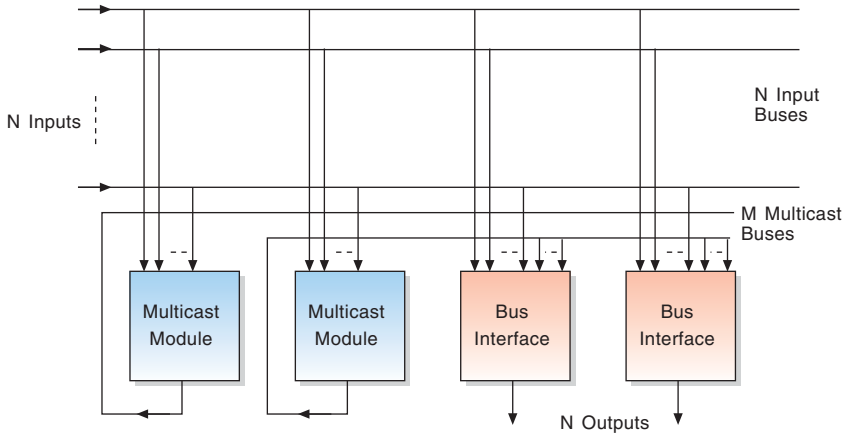


Figure 10.32 Knockout switch with multicast function

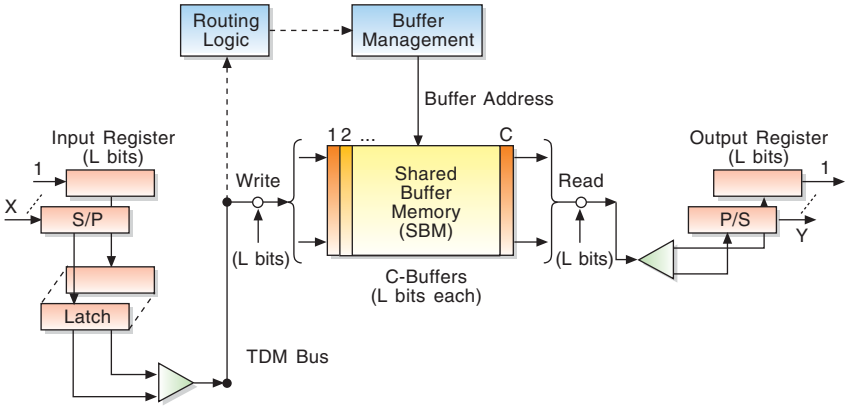


Figure 10.33 Roxanne switching element

The shifting function is illustrated in Figure 10.30 for 8 inputs. It distributes the cells on the buffers (data storages). In order to decrease the necessary operating speed of the output buffer, this buffer is divided into L parts (ref. Figure 10.28). (It may happen that N cells are destined for the same output.)

The shifter provides a circular shifting of the L cells at the inputs to the L outputs. At the first cell time five cells are shifted to the corresponding five outlets. During the next cell time four cells arriving at the first four inlets of the shifter are shifted as shown. During the third cell time the cell at inlet 1 is shifted to outlet 2, and so on.

To expand the Knockout switch the principle illustrated in Figure 10.31 is applied, giving a switch of size $2N$ x $2N$. The number of buses is now $2N$ and the number of buffers is the same. The number of filters and concentrators is now $4N$. The concentrators have $N + L$ inputs and L outputs.

The broadcast function is easily obtained since an input has access to all outputs. A special routing identification can be given to the broadcast cells to allow the filters to recognise those cells.

The multicast function is intended to be solved as illustrated in Figure 10.32. Modules called multicast modules have been added to recognize the multicasting cells. The outputs from these modules are fed back via M multicast buses to the inputs of the other regular modules.

An important functional unit is the cell duplicator in the bus interface of a multicast module. This unit makes a number of copies of the incoming cell, giving each of the copies its own destination (output) address.

The Roxanne switch is based on the central queuing principle. The switch was developed by Alcatel and the basic switching element is shown in Figure 10.33. It has *X* inlets and *Y* outlets. After conversion from serial to parallel, the cells are latched in a register. Via a TDM bus, the *X* cells are transferred to the central queue. The transfer uses one cell time. With *X* cells this means a transfer time of t_{cell} / X. Before writing the cells into the buffer, some routing operations in the routing logic have to be done. Then the cells are written into the buffer under control of the Buffer Management unit. This unit also controls the rest of the processes, administration of free and occupied locations in the buffer, queuing discipline, etc.

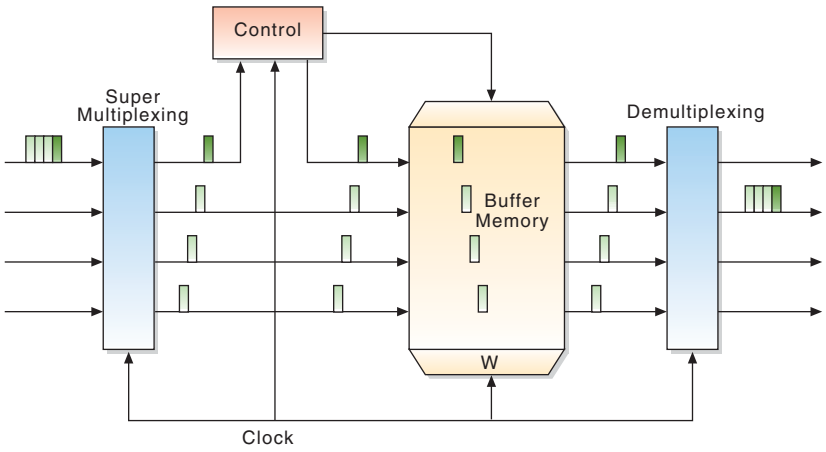
The Coprin switching element is a CNET (France Telecom) solution which was designed to switch data, voice and video. Figure 10.34 shows the principle. It is based on central queuing.

Besides the input functions (synchronization, etc.) the switch, as shown, is based on four units called super multiplexing, buffer memory, demultiplexing and control. The basic functions are:

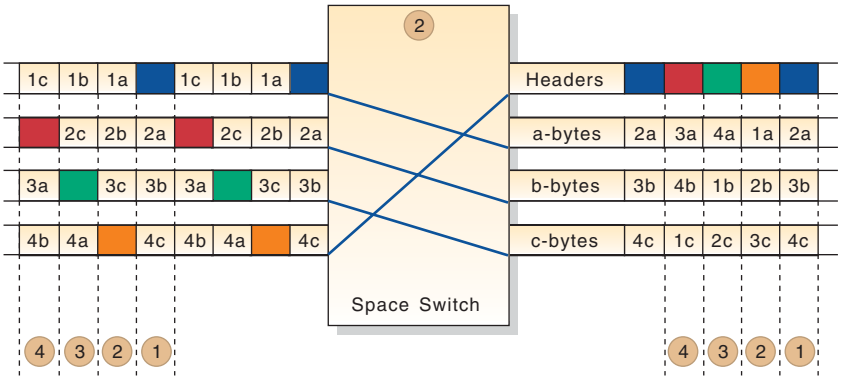
- The super multiplexing unit transposes the incoming cell stream to parallel data streams so that the headers are multiplexed into a special stream and the rest of the cells on another data stream.
- The demultiplexing unit performs the opposite operation of super multiplexing.
- The buffer memory stores the cells and takes care of the organisation of the incoming data.
- The control memory manages the buffer memory (administrating free and occupied locations, etc.).

Figure 10.34b-c shows in more detail the function of the super multiplexing unit with four inputs. The demultiplexing unit functions like the super multiplexing unit but in the "opposite direction".

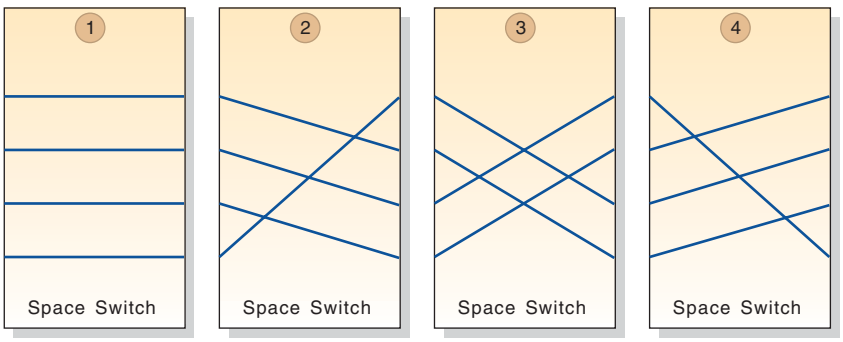
The memory buffers and translation tables are under the control of the control unit.



a) Units in the Coprin switch



b) Super Multiplexing Function in Coprin



c) Four States of the Space Switch

Figure 10.34 The Coprin switch

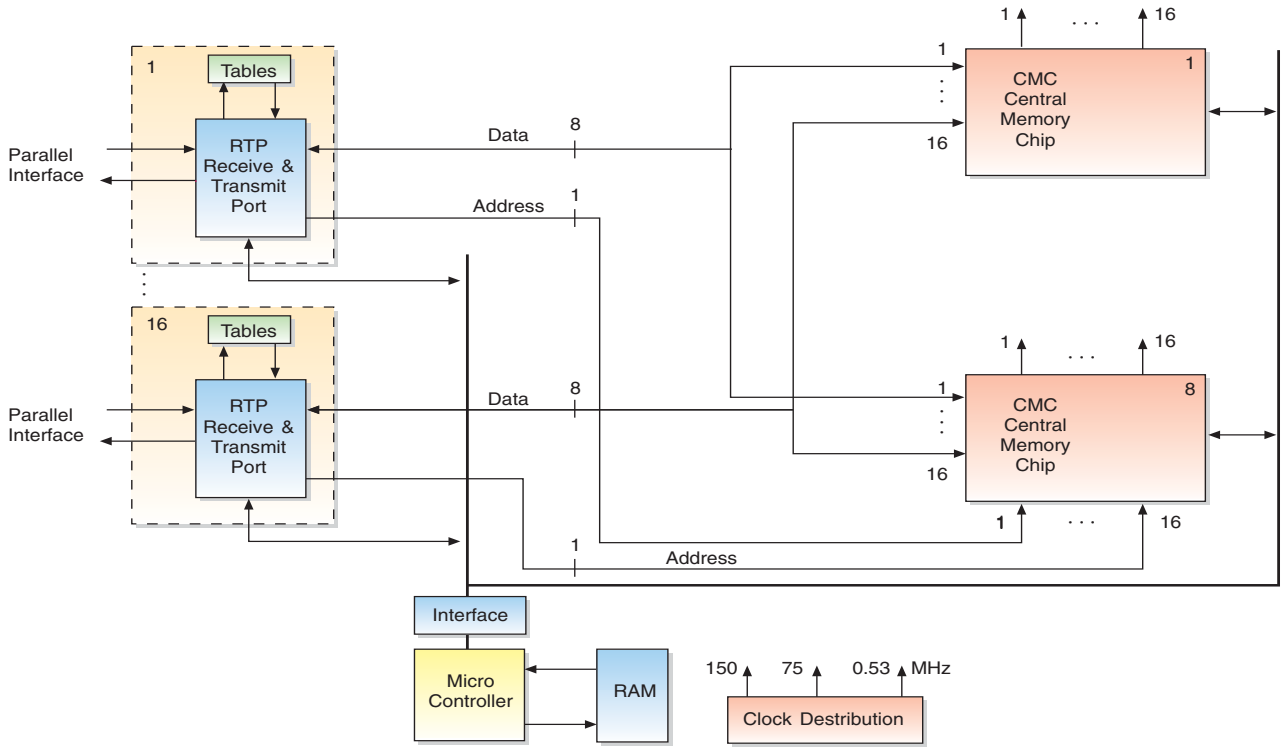


Figure 10.35 Basic building blocks of the Athena switching element

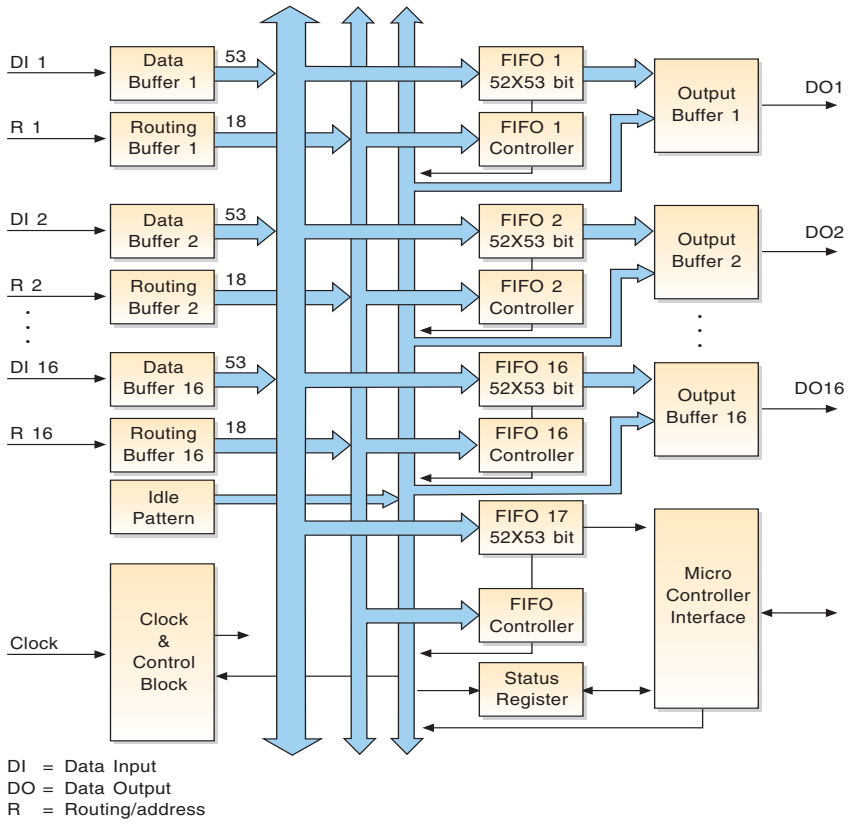


Figure 10.36 Structure of Central Memory Chip of the Athena switch

Another switching principle is used in the Athena switch shown in Figure 10.35 which has the following characteristics:

- Output buffer
- One receive/transmit port RTP per inlet
- Bus between in-/out ports and buffer (bitrate about 9.6 Gb/s)
- Buffer divided into (here) 8 banks CMCs, giving a corresponding reduction in bitrate per in-/outlet
- Each RTP has one outlet to each of the memory banks (8 outlets in the example)
- Each RTP sends the memory address to the buffer.

With 8 CMCs, each of them has to handle 1/8 of the total data and each CMC has to receive identical routing information from every RTP. The routing data is distributed by a physical address bus from each RTP to the 8 CMCs. The addresses specify to which outlets the cells have to be sent. To manage the multicast function, an extra routing bit from each RTP is added to indicate whether or not a cell has to be copied. A microcontroller handles functions like administration of the routing tables, etc.

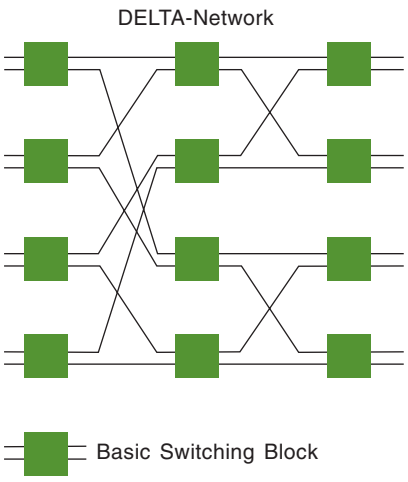


Figure 10.37 Switching fabric based on MIN

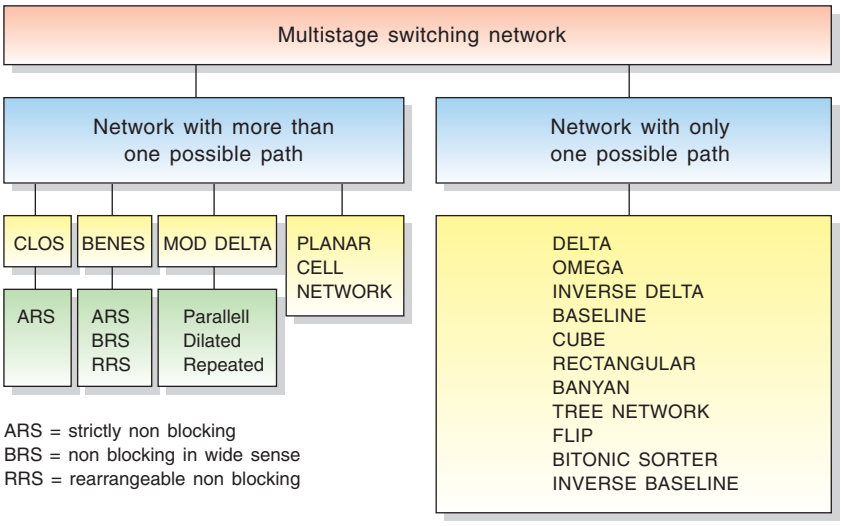


Figure 10.38 Possible network structures for switching of ATM data

Figure 10.36 shows the structure of a CMC. Each CMC has to queue a part of the cell stream (1/8 of the data in this example) and has one data line and one address routing line from each RTP. The corresponding data are stored at the input of the CMC and, via an internal bus, transported to the FIFO selected by the routing information in the Routing Buffer. The switch gives some cell losses, but the loss rate is kept below 10^{-10} for normal cells and below 10^{-15} for high priority cells (when no more than 10 % of the cells have high priority).

10.6.5 ATM switching fabrics

Larger switching networks for ATM data can be constructed by connecting together the elements described above to a MIN (multistage interconnection network) like the one illustrated in Figure 10.37. The resulting network can, like switching networks for switching of "non-ATM data", be classified into two main groups:

- Network with one possible path between a specified pair of inlet and outlet (partial permutation network)
- Network with more than one possible path between an inlet and an outlet pair (complete permutation network).

Section 7 describes several examples of networks belonging to these groups.

The drawback of networks with only one possible path is the possibility of congestion/contention. An advantage is the (relatively) simple routing (self routing) in such networks. It should, however, be noted that complete permutation networks can also be made self routing by distributing the control to the switching elements.

To avoid contention in such partial permutation networks, the methods illustrated in Figure 7.33 could be used.

Some potential switching network structures are given in Figure 10.38.

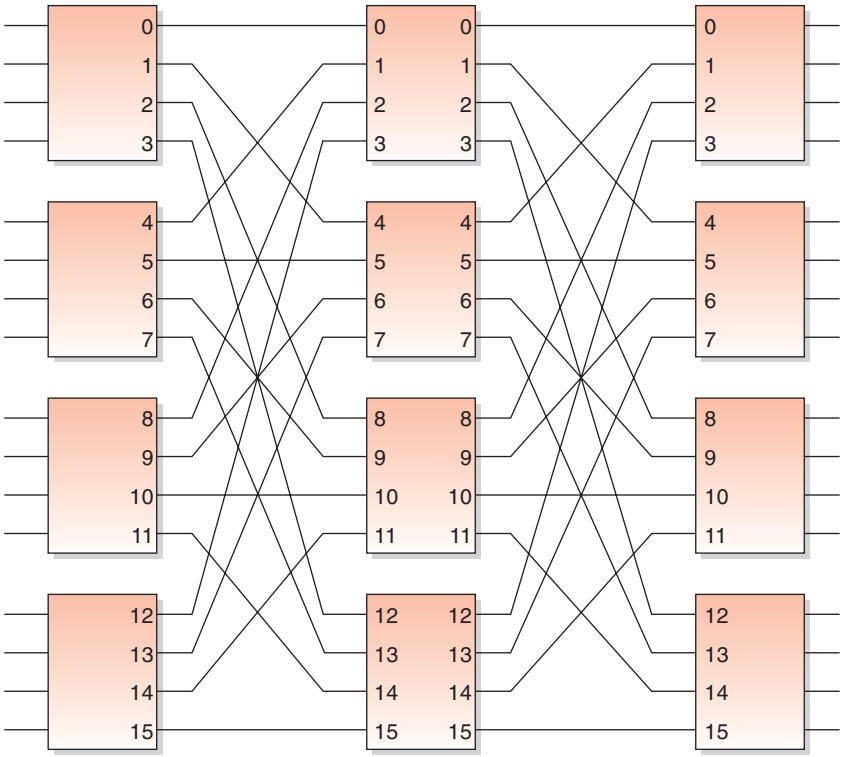


Figure 10.39 Shuffle network with "ATM"-switching elements of size 4 x 4

Figure 10.39 illustrates a structure with a link structure obtained by using 2-shuffle permutations and switching elements of size 4 x 4.

Bibliography

1 Anido, G J, Seeto, A W. Multi-path routing techniques for a class of multistage fast packet switches. *IZS '88,* Zurich, March 1988.

2 Arakawa, N, Noiri, A, Inoue, H. ATM switch for multi-media switching system. *ISS '90,* Stockholm, June 1990.

3 Bauwens, J, dePrycker, M. Broadband experiment using asynchronous time division techniques. *Electrical Communication,* 61(1), 123–130, 1987.

4 Birdsall, T G, Ristenbatt, M P, Weinstein, S B. Analysis of asynchronous time multiplexing of speech sources. *IRE Transactions on Communications Systems,* December 1962.

5 Boettle, D, Klein, M. High speed (140 Mb/s) switching techniques for broadband communications. *Proceedings 1986 International Zurich Seminar on Digital Switching,* 97–100, 1986.

6 Bonatti, M, Fratta, L, Tomasi, C. A call set-up architecture for very large switching systems. *Globecom '86,* Houston, December 1986.

7 Chin, C, Hwang, K. Packet switching networks for multiprocessors and data flow computers. *IEEE Transactions on Computers,* 33(11), 1984.

8 Coudreuse, J P, Servel, M. Prelude : an asynchronous time-division switched network. *ICC '87,* Seattle, June 1987.

9 Denissen, F, Desmet, E, Petit, G H. The policing function in an ATM network. *IZS '90,* Zurich, March 1990.

10 dePrycker, M. *Asynchronous transfer mode : solution for broadband ISDN.* New York, Ellis Horwood.

11 dePrycker, M, Bauwens, J. A switching exchange for an asynchronous time division based network. *ICC '87,* Seattle, June 1987.

12 dePrycker, M, Bauwens, J. A broadband central office for ATM. *International Journal of Digital and Analog Cabled System,* May 1988.

13 dePrycker, M, de Somer, M. Performance of a service independent switching network with distributed control. *IEEE Journal on Selected Areas in Communications,* 5(8), 1987.

14 dePrycker, M, de Somer, M. An ATD broadband switching exchange with distributed control. *CEPT Seminar on BISDN,* Albufeira (Portugal), January 1987.

15 dePrycker, M et al. An ATM switching architecture with intrinsic multicast capabilities for the Belgian broadband experiment. *ISS '90,* Stockholm, May 1990.

16 deSomer, M. Building a mathematical model describing the behaviour of a variable bitrate switching network. *ITC Seminar,* Brussels, May 1986.

17 Devalt, M, Quinquis, J P, Rouaud, Y. Asynchronous time-division switching : a new concept for ISDN nodes. *Proceedings ISS '81,* Montreal, 42b.5.1–42b.5.7, 21–25 Sep., 1981.

18 Dieudonne, M, Quinquis, M. Switching techniques for asynchronous time division multiplexing (or fast packet switching). *ISS '87,* Phoenix, March 1987.

19 Erickson, J et al. A 1.7 Gigabit-per-second, time-multiplexed photonic switching experiment. *IEEE Communications Magazine,* 25(5), 56–58, 1987.

20 Gonet, P, Adam, P, Coudreuse, J P. Asynchronous time-division switching : the way to flexible broadband communication networks. *Proceedings 1986 International Zurich Seminar on Digital Switching,* 141–147, 1986.

21 Hayward, G A et al. CMOS VLSI applications in broadband circuit switching. *IEEE Journal on Selected Areas in Communications,* SAC-5(8), 1231–1242, 1987.

22 Henrion, M et al. Switching network architecture for ATM based broadband communications. *ISS '90,* Stockholm, June 1990.

23 Huang, A, Knauer, S. Starlite : a wideband digital switch. *Globecom '84,* 121–125, 1984.

24 Huber, M N, Rathgeb, E P, Theimer, T H. Self routing banyan networks in an ATM-environment. *ICCC '88,* Tel Aviv, Oct. 1988.

25 Kim, H S, Leon-Garcia, A. Performance of buffered banyan networks under non-uniform traffic patterns. *Infocom '88,* New Orleans, March 1988.

26 Kleinrock, L. *Queueing systems, vol. 1.* New York, John Wiley, 1975.

27 Koinuma, T et al. An ATM switching system based on a distributed control architecture. *ISS '90,* Stockholm, June 1990.

28 Kulzer, J J, Montgomery, W A. Statistical switching architectures for future services. *Proceedings ISS '84,* Florence, 7–11 May, 43.1.1–43.1.6, 1984.

29 Langer, K et al. Broadband switching network and TB switching technology for 70 Mb/s. *Proceedings 1986 International Zurich Seminar on Digital Switching,* 89–96, 1986.

30 Lenart, J, Jou, L, Su, S F. A high speed ring-based switching concept. *Proceedings 1987 International Switching Symposium,* 1, 43–47, 1987.

31 Listanti, M, Roveri, A. Switching structures for ATM. *Computer Communications,* 12(6), 1989.

32 Maxemchuk, N. Routing in the Manhattan street network. *IEEE Transactions on Communications,* COM-35(5), 503–512, 1987.

33 Nojima, S et al. Integrated services packet network using bus matrix switch. *IEEE Journal on Selected Areas in Communications,* SAC-5(8), 1284–1291, 1987.

34 Rathgeb, E. Theimer, T, Huber, M. Buffering concepts for ATM switching networks. *Globecom '88,* Hollywood, Nov. 1988.

35 Richards, G W, Hwang, F. A two-stage rearrangeable broadcast switching network. *IEEE Transactions on Communications,* 33(10), 1985.

36 Saito, Y, Sahita, Y, Morita, S. A study on broadband time division switch. *Proceedings 1986 International Zurich Seminar on Digital Switching,* 101–108, 1986.

37 Sakurai, Y et al. Large scale ATM multi-stage switching network with shared buffer memory switches. *ISS '90,* Stockholm, June 1990.

38 Suzuki, H, Takeuchi, T, Yamaguchi, T. Very high speed and high capacity packet switching. *International Communication Conference Proceedings,* 749–754, 1986.

39 Thomas, A, Coudreuse, J P, Servel, M. Asynchronous time-division techniques : an experimental packet network integrating video communication. *Proceeding ISS '84,* 32c.2.1–32c.2.7, Florence, 7–11 May, 1984.

40 Turner, J S, Wyatt, L F. A packet network architecture for integrated services. *Globecom 1983,* 45–50, 1983.

41 Wu, L T, Huang, N C. Synchrounous wideband network : an interoffice facility hubbing network. *IZS '86,* Zurich, March 1986.

42 Yamanaka, N, Kikuchi, S. High-speed time-division switch for broadband ISDN. *Globecom '87,* 3, 50.7.1–50.7.5, 1987.

43 Yamanaka, N, Miyanaga, H, Yamamoto, Y. High-speed time division switch for 32-Mbit/s bearer rate signals. *IEEE Journal on Selected Areas in Communications,* SAC-5(8), 1249–1255, 1987.

44 Yeh, Y S, Hluchyj, M G, Acampora, A. The knockout switch : a simple, modular architecture for high-performance packet switching. *ISS '87,* Phoenix, March 1987.

45 Yeh, Y, Hluchyj, M, Acampora, A. The knockout switch : a simple modular architecture for high performance packet switching. *IEEE Journal on Selected Areas in Communications,* SAC-5(8), 1274– 1281, 1987.

11 Photonic switching

11.1 Introduction

Optical (or photonic) technology is used in the transmission part of telecommunications and is going to give a dramatic increase in the users' bitrate (bandwidth). This is appropriate with the ever increasing need for data transfer. A short presentation, like the one in Table 11.1, shows the demand for bitrates of some existing and future communication services. The table also shows the corresponding number of channels given by an optical fibre with a bitrate of 10 Tb/s.

The main advantages of light (the carrier of information in photonic components) are mentioned in Section 5.

A short overview of components necessary to construct optical switches in Section 5 is given, and some further discussion will be given in this section, mostly with respect to logic.

A serious problem in photonic switching is the variation of the polarisation of the light. This means that, in a practical solution, the components have to be made independent of polarisation and so give more complex and uneconomical components.

11.2 Photonic logic

An idealised logic element has a principal characteristic as shown in Figure 11.1a. Optical input power greater than a limit P_t changes the outlet power from P_0 to P_1. $P_1 > P_0$ means amplification. With standard levels P_0 and P_1, such a characteristic is essential for a logic element.

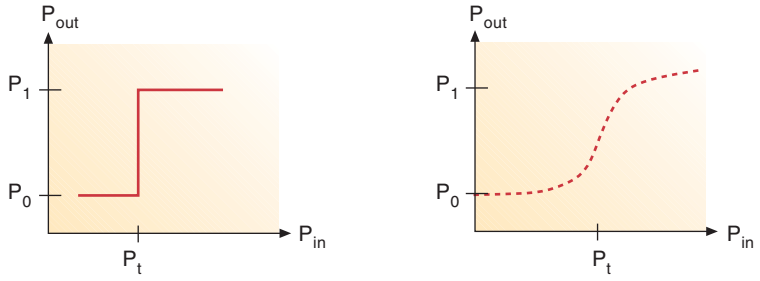
The element operates as an "and" gate with input signal k_x and k_y each with level P_k and with $2P_k \geq P_t$ (see Figure 11.2a). A logic "or" gate can be realised by using input signal $P_k \geq P_t$ (Figure 11.2b). In both cases a bias power P_m can be used to give a suitable working point.

The characteristic of a practical element is shown in Figure 11.1b. The element can be used as an amplifier and it is then favourable with a piecewise linear characteristic.

Table 11.1 Some services and associated bitrates

Service	Bitrate bit/s	Approximate number of channels
Picture phone (coded)	64 k – 384 k	100 M – 10 M
High speed data communication	2 M – 140 M	5 M – 0.1 M
TV (normal)	34 M – 140 M	0.4 M – 0.1 M
HDTV	140 M →	0.1 M –
ATM based on 600 Mb/s	600 M	160 000

k = 10^3, M = 10^6, T = 10^{12}, Bitrate of transfer medium = 10 Tb/s



a) Idealised curve b) Practical curve

Figure 11.1 Characteristics of a logical element

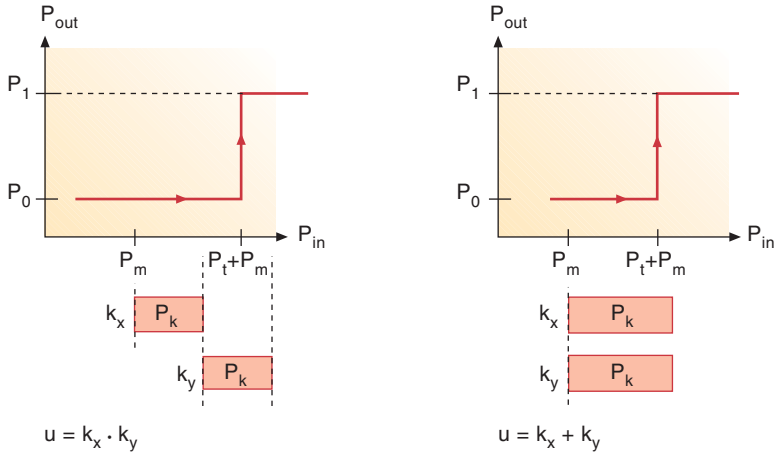


Figure 11.2 Characteristics of logical gates

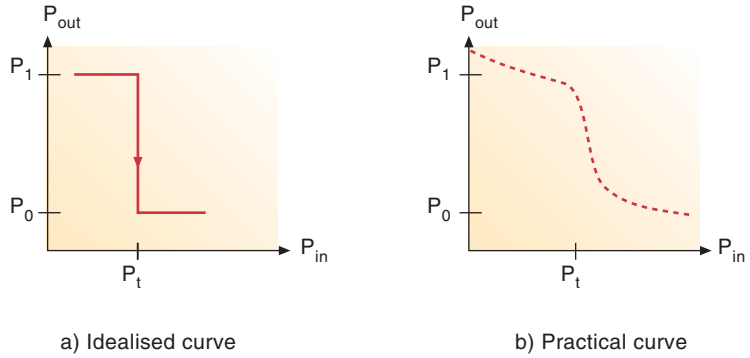


Figure 11.3 Characteristics of an inverter

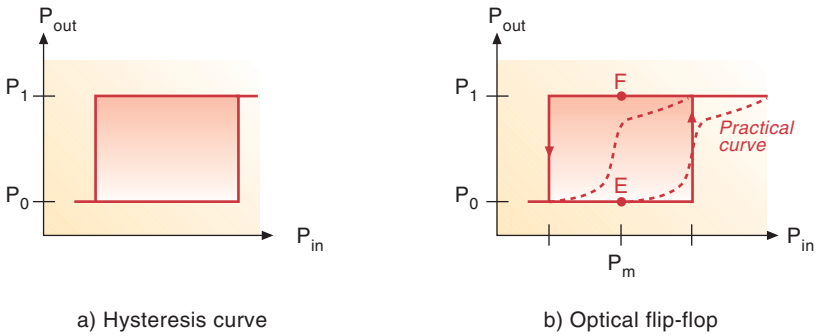


Figure 11.4 Characteristics of optical latch

Figure 11.3a shows the idealised characteristic of an inverter. P_{in} greater than a limit P_t gives $P_{out} = P_0$. P_{out} will change to P_1 if P_{in} is reduced.

A practical element has a characteristic as in Figure 11.3b.

Figure 11.4a shows an (idealised) hysteresis curve needed for implementing an optical latch (flip-flop).

For input power P_m there are two stable points E and F. By adding an (optical) pulse when in state E it is possible to reach state F when removing the pulse. By reducing the bias P_m the element which is in state F, goes to state E after the reduction.

Photonic logic can be divided into two main groups:

- Devices working with both optical and electrical/acoustical/thermal power
- Devices working with only optical power.

A component in the first group is the proposed SEED structure (SEED = Self-Electro-optic Effect Device) shown in Figure 11.5a. A SEED requires both optical and electrical power and consists of a P_{in} diode with Multiple Quantum Well (MQW) material in the intrinsic region and a resistor connected between the diode and a voltage level V. A possible characteristic curve for this device is shown in Figure 11.5b. The structure has two inputs: a signal beam and a bias beam. With only the bias beam, the structure is nearly transparent and almost all the bias energy is passing through the device. When the signal beam is present, the device absorbs energy and little energy passes the element. When a sufficiently strong current is achieved, the positive feedback allows the device to retain its state after the light source is removed.

It is expected that the possible integration will enable 10,000 – 100,000 SEEDs in an area of 1 cm^2. Optical energy to drive the SEED is reported to about 4 fJ/µm^2 and electrical energy to about 16 fJ/µm^2.

Important parameters are the fan-in/fan-out parameters. The tolerances illustrated in Figure 11.6 determine the usable regions of these parameters.

Fan-in is given by:
$n_{in} = (P_{imax} - P_t) / P_s$

Fan-out is given by:

$n_{out} = (P_{1min} - P_{0max}) / P_s$

Example of a photonic component in the second main group is the Non-Linear Fabry-Perot Etalon (NLFPE) used as a gate or bistable component. This is illustrated in Figure 11.7. A non-linear material is placed between two mirrors. Without the signal beam P_s the (optical) bias beam P_t passes through the element. The signal beam (together with P_t) changes the refraction index of the non-linear material so that the optical power is reflected. The transfer curve is like the curve of the SEED element and in this example operates as a NOR gate. However, it can also be designed to operate as AND, OR, NAND or XOR gate.

Optical switching energy is about 40 fJ/μm^2 and reported switching time about 100 ps.

A structure using the NLFPE is shown in Figure 11.8. It is a pulsed device using two different wavelengths λ_1 and λ_2. The structure is known under the name "Optical Logic Etalon", OLE. Input data uses λ_1 and the clock uses λ_2. As shown in the figure the input power occurs a short time before the clock power. Ideally, the input power is chosen so that "1" corresponds to an absorption peak of the non-linear material (in OLE). Then almost all energy of the incoming data power is absorbed. This absorption changes the index of refraction of the non-linear material and so shifts the resonant peak by changing the optical path length of the cavity, and the OLE becomes transmissive. When the clock power arrives, a short time later, it will pass through to the output. With no data power occurring, the OLE becomes non-transmissive and so (ideally) no clock power is transferred to the output.

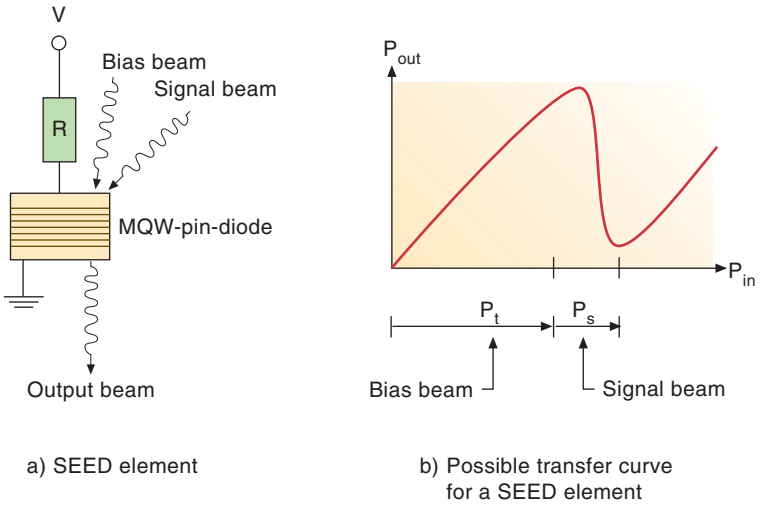


Figure 11.5 SEED element

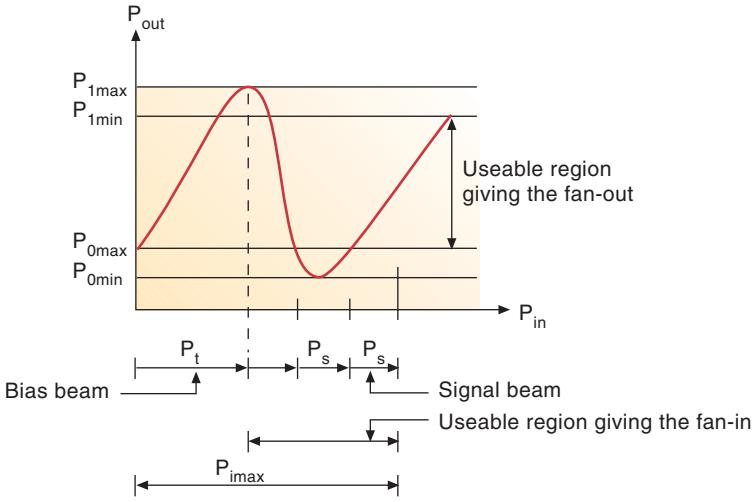


Figure 11.6 Fan-in / fan-out regions for SEED element

11.3 Storage components

To implement a switching system there is generally a need to store data in some form (digital or analogue), so a convenient storage medium is essential. Compared to electronics, however, this is not solved satisfactorily by use of photonic technology. The today's solution can be divided into two main groups:

- Optical delay lines
- Semiconductor based storage elements.

An optical delay line can be implemented by using an optical fibre and a switch to couple light in and out of the fibre as

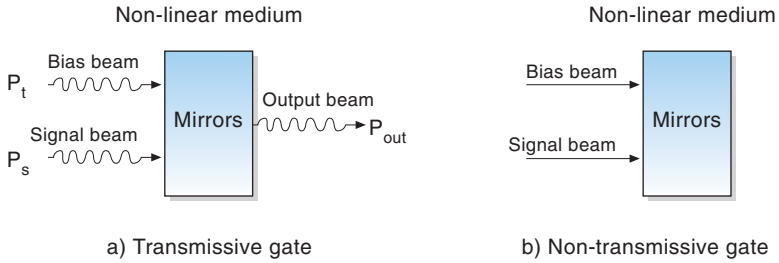


Figure 11.7 The Non-Linear Fabry-Perot Etalon optical gate

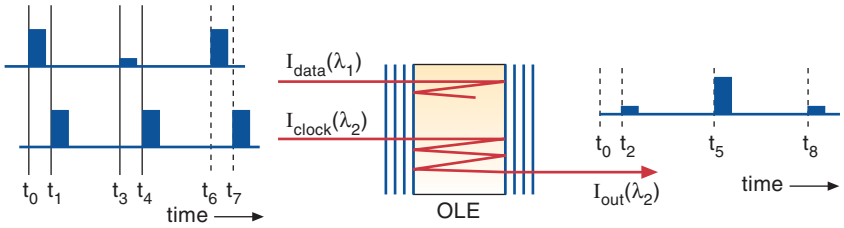


Figure 11.8 OLE as a NOR gate

illustrated in Figure 11.9. Light coupled into the fibre circulates as long as the switch is in the straight state. When light arrives at the switch, it is disconnected if the switch is in the cross state. The time to pass the fibre depends on the length of the fibre and on the speed of the light. The optical switch can be a directional coupler.

The basic construction is quite simple but there are some practical problems. The most serious one is the loss in the switch and the fibre since this limits the number of round trips of the light and so limits the delay. In later years, however, the possibility of using the optical fibre amplifier has appeared, which eliminates the loss problem. Still, such storage elements cannot be integrated in the same way as the electronic memory, and they give a large and inconvenient storage medium (compared with the electronic memory).

Other methods to implement optical storages are to use bistable laser diodes. The SEED elements can also be used as bistable components and then also for storage purposes.

All these components use electrical power and can be classified as belonging to the first main group (i.e. elements using both optical and electrical power). This can be a drawback in some situations.

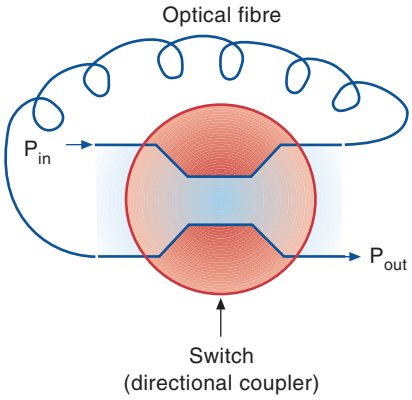


Figure 11.9 Optical delay line

11.4 Optical space switching

Structures for optical space switching are much the same as when using electronic elements. Of course, attention must be paid to the special properties of photonic components. Optical switching elements can, like for electronic elements, be divided into the relation group and the logic group.

The directional coupler is the best known optical/photonical relation element and can be used to construct various types of switching structures. The directional coupler (here called a b-element) has the "straight" state and the "cross" state as illustrated in Figure 11.10, which also shows how the input power is split on the outlets.

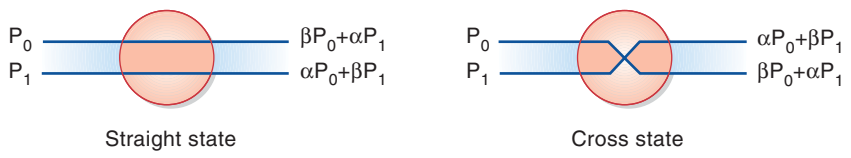


Figure 11.10 Possible states of the β-element

A few examples of the most common switching structures based on β-elements are shown in Figure 11.11.

Figure 11.11a shows a single quadratic matrix. When connecting inlet i and outlet j, the following operations have to be executed:

- Switch sw_{ij} is set to the straight state
- Switches sw_{i0}, sw_{i1}, ..., $sw_{i(j-1)}$ are set to cross state
- Switches sw_{0j}, sw_{1j}, ..., $sw_{(i-1)j}$ are set to cross state.

The number of β-elements in a matrix of size $(n + 1)$ x $(n + 1)$ is:

$$K_{n+1} = (n + 1)^2$$

The number of elements k_s the signal has to pass in a connection will vary between the limits:

$$1 \leq k_s \leq 2n + 1$$

Using the notation in Figure 11.10, the input power P_0 on input i which will be transferred to output j can be written:

$$P_{ij} = P_0\beta^i\beta^{j+1} = \beta^{i+j+1}P_0$$

Example:

$\beta = 0.99$ and $i = 0, j = 0$ gives:
$P_{00} = \beta P_0 = 0.99P_0$

$\beta = 0.99$ and $i = 7, j = 7$ gives:
$P_{77} = \beta^{15}P_0 = 0.86P_0$

Additional losses are due to the connection between elements and between inlets and outlets and the "outer world". The losses between the elements can in most cases be neglected relative to losses at inlet and outlet connections.

Another important parameter is the crosstalk. Referring to Figures 11.10 and 11.11a the contributions to the crosstalk from surrounding switching elements are given by terms of the form $\alpha^x\beta^y$ where x and y are integers > 1. Greatest contributions are terms having $x = 1$ since $\alpha^2 << \alpha^1 << 1$ (and $\alpha << \beta$). It can be assumed that no power is connected back in a β-element, i.e. from one inlet to the other inlet on the same element.

Example 1:

For $sw_{00} = 1$ (switch is in the straight state = connection between inlet 0 and outlet 0) and power on the other inlets, the contribution can be written:

$$\Delta P_{00} = \alpha^2 P_1 + \alpha^2\beta P_2 + \alpha^2\beta^2 P_3 + \alpha^2\beta^3 P_4 + \alpha^2\beta^4 P_5 + \alpha^2\beta^5 P_6 + \alpha^2\beta^6 P_7$$

$P_i = P_0$ for all i gives for the matrix of size $(n + 1)$ x $(n + 1)$:

$$\Delta P_{00} = (1 + \beta + \beta^2 + ... + \beta^{n-1})\alpha^2 P_0 = \left(\sum_{i=0}^{i=n-1} \beta^i\right) \cdot \alpha^2 P_0$$

The power connected from inlet 0 to outlet 0 is given by:

$$P_{00} = \beta P_0$$

Example 2:

For $sw_{70} = 1$ (switch is in the straight state = connection between inlet 7 and outlet 0) and power on the other inlets, the contribution can be written:

$$\Delta P_{70} = \alpha P_0 + \alpha\beta P_1 + \alpha\beta^2 P_2 + \alpha\beta^3 P_3 + \alpha\beta^4 P_4 + \alpha\beta^5 P_5 + \alpha\beta^6 P_6$$

$P_i = P_0$ for all i gives, for a matrix of size $(n + 1)$ x $(n + 1)$,:

$$\Delta P_{70} = (1 + \beta + \beta^2 + ... + \beta^n)\alpha P_0 = \left(\sum_{i=0}^{i=n} \beta^i\right) \cdot \alpha P_0$$

The power connected from inlet 7 to the outlet 0 is now given by:

$$P_{70} = P_0\beta^7\beta^{0+1} = \beta^8 P_0$$

Comparing the result of these two simple examples show that both loss and crosstalk depend on the position of the element used for the connection or on which inlet and outlet are to be connected. In the general case it is difficult to calculate the crosstalk exactly. It is possible, however, to find a "worst case" when terms containing α^2 are neglected. It turns out that the "worst case" for a matrix of size $(n + 1)$ x $(n + 1)$ as in Figure 11.11a is when $sw_{n0} = 1$, and with power on all other inlets this gives:

$$\Delta P_{n0} = (1 + \beta + \beta^2 + ... + \beta^n)\alpha P_0 = \left(\sum_{i=0}^{i=n} \beta^i\right) \cdot \alpha P_0$$

$P_i = P_0$ for all input numbers = i.

A possible drawback with the configuration in Figure 11.11b is the crossing of the waveguides in the switch. However, if the crossing angle is over a certain value, the losses can be acceptable.

The drawback of crossing waveguides is avoided in the "planar" structure in Figure 11.11c. The total number of elements

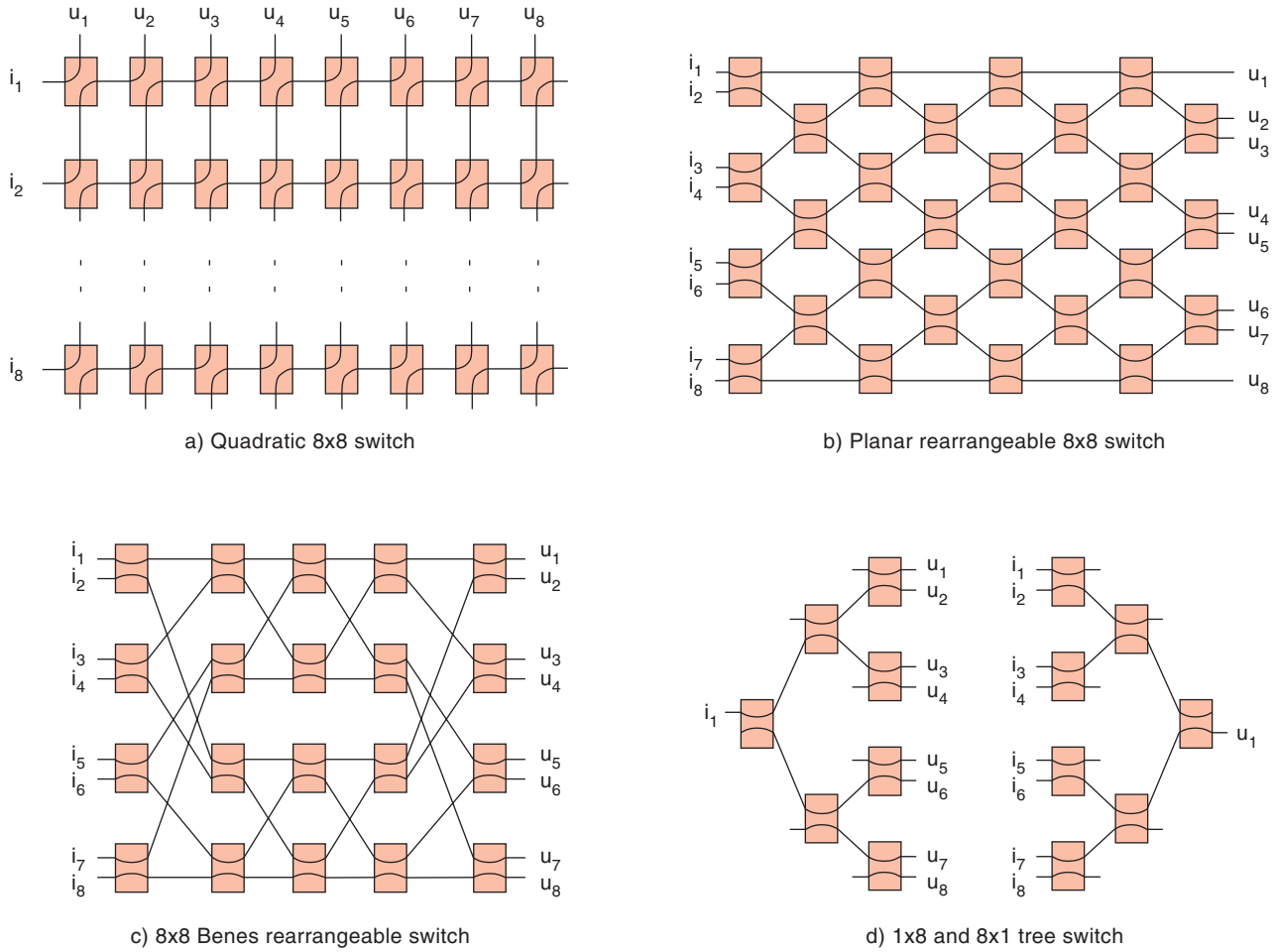
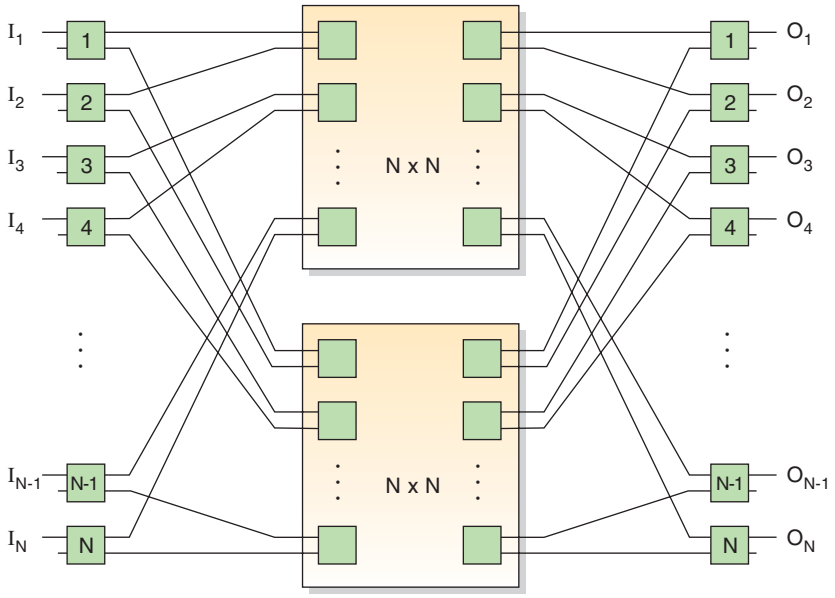
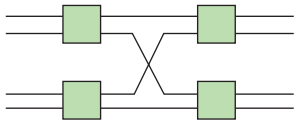


Figure 11.11 Switching structures based on optical β-elements



a) Dilated Benes network



b) Last recursion stage

Figure 11.12 Construction of dilated Benes network

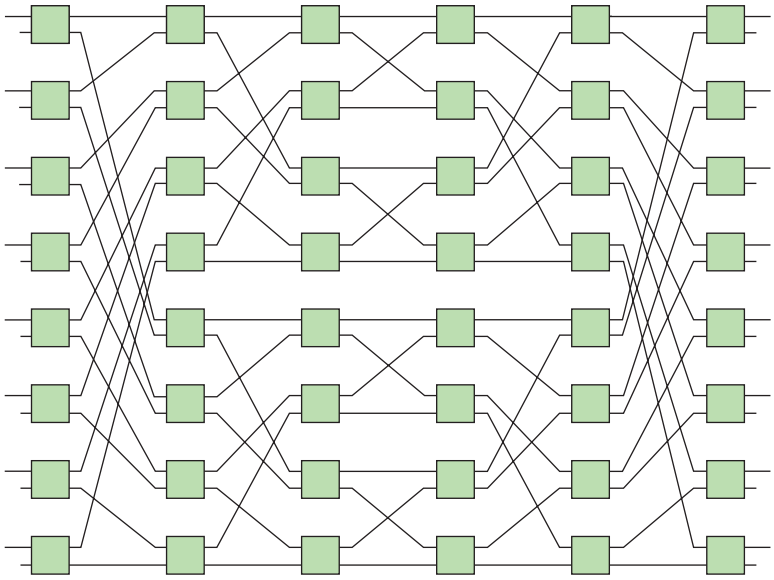


Figure 11.13 8 x 8 dilated Benes network

(β-elements) in a planar switch of size $(n + 1)$ x $(n + 1)$ is:

$$K_n = n(n + 1) / 2$$

The number of elements a signal has to pass is:

$$(n + 1) / 2 \leq k_s \leq (n + 1).$$

The expression for the contribution to crosstalk in such a switch, can be written as a series:

$$\Delta P = \alpha\beta^x + \alpha^2\beta^y + \alpha^3\beta^w + ...$$

where

α and β are as above and x, y, w, ... are integers > 1.

Neglecting terms with higher order of α (i.e. α^2, α^3, α^4, etc.), it is seen that crosstalk is directly proportional to α, so the crosstalk is no better than the crosstalk in a single element.

At the cost of more components (β-elements), some methods have been proposed to get a better crosstalk value. An example of such a network is shown in Figure 11.12a and is named dilated Benes network. This network is constructed much in the same way as shown in Section 7 and repeated in Figure 11.12a. The difference is that the outer stages use only one inlet/outlet and that the recursion stops with 4 x 4 subnetworks which is constructed with two stages of 2 x 2 switches as shown in Figure 11.12b. Another difference is the routing of paths. In this case no 2 x 2 switch carries more than one path.

Due to the construction and routing algorithm, the contribution to the crosstalk contains only terms proportional to α^2. A drawback is the set-up time (as for all rearrangeable networks).

Figure 11.13 shows a dilated Benes network of size 8 x 8.

The number of stages in this network is:

$$t = 2\log_2(n + 1)$$

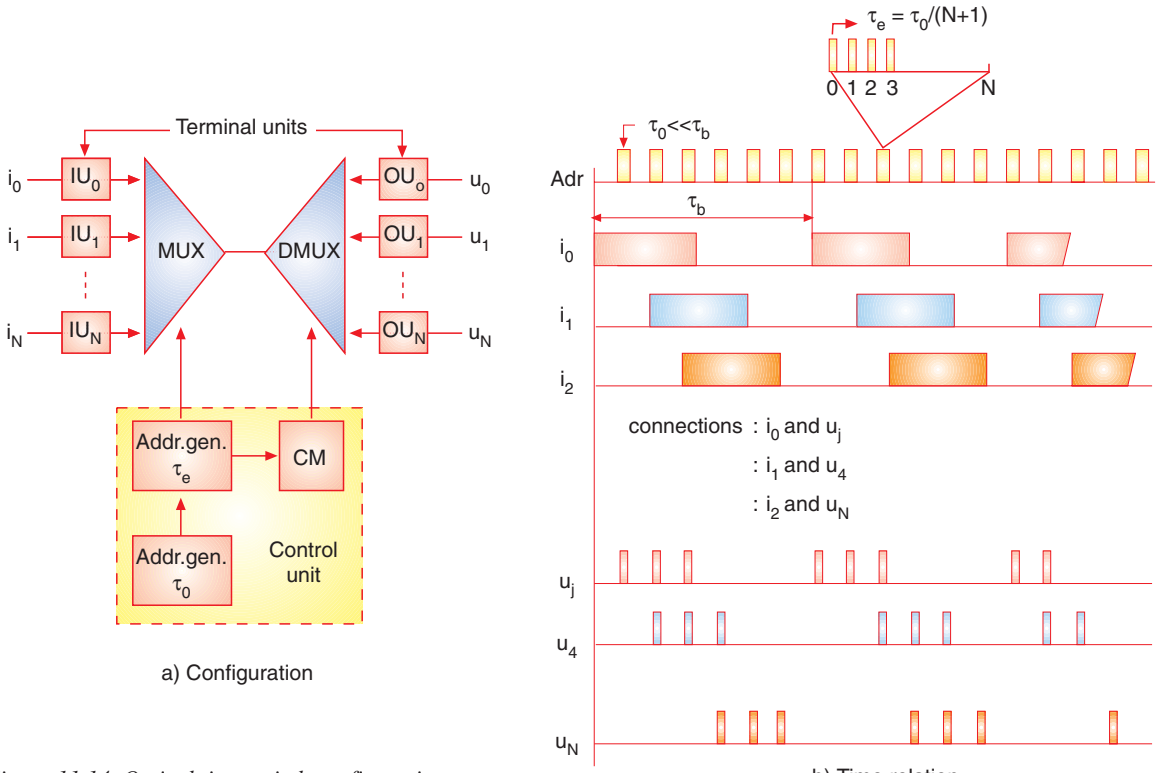


Figure 11.14 Optical time switch configuration

The tree structure in Figure 11.11d is of interest since it gives improved crosstalk and has no crossing of waveguides within the tree. The number of trees in a structure of size $(n + 1)$ x $(n + 1)$ is given by:

$$t = 2\log_2(n + 1)$$

where

$n + 1$ = number of outlets from each tree = number of inlets/outlets.

The total number of switches (β-elements) can be written as:

$$K_s = 2(n+1)\left(\sum_{x=1}^{x=\log_2(n+1)} (n+1)/2^x\right)$$

The number of switches (β-elements) serially connected is given by:

$$k_s = 2\log_2(n + 1).$$

The distribution of the power on the outlets of a tree can be written as:

$$P_i = \binom{(t+1)}{i} \alpha^i \beta^{(t+1-i)} P_0$$

where

$\alpha + \beta$ = 1; $\alpha \ll \beta$

$t + 1$ = number of stages in the tree

P_0 = power on the input of the tree

$\binom{(t+1)}{i}$ = the binomial coefficient.

Example:

$t + 1$ = 3 (3 stages correspond to a tree with 8 outlets)

$p_0 = \beta^3 \cdot P_0$

$p_1 = 3\alpha\beta^2 \cdot P_0$

$p_2 = 3\alpha^2\beta \cdot P_0$

$p_3 = \alpha^3 \cdot P_0$

For one connection, an inlet tree and an outlet tree are involved in such a tree structure.

The contributions to the crosstalk from inlet trees not involved in the connection are difficult to calculate exactly. A "worst case expression" can be given as follows:

$$\Delta p = (\alpha^2\beta^{t-1}2^n)P_0$$

where

α and β are as above and have the same values for all elements

t = number of stages = $\log_2(n + 1)$

$n + 1$ = number of inlets = number of outlets

P_i = P_0 for all i = 0, 1, 2, ..., n

It can be shown that the unwanted power transfer depends only on the position of the outlet and not on which inlet and outlet trees the connection involves.

Within a tree (ref. Figure 11.11d) there is no crossing waveguides, but when connecting the trees to a network, there will generally be several such crossings. This can be avoided if each tree is made on a separate substrate. By choosing the right physical parameters it should be possible to place the substrates as shown so that only free air connection can be used.

Table 11.2 shows sizes of the networks for an example of allowed total loss and given loss per β-element. It is seen that

Table 11.2

Type of network (with β-elements)		Number of inlets N	N for A = 20 dBm	
			a_c = 2 dBm	a_c = 1 dBm
Quadratic matrix	(Figure 11.11a)	$N = \lfloor (A + a_c)/2a_c \rfloor$	5	10
Rearrangeable network	(Figure 11.11b)	$\lfloor \log_2 N = (A + a_c)/2a_c \rfloor$	45	1448
Planar network	(Figure 11.11c)	$N = \lfloor A/a_c \rfloor$	10	20
Active input and output tree network	(Figure 11.11d)	$\lfloor \log_2 N = A/2a_c \rfloor$	32	1024
Active splitter and passive (tree) 3-dB-combiner		$\lfloor \log_2 N = A/(a_c + 3) \rfloor$	16	32
A = total loss a_c = loss per element				

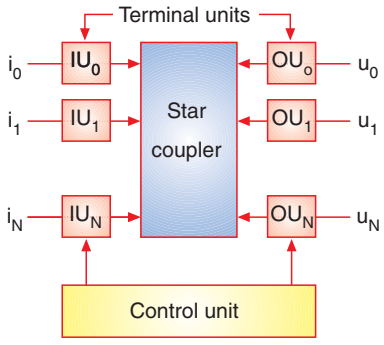


Figure 11.15 Optical time switch with star coupler

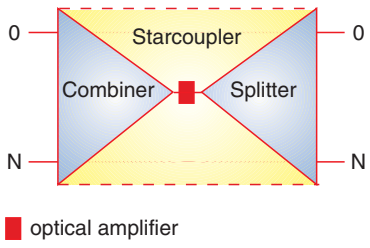


Figure 11.16 Optical time switch with combiner and splitter

the element loss is very important in the construction of large networks.

11.5 Optical time switching

In time switching a time slot is assigned a specific connection, and it is always assumed that the switching times in such networks are so small that the data can be transferred and detected with sufficient reliability.

A simple system to achieve time switching is shown in Figure 11.14. It consists of:

- An optical multiplexer, MUX, and an optical demultiplexer, DMUX, implemented, for example, by (optical) directional couplers
- Address generators, generating time pulses τ_e and τ_0
- An electronic control storage, CM
- Input units, IU, and output units, OU.

All inlets have to be scanned during the time interval τ_0 so this interval must be divided into intervals given by:

$\tau_e = \tau_0 / (N + 1)$;
$N + 1$ = number of inlets
(= number of outlets)

The process is as follows for establishing connection between i_0 and u_j:

- The MUX is connected to i_0 during the first address pulse τ_e
- The control store is addressed at the same time and gives the address u_j to DMUX such that output u_j is addressed and input i_0 and output u_j are connected during the time τ_e
- During the next τ_e pulse, input i_1 and, say, output u_4 are connected as above (address u_4 given by the control store).

The process is repeated in a cyclic way. It is assumed that the signal is regenerated to normal bit length and level in the outlet unit. To transfer analogue signals, this method demands filter and a sampling frequency ($1 / \tau_0$) satisfying Shannon's theorem.

Figure 11.15 shows another simple method based on:

- A directional coupler, sw_i, for each inlet and outlet (within the terminal units)
- A star coupler of size $(N + 1)$ x $(N + 1)$
- A detector for each outlet (within the terminal unit)
- A control unit.

The switch uses the broadcasting principle so that the power from an input, addressed by the control unit, is distributed to all outlets. The control unit controls which outlet the power is transferred to by setting the outlet switch to the proper state. The functional behaviour is much the same as above (Figure 11.14). The drawback of this method is the losses due to the distribution of power to all other outlets.

The star coupler can be replaced by a fan-in and fan-out circuit (Figure 11.16) implemented by elements described in chapter 5. An advantage is that an optical amplifier between the fan-in and fan-out can compensate the losses.

The fact that no memory is needed for the two methods above is an important advantage since no practical usable memory components yet exist when the number of inputs/outputs increases to a usable size and the incoming bitrate is high.

It should also be noted that since the fibre is nearly an ideal transmission medium, it should be possible to multiplex analogue sources and then transmit the samples on a fibre with reasonable length (several kilometres) to the users and thereby avoiding the analogue/digital conversions.

For a given τ_0, the maximum number of inlets/outlets these methods give can be written as:

$$N_{\max} = \lfloor \tau_b / k\tau_0 \rfloor$$

where $k \geq 1$ expresses that more than one pulse may have to be detected before the response is accepted, and τ_b is the bit duration.

Example:

τ_0 = 1 ns corresponding to a bitrate of 1 Gb/s and k = 2 gives:

1. τ_b = 488 ns (first order PCM):

   N_{max} = 244 with serial transfer

   N_{max} = 1952 with 8 bit parallel transfer

   Time per transfer: τ_e = 4.09 ps

2. τ_b = 1667 ps (ATM):

   N_{max} = 1 with serial transfer

   N_{max} = 424 with 424 bit (= 1 cell) parallel transfer

   Time per transfer: τ_e = 2.35 ps

It is seen that these methods put high demands on the components controlling the processes (switches, etc.) when the number of inputs/outputs increases to a usable size and the incoming bitrate is high. At high bitrates, more attention must also be given to varying time delays resulting in difficulties to synchronize the processes. Another serious problem is that today's level of optical technology needs electronics to drive the switches, and this is a bottleneck at high bitrates.

11.5.1 Optical time slot interchanger – OTSI

An important functional component in time switching is the time slot interchanger (TSI) which interchanges data belonging to different time slots.

The basic unit of a TSI is the data storage and this is the most difficult unit to implement in a practical way with optical components. A possible data storage is the fibre delay line where data are inserted into a fibre (waveguide), circulated there and then retrieved at an appropriate time.

The task of the time slot interchanger is to interchange the data in the time slots which can be done as indicated in Figure 11.17 for an OTSI with 4 time slots.

(This corresponds to a switch of size 4 x 4.) It consists of:

- Demultiplexer (DMUX) and multiplexer (MUX) implemented by optical switches (β-couplers)
- Optical delay lines with delay Δτ, 2Δτ, 3Δτ and 4Δτ
- Control units 1 and 2.

The control unit 1 mainly serves as a time generator and scans the input to MUX sequentially.

Figure 11.17 shows an example of the data interchanges.

The following operations have to be executed:

- During time slot t_0 data d_0 is sent to delay Δτ
- During time slot t_1 data d_1 is sent to delay 2Δτ
- During time slot t_2 data d_2 is sent to delay 3Δτ
- During time slot t_3 data d_3 is sent to delay 4Δτ.

Instead of delay lines with permanent delay, we can use delay lines with variable delay. The optical signal circulates in such a line as long as necessary and is coupled in and out via a switch when necessary.

Another method to achieve time slot interchanging (in the sense explained above) is illustrated in Figure 11.18.

The interchanger consists of:

- N optical switches of type (1 x 2)
- N programmable optical delay lines

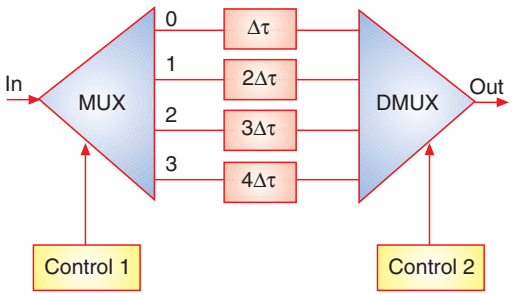


Figure 11.17 Time switch interchanger with fixed (optical) delay as storage

- A star coupler
- A control unit.

Optical input data consist of time multiplexed channels of total length T. The time frame is divided into N + 1 slots each of duration τ. The first time slot is reserved for an synchronous or reference pulse and the rest (N slots) the data.

It functions as follows:

- First the switches are set to the straight state, i.e. all are connected in series
- Data entering the input of the first switch propagates along the line of switches
- When the front of the S-pulse reaches the control unit, this unit changes the states of all switches in the switch-line
- The data pulse will now be sent to the associated delay lines
- After the programmed delay, the data appear at the output of the star coupler in an order determined by the programmable delay lines.

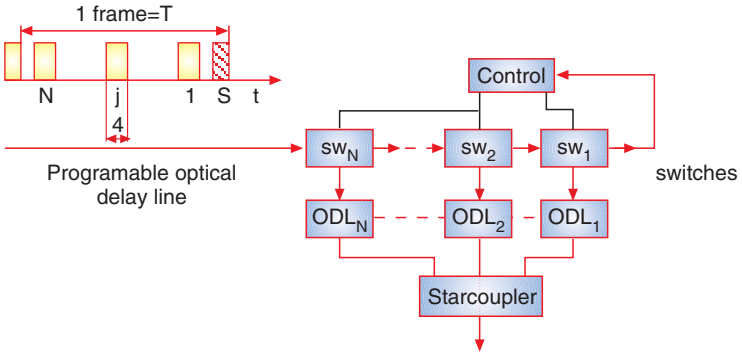


Figure 11.18 Time switch with programmable delay lines as storage

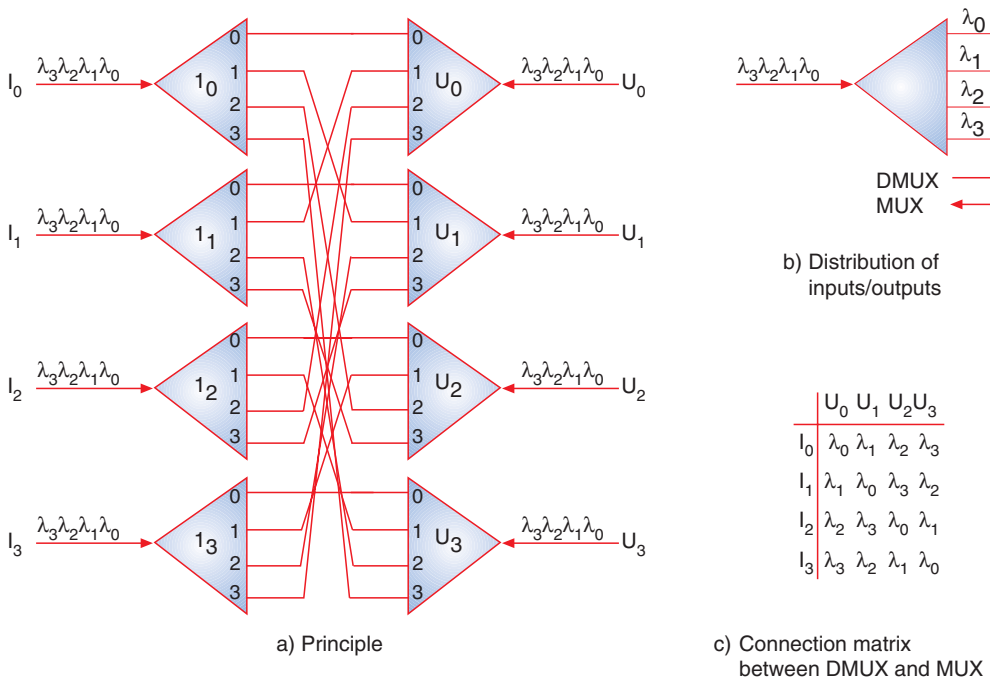


Figure 11.19 Example of passive wavelength routing

11.6 Wavelength routing

Due to the large bandwidth of optical components (spacing of wavelengths of only 0.1 nm gives a bandwidth of about 18 GHz), it is not surprising that the idea of using a number of closely located wavelengths as a basis when constructing communication networks has arisen.

There are, essentially, two ways of doing wavelength routing:

- Passive routing
- Active routing.

Further methods include combining these two methods and also combining them with time and space switching.

11.6.1 Passive wavelength routing

By passive wavelength routing it is here meant that the routing through the network relies on fixed wavelength selective components. There is no switching element by which the route can be chosen.

The method is illustrated in Figure 11.19 for a system with 4 inlets and 4 outlets.

The DMUX and MUX are wavelength multiplexers/demultiplexers implemented from optical filters. Input data are modulated on 4 wavelengths as indicated (by $\lambda_3\ \lambda_2\ \lambda_1\ \lambda_0$). The demultiplexers route the data as shown in Figure 11.19b, via the connections to MUX which executes the multiplexing process. The connections between the demultiplexers and multiplexers are shown in matrix form in Figure 11.19c.

It is very important for correct functioning that the outlets from all of the demultiplexers follow the same order and also that inputs to all multiplexers follow the same order.

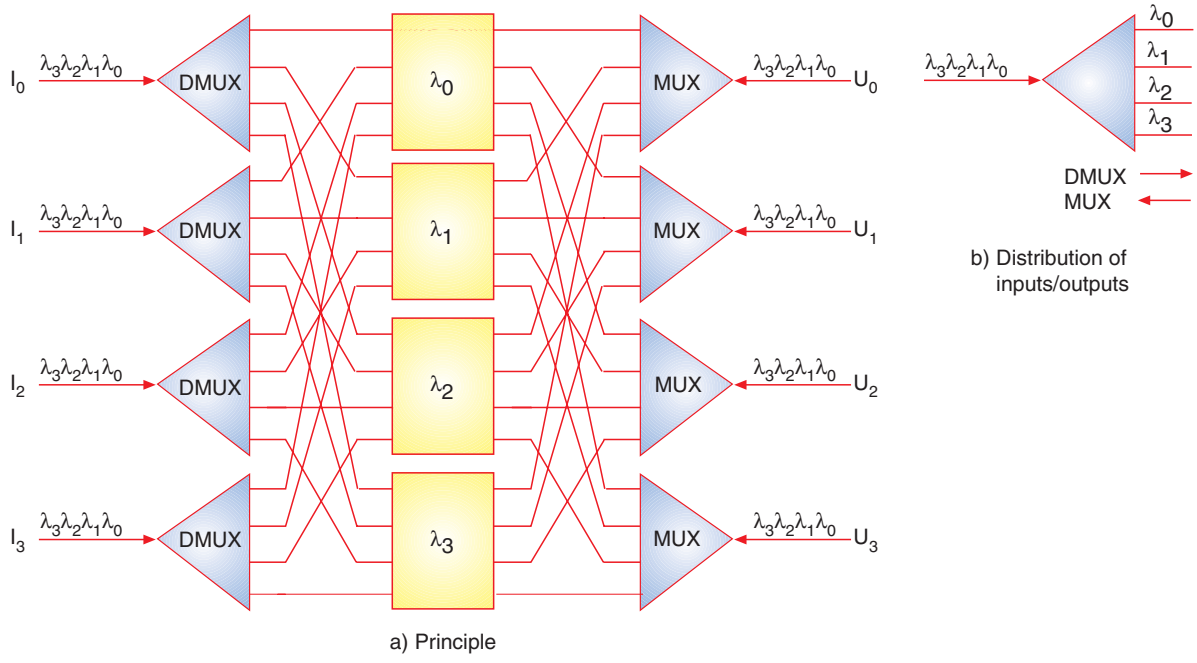


Figure 11.20 Example of active wavelength routing

The outputs must be able to recognise the various wavelengths. Therefore, tuneable (optical) filters have to be used at the outputs.

Number of links is:
$n_l = 4 \times 4 = 16$ (general: $N \times N$;
N = number of inlets).

11.6.2 Active wavelength routing

By using space switches (wavelength-selective or not, or perhaps both) the flexibility and the capacity of the network, measured in number of terminals (inlets/outlets) can be increased. This is illustrated in Figure 11.20. The system consists of wavelength de-/multiplexers as above but has in addition switches, each routing a separate wavelength as indicated. The switches can be wavelength-selective or not.

The configuration gives:

Number of links: $n_L = 2N^2$

Number of switches: $n_{sw} = N$

As above, the outputs must include tuneable filters in order to recognise data. An alternative method is to use a wavelength tuneable switch. Such a switch can be implemented by use of an acousto-optic tuneable filter. This principle applied for a switch is illustrated in Figure 11.21. As indicated, the switch can execute the change-bypass function on individual wavelengths simultaneously. There is, however, a restriction:

- The same wavelength can occur only once on an inlet or an outlet.

An advantage when using wavelength routing is that the maximum total bitrate capacity is BN^2 compared with BN in star coupler configurations (N is the number of inlets/outlets and B is the bitrate of each input).

11.6.3 Network structures based on optical wavelength routing

Many network structures based on wavelength routing are implemented or have been proposed. An example of such structures (the Shuffle Net) is shown in Figure 11.22. It is a bus configuration based on the perfect shuffle principle. In the example 16 wavelengths are used, and, as shown in Figure 11.22b, the logical interconnection pattern resembles the perfect shuffle with one wavelength assigned to each interconnection.

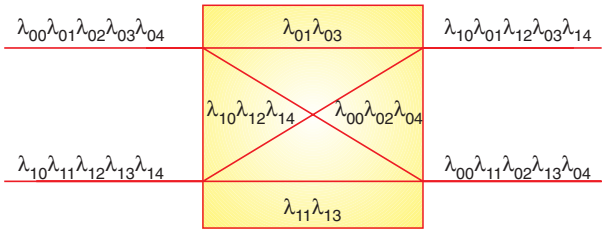


Figure 11.21 Change - bypass wavelength switch

It is seen that the number of wavelengths is $2N$ where N is the number of nodes.

A "shuffle" configuration requiring only two wavelengths independent of the number of nodes is also proposed. The cost is that the interconnection pattern is more complex than the configuration above.

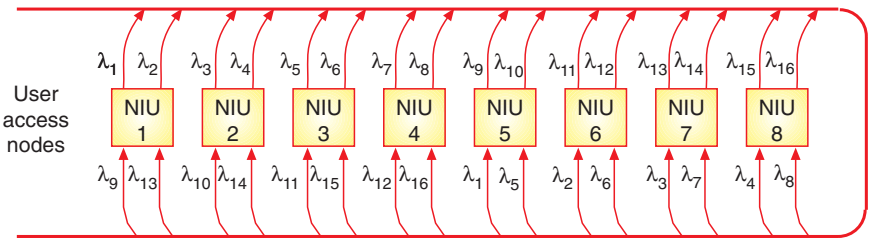
Another proposed and experimental networks are for example, FOX (fast optical cross-connect), HYPASS (hybrid packet switch), STARTRACK, and LAMBDANET.

11.7 Optical wavelength switches – OWSW

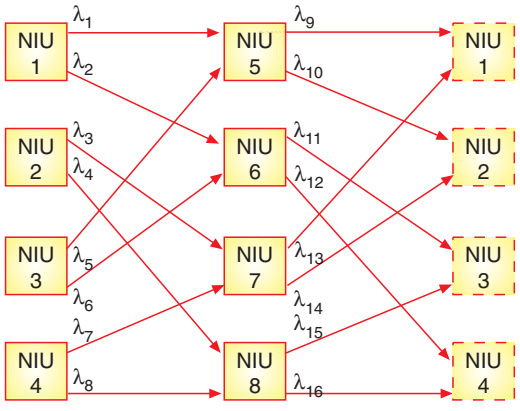
The basic principle of a wavelength switch is illustrated in Figure 11.23. Data d_0 modulated on a wavelength λ_0 comes from the switch modulated on another wavelength λ_1.

Input data modulated on wavelength λ_1 are modulated on wavelength λ_0 at the output of the switch.

There are various methods for implementing a wavelength switch, two of



a) Configuration



b) Logical interconnection in the Shuffle Net

Figure 11.22 The Shuffle Net

Figure 11.23 Principle of wavelength switch

which are illustrated in the Figures 11.24 and 11.25.

In Figure 11.24 incoming data modulated on a wavelength are demodulated and then modulated on the new wavelength from a bank of fixed wavelengths. The new wavelengths are distributed via a photonic switch (OSW).

In Figure 11.25 incoming data are divided and sent to each of the tuneable filters. The filters having access to the modulator with the new wavelength the data are to be modulated onto, are then tuned to this wavelength. The optically controlled modulator can be implemented by using optical non-linear devices.

11.7.1 Network structures based on passive and active optical wavelength routing

An interesting configuration shown in Figure 11.26 uses both passive and active optical wavelength routing to get a multistage switching network. The structure is based on a λ-switch (like the one above) and a passive wavelength routing component with *N* inlets and *n* wavelengths. The total capacity is *Nn*. The λ-switches in the switching stages perform the necessary wavelength translation and the passive routing stage performs the interconnections.

It would also be possible to incorporate optical time switching, passive and active wavelength routing in an optical switching network.

11.8 Photonic switching of ATM data

As pointed out in a preceding chapter, switching of ATM data is based on storage of part of the route (of the connection) in each node. In addition, the contention problem needs more or less storage depending either on the actual switching or the network configuration. At times, this is a drawback when using photonic switching, because of the lack of practical data storage. Storage times vary between the whole conversation time and one or a few cell times.

For shorter time storage, the fibre optic delay line can be used, but when storing data during a whole conversation time, such a solution seems to be very impractical.

The contention problem arises when two cells belonging to different "connections" try to use the same outlet in a switch or in a node. In a switch the contention can be internal or external. In the first instance the problem can be solved by constructing the switch to be strictly non-blocking (Section 7). In this case it means that the number of possible "connections" or transfers of cells is less or equal to the number of cells simultaneously arriving at the input. In the second instance (more than one cell to the same outlet), storage is in principle needed. However, if the switching network is in an end node and is constructed as a non-blocking network, there may not be a need for storage since VPI/ VCI can be used directly for establishing the connections (transfer cells to right destinations).

In transit nodes, however, there will, in general, be a need for storage both for routing data and because of contention.

Using photonic components, it is possible to avoid storage when several wavelengths are used. A proposed solution is shown in Figure 11.27. The data are modulated on wavelength λ_0, VPI/VCI on wavelength λ_1 for the first node, VPI/VCI on wavelength λ_2 for the second node, and so on, up to λ_n for the last node. A filter in each node filters out the header (VPI/VCI) dedicated to the node. The received VPI/VCI is then used directly for setting up the connection (transfer) and the cell is passed to the next node where the same process is repeated but now with the VPI/VCI modulated on wavelength λ_2, and so on.

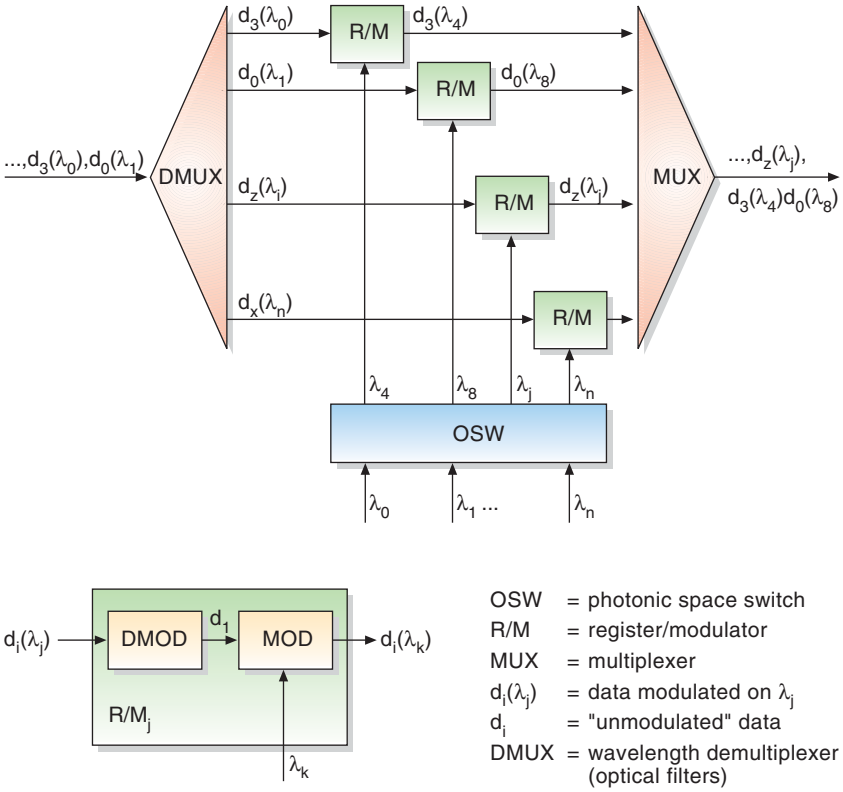


Figure 11.24 Wavelength switch with fixed filters

Bibliography

1 Arrathoon, R, Hassoum, M H. Optical threshold logic elements for digital computation. *Optics Letters,* 10(4), 1984.

2 Ash, E A, Pitt, C W, Wilson, M G. Integrated optical circuits for telecommunications. *Nature,* vol. 261, June 3, 1976.

3 Beesley, J. The foundation of systems IFS-I. *International Switching Symposium,* Boston, 1972.

4 Boissier, A, Glatigny, A, Thepaut, L. Space division optical switching system of medium capacities. *Proceedings SPIE,* 585, 65–70, 1985.

5 Digonnet, M, Shaw, H J. Wavelength multiplexing in single mode fibre couplers. *Applied Optics,* 22(3), 1983.

6 Erickson, J R et al. A 1.7 Gigabit-per-second, time-multiplexed photonic switching experiment. *IEEE Communications Magazine,* 25(5), 56–58, 1987.

7 Gibbs, H M. *Optical bistability : controlling light with light.* New York, Academic Press, 1985.

8 Goodman, J W. Fan-in and fan-out with optical interconnections. *Optica Acta,* 32(12), 1489–1496, 1985.

9 Goodwin, M, Stewart, C. Proton-exchanged optical waveguides in Y-cut lithium niobate. *Electronics Letters,* 19(6), 1983.

10 Haus, H A, Whitaker, N A Jr. Elimination of cross talk in optical directional couplers. *Applied Physics Letter,* 46(1), 1985.

11 Hinton, H S. Photonic switching using directional couplers. *IEEE Communications Magazine,* 25(5), 16–26, 1985.

12 Hinton, H S. Photonic switching technology applications. *AT&T Technical Journal,* 66(3), 1987.

13 Hinton, H S. Architectural considerations for photonic switching networks. *IEEE Journal on Selected Areas in Communication,* 6(7), 1988.

14 Hluchyj, M G, Karol, M J. Shufflenet : an application of generalized perfect shuffles to multihop lightwave networks. *INFOCOM '88,* 379–390, 1988.

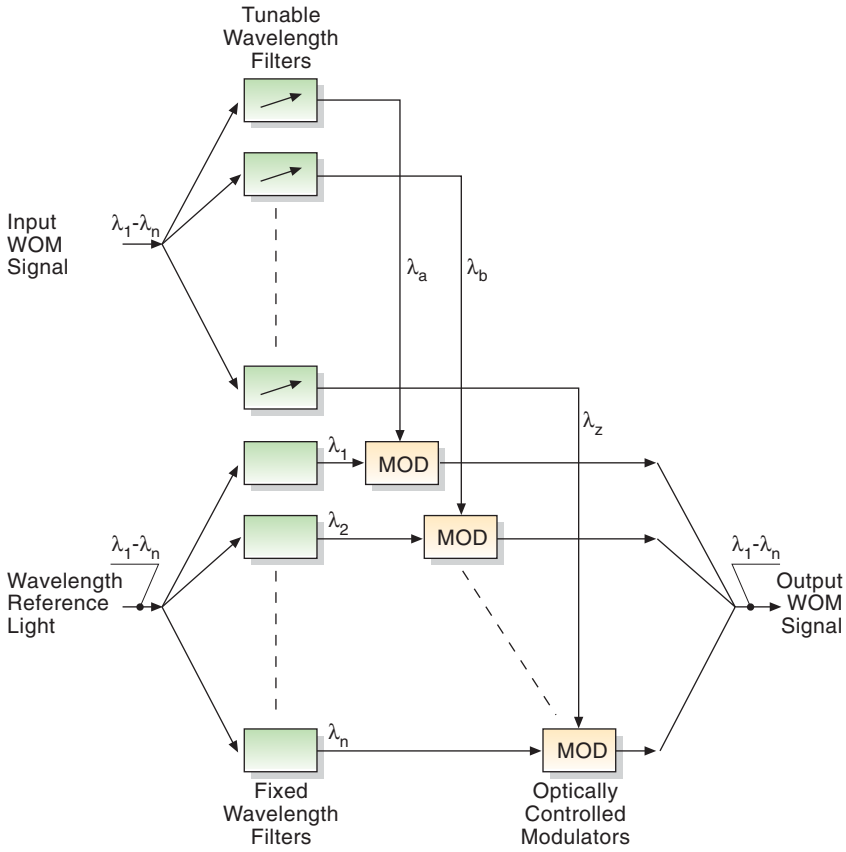


Figure 11.25 Wavelength switch with tuneable filters

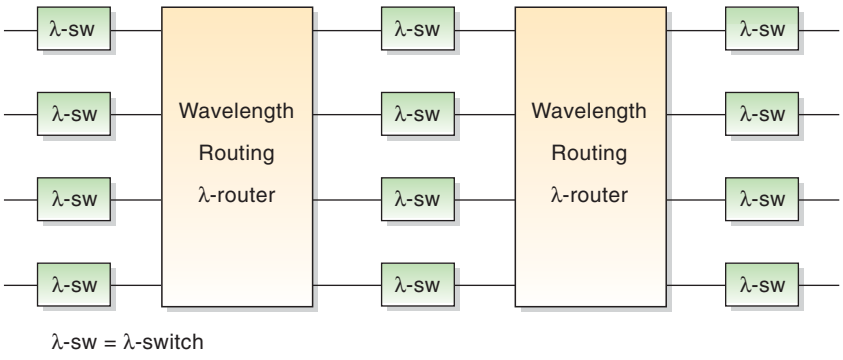


Figure 11.26 Structure with passive and active opitcal wavelength routing

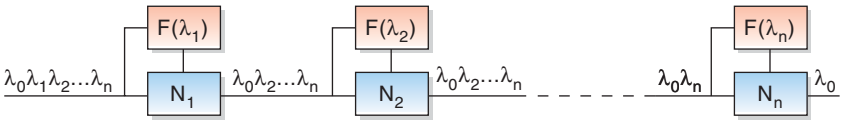


Figure 11.27 Principle to avoid storage of routing data in switching of ATM data

15 Huang, A, Knauer, S. Starlite : a wideband digital switch. *Globecom '84,* 121–125, 1984.

16 Hunsperger. *Integrated optics : theory and technology.* Springer, 1982.

17 Jahns, J, Murdocca, M J. Crossover networks and their optical implementation. *Applied Optics,* 27, 3155, 1988.

18 Kahn, A H, Nejib, U R. Optical logic gates employing liquid crystal optical switches. *Applied Optics,* 26, 270, 1987.

19 Kogelnik, K. *Review of integrated optics.* Holmdel, New Jersey, Bell Telephone Laboratories.

20 Lea, C. A directional-coupler-based broadcast switching network. *Globecom '87,* 3, 2044–2048, 1987.

21 Lentine, A L et al. The symmetric self electrooptic effect device. *Conference on Lasers and Electrooptics, Technical Digest Series.* Washington, DC, Optical Society of America, 249, 1987.

22 Mandel, P, Smith, S D, Wherrett, B S (eds.). *From optical bistability towards optical computing : the EJOB project.* The Netherlands, North Holland, 1987.

23 Marcatili, E A J. Dielectric rectangular waveguide and directional coupler for integrated optics. *Bell System Technical Journal,* 2071– 2102, Sep. 1969.

24 Midwinter, J E, Smith, P W (eds.). Photonic switching. *IEEE Journal on Selected Areas in Communications,* SAC-6(7), 1988.

25 Miller, D A B et al. The quantum well self-electrooptic effect device : optoelectronic bistability and oscillation and self-linearized modulation. *IEEE Journal on Quantum Electronics,* QE-21, 1462, 1985.

26 Miller, S E. Integrated optics : an introduction. *Bell System Technical Journal,* 2059–2069, Sep. 1969.

27 Mortimore, D B. Wavelength-flattened 8 x 8 single-mode star coupler. *Electronics Letter,* 22(22), 1986.

28 Murdocca, M J et al. Optical design of programmable logic arrays. *Applied Optics,* 27, 1651, 1988.

29 Murdocca, M. *A digital design methodology for optical computing.* The MIT Press, 1990.

30 Nordbryn, A. *Optisk svitsjing-forprosjekt.* Sentralinstitutt for industriell forskning, oktober 1984. (Rapport 83 11 05-1.) (In Norwegian.)

31 Ogawa, Y, Ito, H, Inaba, H. Bistable optical device using a light emitting diode. *Applied Optics,* 21(11), 1982.

32 Okumura, K et al. Optical logic inverter and elements using laser or light-emitting diodes and photo detectors in bistable system. *Optics Letters,* 9(11), 1984.

33 Padmanabhan, K, Netravali, A N. Dilated networks for photonic switching. *IEEE Transactions on Communication,* COM-35(12), December 1987.

34 Prise, M E et al. Module for optical logic circuits using symmetric self-electrooptic effect devices. *Applied Optics,* 29, 2164, 1990.

35 *Proceedings IEEE 72,* 758–974, 1984. (Special issue on Optical Computing.)

36 *Proceedings of the Royal Society's Topical Meeting on Optical Bistability, Dynamic Nonlinearities and Photonic Logic, Phil. Trans. Roy. Soc. Lond.,* A313, 187, 1984.

37 Prucnal, P, Blumenthal, D, Perrier, P. Photonic switch with optically self-routed bit switching. *IEEE Communications Magazine,* 25(5), 50–55, 1987.

38 Sawano, T et al. Photonic switching system for broadband services. *Globecom '87,* 3, 51.5.1–51.5.5, 1987.

39 Schmidt, R V, Alferness, R C. Directional coupler switches, modulators, and filters using alternating db techniques. *IEEE Transactions on Circuits and Systems,* CAS-26(12), 1099–1108, 1987.

40 Shimazu, Y, Takada, T. High speed time switch using GaAs LSI technology. *IEEE Journal of Selected Areas in Communications,* SAC-4(1), 32–38, 1986.

41 Smith, P W. Hybrid bistable optical devices. *Optical engineering,* 19(4), 1980.

42 Smith, P W, Tomlinson, W J. Bistable optical devices promise subpicosecond switching. *IEEE Spectrum,* June 1981.

43 Smith, S D. Optical bistability, photonic logic and optical computation. *Applied Optics,* vol. 25, 1550–1564, May 15, 1986.

44 Spanke, R A. Architectures for large nonblocking optical space switches. *IEEE Journal of Quantum Electronics,* QE-22(6), 1986.

45 Spanke, R A, Benes, V E. N-stage planar optical permutation network. *Applied Optics,* 26(7), 1987.

46 Tamir, T. *Integrated Optics.* Berlin, Springer, 1975.

47 Thompson, R A, Giordano, P P. An experimental photonic time-slot interchanger using optical fibers as reentrant delay-line memories. *IEEE Journal on Lightwave Technology,* LT-5(1), 154–162, 1987.

48 Prucnal, P R, Perrier, P A, Chbat, M W. Demonstration of an optical self-clocked time-division switch. *Fiber and Integrated Optics,* 8, 281–291, 1989.

On optical switching — a proposal for an all optical ATM switch

BY STEINAR ANDRESEN

This article starts with a few historical remarks to the evolution of switching technology. It is shown that switching in space and time can be accomplished on a systematic basis using fixed delay lines. In the optical domain such delay lines can be established by using some metres (or kilometres) of high quality optical fibre.

The author also shows that the principle may be extended to cover wavelength (or frequency switching).

The first half of the article deals with Benes networks; they are symmetrical multistage link networks capable of implementing an arbitrary permutation of input signals on the output (demands proper setting of network parameters).

In the second half, the author examines the possibility of building a simple switch for ATM based signals. A novel scheme proposing utilisation of both Hardamad coding and wavelength modulation is proposed. For the solution chosen all input information from a given cell period at the input side is accessible in *r* consecutive cell periods on the output side. It is still not proven that scheme can be implemented with the electro-optic or all optical circuitry given today. However, as the proposed scheme is simple and elegant and can be scaled to cater for huge systems it might be worth while to launch a feasibility project.

Introduction

In telecommunication there is always a demand for optimizing the use of media. This has given birth to a number of multiplexing techniques.

For transmission purposes FDM was the first choice for analogue systems. However, switching based on FDM was impractical with the type of technology at hand at that time. Thus, switching was carried out in the space domain only. This meant that the signals had to be demultiplexed before switching, and then (re)multiplexed to be carried along on a new transmission link. Integrated schemes for transmission and switching arose with the advent of digital electronics, which made it economical to utilize Time Division Multiplexing (TDM), thereby effectively "collapsing" the physical size of a switching network. One of the reasons for the economy obtained on TDM switching was the cheap and abundant supply of RAM VLSI chips which makes temporary storage of the signals an easy task.

In the future, transmission schemes based on the utilization of optical fibres seem most promising with respect to cost and bandwidth. In order to obtain good economy even here, the signals should preferably be switched without the need for translation to another modulation mode.

The situation within the optical domain so far differs from the electronic domain in two aspects:

1 There is no feasible "RAM" technology that makes temporary storage and random retrieval of optical signals easy.

2 FDM based switching seems feasible, i.e. routing based on selective wavelength filters and/or tuneable signal receivers.

The author of this article was engaged in the early 1970s in a project where different proposals for PCM switching were studied. This was before the time of cheap and fast RAM technology, so temporary storage of digital signals had to be implemented by delay lines. In our study we used a SAW (Surface Acoustic Waves) as storage device. Of interest to this article, however, is the scheme developed for switching, (1) and (2): Instead of time switches we used a set of delay lines in combination with regular space switches. This principle can also be used with optical switching. For delay values of interest, strands of optical fibres can probably be substituted for the SAW lines, thus making TDM switching structures easily feasible.

Benes networks (symmetrical multistage networks)

An example of such a network is shown in Figure 1. It depicts a rearrangeable Benes network for 16 channels total (2 systems of 8 time-slots each). It is implemented by 2 x 2 matrices or binary matrices. This scheme can be extended to cover larger number of stages and time-slots as shown in Figure 2.

An overview of different schemes is given in (1). Theoretically, the same principle can be used with any multiplexing method where the modulation process is capable of working "modulo *n*" (*n* is some integer number). A similar example using space and wavelength multiplexing is shown in Figure 3.

The switch consists of 3 stages of 4 x 4 matrices. Every signal path carries 4 different wavelengths, the switching matrices can switch signals in each wavelength plane differently. The upper scheme shows the entire system "folded out". The drawing underneath shows one of the wavelength planes. In order to work, wavelength conversion has to be accomplished in the link stage.

In the figure the translators are supposed to be two-way translators (to and from the other wavelength). Analogous to the time displacement schemes given in Figures 1 and 2 it is also possible to design systems that utilize wavelength displacement (modulo *n*). The wavelength displacement scheme can be modelled as a broadband common frequency translation, with the peculiarity that it should work modulo some *n*. Consider a band *B* ranging from f_1 to f_n (i.e. $B = f_n - f_1$). Let this band be divided into *n* carrier frequencies with distance Δ_f. If all signals are "lifted" one step, i.e. signal with carrier f_1 is transferred to the carrier f_2 ($f_2 = f_1 + \Delta_f$) etc., the uppermost frequency will become lifted out of range by the

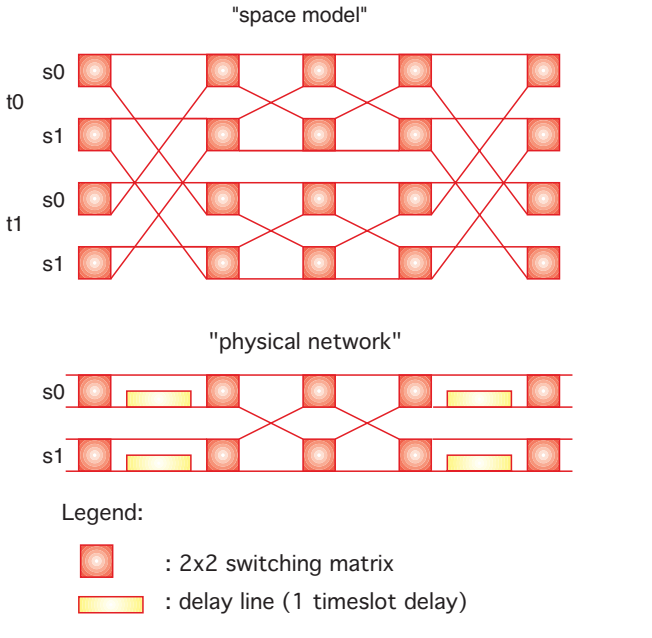


Figure 1 A combined space and time multiplexed network

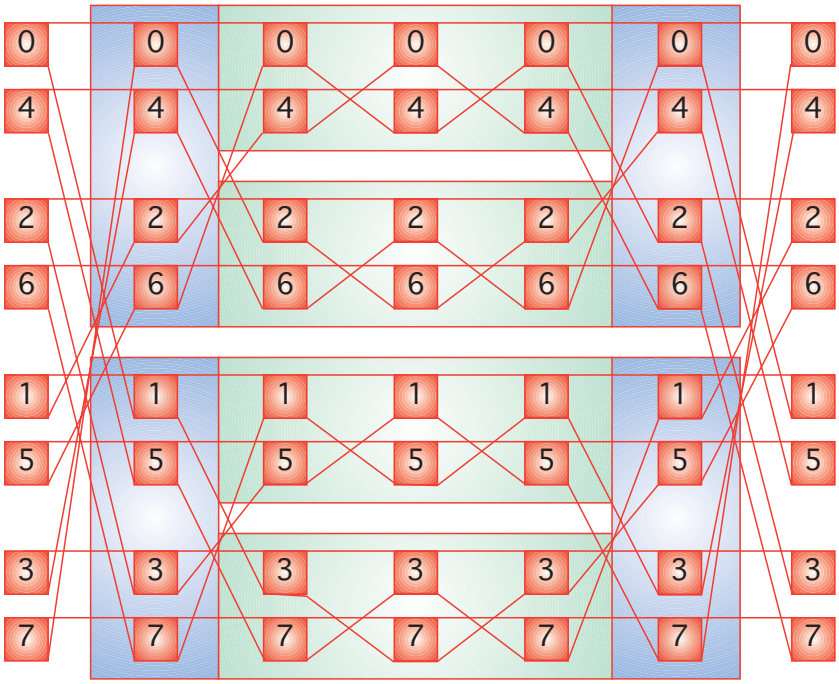


Figure 2 Space and time 7 stage switched network (2 systems with 8 time-slots)

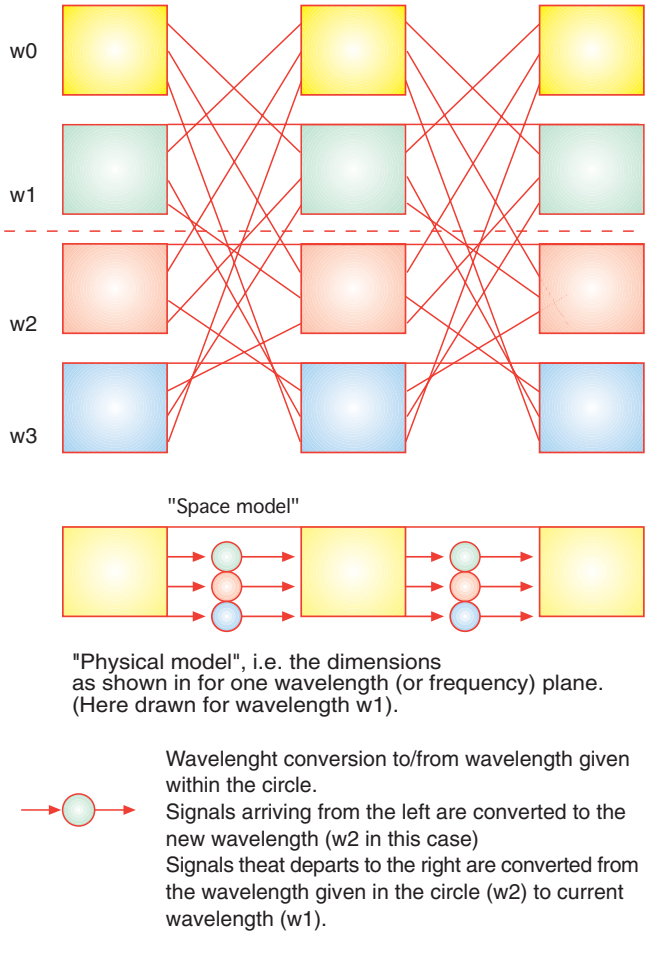


Figure 3 Switching in a "space" and "wavelength" system

process and has to be moved down $n \cdot \Delta_f$ Hz. A possible implementation scheme could then be as follows:

1 Shifting the signals up the required number of bands: Broadband remodulation

2 Band filtering: Diverting signals that are out-of-band

3 Remodulation (and filtering) of out-of band signals to shift them down $n \cdot \Delta_f$ Hz. (This will get them back into permissible range.)

What is actually needed is a broadband (or multiband) frequency shifter (for step 1 and 3) and a band filter (step 2). Such a system is shown in Figure 4.

Whether this is feasible is not known. However, to obtain a shared use of the switches we also have to invent some kind of switching matrix that implements connection patterns specific for each wavelength plane. Whereas it is not unfeasible to design signal processing schemes that would allow this, the solutions do not seem economical! The reader is invited to play with the possibilities to see what can be established. (As a first exercise: try to construct signal processing scheme for a link-stage that takes e.g. 4 signals given on separate terminals but with the same wavelength, and maps them onto a single terminal but with 4 distinct wavelengths in the next switching stage – this corresponds to a "space/ frequency twist" and can easily be derived.)

A simple version of the STW scheme with just two frequencies (which may more easily be constructed) is shown in Figure 5.

Within each carrier a TDM scheme with two time-slots is used. The switch is rearrangeable and capable of realizing an arbitrary permutation of the set of 16 input channels onto the output.

The description above assumes switching within a periodic frame period. In a regular PCM system this period is 125 µsec. With a propagation speed of roughly $2 \cdot 10^8$ m/s within the fibre, this corresponds to approximately 25 km.

In order to be able to switch between any time-slot we would need to establish signal paths that implement delays at least 125 µsec minus one time-slot. This can be done in many ways, by using a scheme similar to Figure 3. The longest single line would need to be 12.5 km, which may be impractical (but not impossible).

Switching of ATM signals

In the rest of this paper we leave the principle of Benes networks and will instead concentrate on a scheme for doing ATM switching. The transmission systems carrying ATM signals may have an overall frame period defined, but the information from a single connection is not assigned to a fixed time-slot within this period. Instead, at the receiving side we must sort the information according to the virtual channel number which is sent along with each "frame" or cell of information. The cell length is standardized to 53 bytes (8 bits words). Building a switching node for a set of such streams would necessitate the means for:

1 Identifying the virtual channel numbers in the incoming cells

2 Do a look-up in a connection register to decide which physical line the information is meant for and also read the virtual channel assigned to this connection on the next link

3 Substitute new virtual channel information for the old one and route the information to the assigned line.

Because no prior fixed assignment to "time-slots" are made, information from two or more different sources may contend for the same outgoing slot on a given output. The resulting congestion can be relieved if we can delay the incoming signals with an arbitrary number of "cell" periods.

In order to enhance the capacity of the fibres it is customary to let them carry a number of wavelength multiplexed systems. We must then also be prepared to switch the signal onto a new wavelength. At the outset we will drop this assumption and will instead utilize wavelength tuning just as a means to obtain "space switching" between different optical fibres. This principle is documented by many authors, a good introduction is given in (3).

For the rest of this paper it is assumed that every input fibre carries a single wavelength signal. ATM signals are modulated onto this carrier. For simplicity it is assumed that the cell rate is the same on all systems. The time between two consecutive cell starts is called a cell period. The actual modulation frequency may be chosen so that a guard interval is obtained between each cell. (For example, the bitrate can be set to a value giving 54 bytes in each cell period. This would give a guard interval of one byte.) Such a guard interval can be practical for switching purposes and synchronization.

It is also assumed that all components work in linear mode, that is signals from several sources may be superimposed without causing signal saturation.

Our suggestion is to use a set of delay lines just to buffer the signals, i.e. to make them accessible in more than one "time-slot". This gives greater freedom regarding fibre access and can be realized very economically.

Consider e.g. 6 incoming fibres each carrying a signal that is modulated with a rate of 2 Gb/s. If we divide this bit stream into ATM cells with a 1 byte guard interval between the cells, we will have a cell rate of about 4.6296 Mcells/s. By use of wavelength switching, e.g. fixed sender frequency and tuneable receiver frequencies, it is possible to switch ATM cells from one of the incoming fibres onto an arbitrary outgoing one. A bank of x delay lines ensures that all incoming signals are accessible in x time-slots. Such a scheme is illustrated in Figure 6.

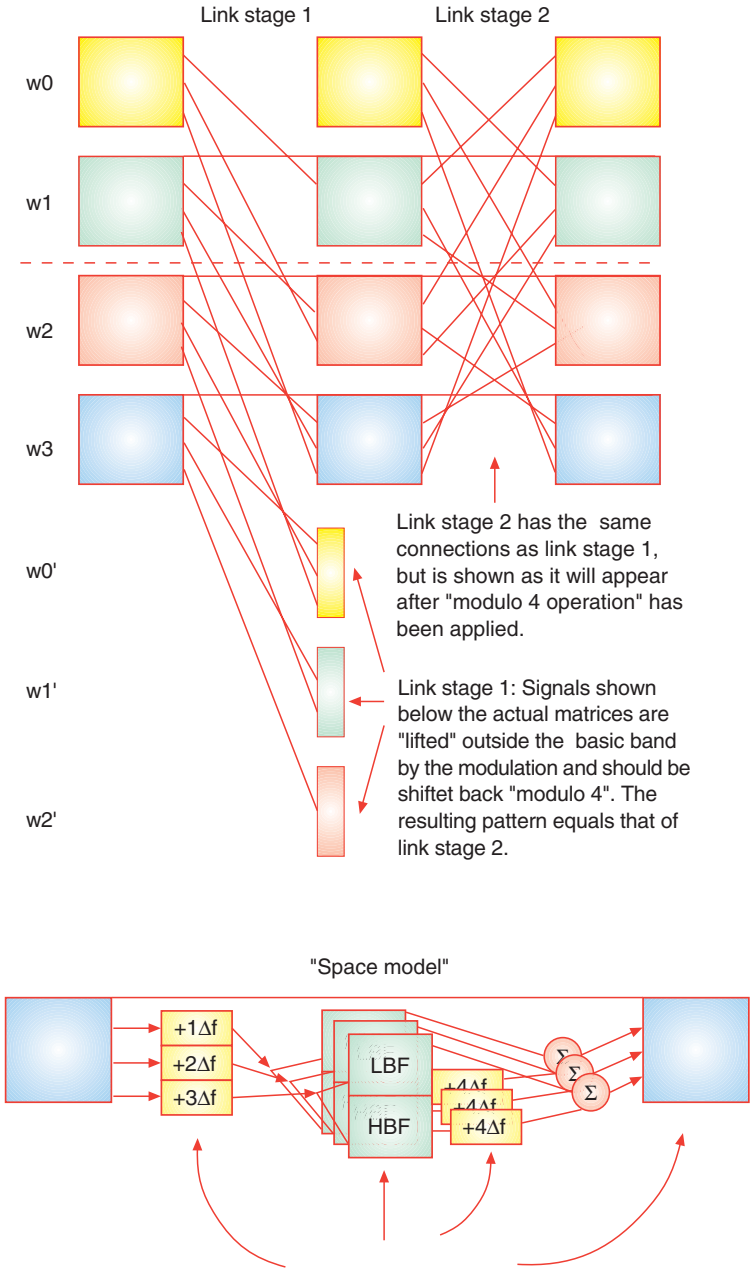


Figure 4 Wavelength shifting "modulo n"

Here x is set at 8. Six incoming systems are given different wavelengths by the modulators of the input stage. The signals are sent through a passive star coupler into 8 different delay lines. At the output of each delay line all the incoming signals will be available with a delay of 0, 1, or 7 time units, respectively. The time unit should correspond to the time between the starts of two consecutive ATM cells on a single system. (This parameter is considered a system wide standard.)

At the output side, the output has to pick the delay line for its source signal and then tune into the wavelength corresponding to the *input* line. The first of these actions demands a spatial switch.

Please note that the element as shown above can be structured in two ways:

OT (Output-tune)
(As given) by using tuneable receivers.

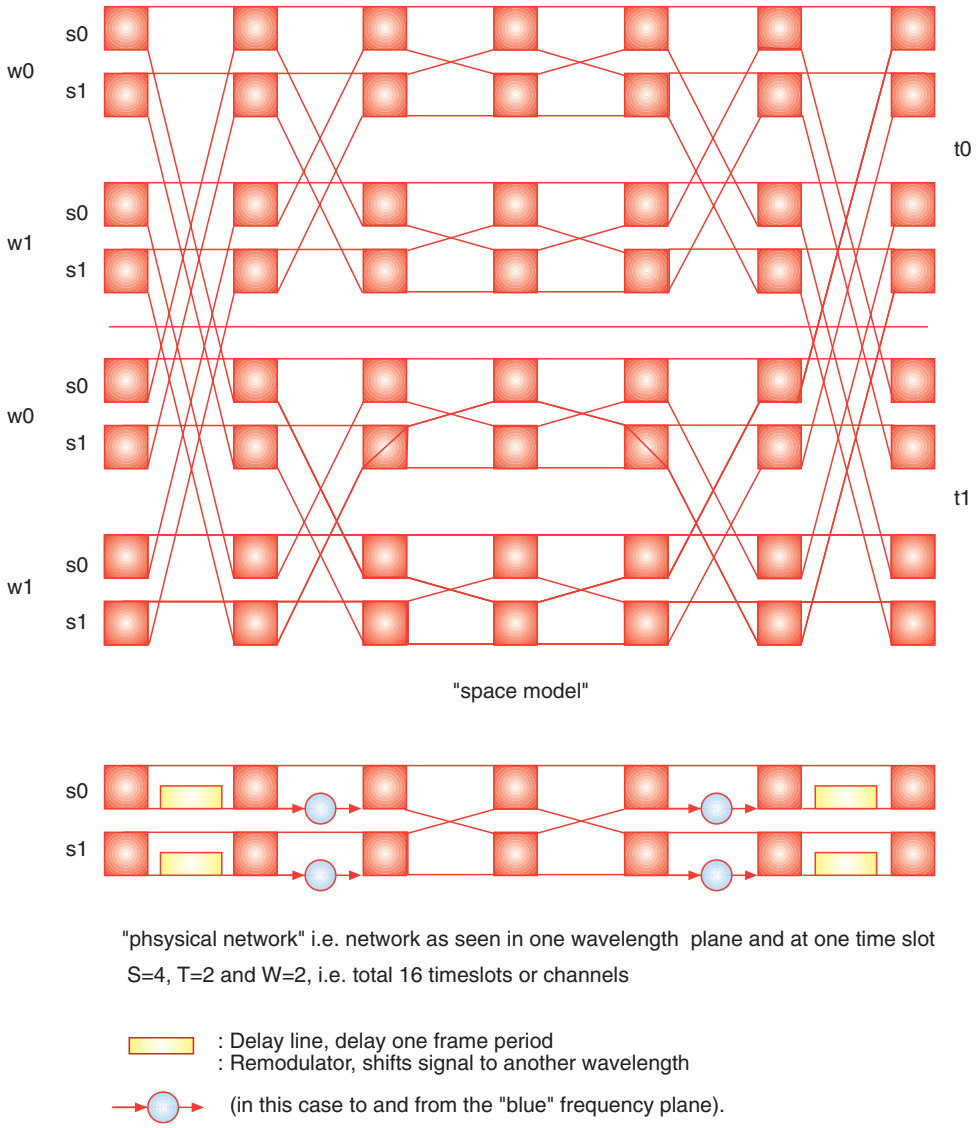


Figure 5 Switch for two physical systems each one carrying signals on two different wavelengths (carrier frequencies)

IT (Input-tune)
By using tuneable senders. Using the same figure this can be illustrated by having the input on the right hand side and output on the left (reversing the direction of information flow).

IT may give rise to the simplest implementation, but OT is the most general one because it allows broadcasting (or multicasting) and also gives more control over priorities as this can be decided upon at the output. In scheme IT, the output port is assigned at the input side and cannot be altered. (It is not possible to "overwrite" a cell launched earlier.)

The spatial switch can be avoided if we introduce a second stage of wavelength modulators/demodulators as shown in Figure 7.

At the output of the delay lines a single set of signals are demodulated per time-slot and then remodulated to a frequency corresponding to the output filter. This solution is chosen in this case to avoid using more than 6 wavelengths. In doing this we restrict the possibility of broadcasting. (An input cell cannot be copied to more than one output *simultaneously*. If an input cell should become copied to more than one output the copies must be retrieved from different delay lines. That is, the copies will experience different delays.)

The remaining problem consists mainly of devising a routing protocol that can be implemented efficiently.

Routing schemes

In this explanation we do not treat the establishment of a connection, but explain the switching process of existing connections. If a set of virtual connections is established, the control logic must read the virtual channel field of every incoming packet and then (after making a look-up in a local connection register) substitute a new virtual channel number for the old one and route (switch) the cell to the right output.

Input tuning

A simple scheme of type IT can be modelled as follows:

With a type input-tune scheme we assign the path for the output signal at once, a single copy of it will travel through the network. Hence no broadcasting/multicasting function is available. Also it is not possible to let a "late-comer" with higher priority take the slot assigned for the cell. ("What is done is done".) There is no need to communicate control information across the switch. However, at the input side we must assure that we do not route two (or more) separate signals to appear at the same output simultaneously. This means we must establish a "global" knowledge of decisions taken. It might prove difficult to obtain this coordination on a simultaneous basis.

Output tuning

The following is known (as seen from a specific output):

The last r cell requests from each input (with priorities) are known, some of these are probably already served. Within this, the set of unserved requests for the output given is of interest to us. Let us call this set R. A host of different strategies can be worked out for this scheme. Two approaches are given here:

> To decide once for every time-slot which request to service just now,

or

> To try to establish (and update for every time-slot elapsed), a pre-assigned schedule for the requests. In reality, this means to work continuously on a "flight plan" for the individual cells (requests).

It may be possible that the latter scheme may yield better results, but for the sake of simplicity the first approach is taken here.

Several versions of traffic schemes can be envisaged, e.g.

1 Systems with only one common class of traffic (one priority)

2 Systems with two traffic classes:

A Time critical traffic, cells that can be delayed up to *x* cell periods per node. A cell is thrown away if this cannot be fulfilled.

B Reliable non-time-critical traffic. May be retransmitted from "somewhere" if lost. There is time-out counters working within a handshake scheme at the upper OSI layers.

We will restrict ourselves to case 1 (systems with just one common class of traffic).

The dimensioning aspects are left to others as in this paper we just want to illustrate some possible implementation schemes.

A sketch of such a system is given below.

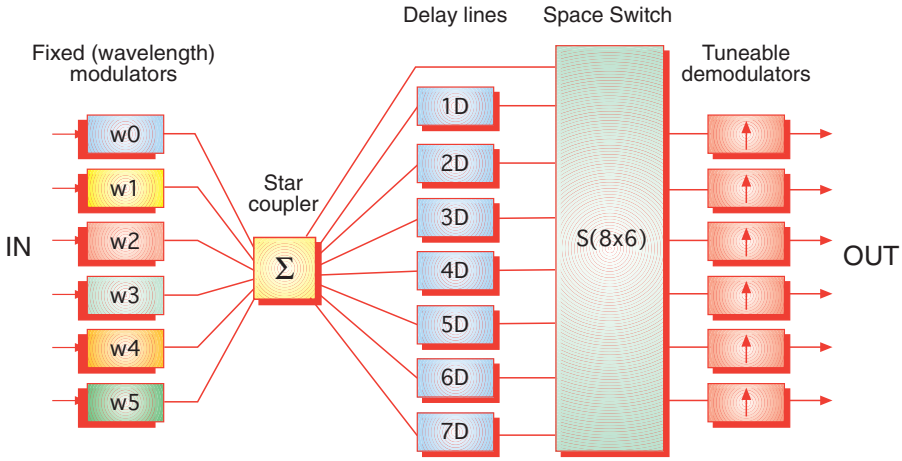


Figure 6 An ATM switching node

ATM switch using orthogonal modulation (wavelength modulation and Hardamad coding)

The switch proposed here has essentially the same structure as switches shown in Figures 6 and 7. But Hardamad coding (4) is used as a substitute for the space switch on the output stage (ref. Figure 6). The output of every delay line is given a distinct Hardamad code; that is, the final choice concerning output access is done at the output (output-tuning). Each output has access to the *r* last cells of all incoming systems (*r* is the number of delay lines).

The resulting system structure is illustrated in Figure 8. The working principle is detailed below. It must be emphasized that the system structure is based on speculative assumptions about what would be a feasible technology within a few years! If the principles prove feasible, parameter values can easily be adjusted to high capacity nodes.

A short stepwise explanation of the working principles are given below:

1 At the input side of the switch:

1.0 "Skew" the information, so that at least a number of different lines

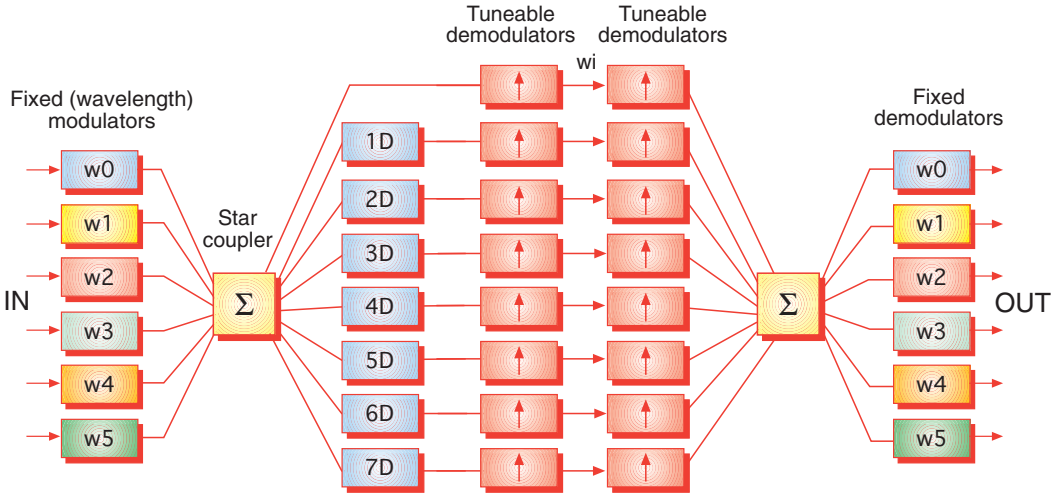


Figure 7 Same switch as in Figure 6, but with wavelength modulation used to accomplish the "space switching"

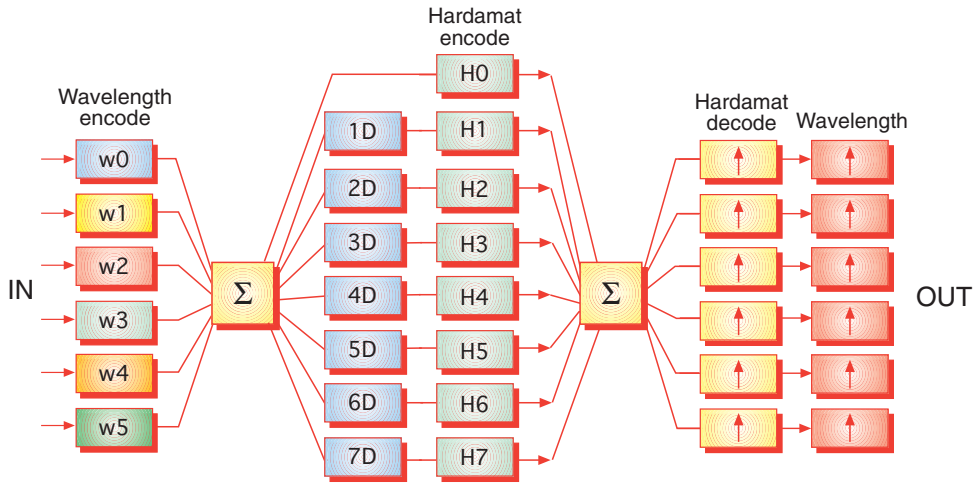


Figure 8 ATM switching realized with orthogonal modulation (simplified scheme)

may be treated in sequence by common circuitry. (This can be done by introducing short delay lines, one for each system.) The delay for each system is different, the delay is incremented in small steps – fractions of a cell period. The resulting "skewness" can be corrected by another set of delay lines with values complimentary to the one in the first stage.

1.1 Read the incoming header (do not remove it). The cell is "stored" by propagating through a fibre. All inputs have such buffers (in front of the wavelength (re)modulator). This buffer is introduced to assure sufficient processing time for the next operation, see below.

1.2 Use the virtual channel number to determine physical output line and new virtual channel number (look-up in a connection register).

1.3 At the output of the input buffer: synchronize onto the arriving cell and produce a new one on the specific wavelength of this input. The new virtual channel number (derived above) should be substituted for the old one during this process.

2 Within the switch:

2.1 The information from all inputs are sent to a star coupler (first all the signals are fed onto a common media). Then the composite signal is branched off to different delay lines. The delay lines in principle delay the incoming composite signal 0, 1, 2 and r "cell periods".

Delay	Input							
	0	1	2	3	4	5	6	7
0	0	1	0	1	0	0	1	0
1	0	0	1	0	0	0	0	1
2	1	1	0	0	0	0	0	0
3	0	0	0	0	0	0	0	1
4	0	1	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0

Figure 9 Request matrix (for a given output)

2.2 The composite signal that is fed into each delay line becomes coded by a Hardamad coding process. A unique Hardamad code is used for each delay line. The Hardamad coding normally relies on successive inversions of the amplitude of the signal. This may be complicated to accomplish for a composite signal consisting of many wavelengths. A possible solution could be to find some way to implement 2-PSK modulation (common to all sub-carriers). A phase switch from 0° to 180° for all sub carriers would correspond to an amplitude inversion. Please note that as the wave period length of light is very short compared to the bit period (1 to 10^5), potentially a very larger number of Hardamad codes could be used. In order to distinguish between, say, 70 to 100 different delay values we have to use codes with a length of at least 128 symbols. This means that each bit becomes divided into 128 equal length periods where it is either transmitted with phase 0° (original) or 180° (inverted). The effect of imperfect modulation in transition phases (when shifting from 0° to 180° or vice versa) can possibly be avoided by sampling the output signals in the mid period (of the Hardamad symbols).

The modulation could be done either before or after the delay lines.

3 At the output side:

3.1 At the output side. All the wavelength- and Hardamad-coded signals are sent through a new star coupler and then to the different outputs. Each output will thus receive all input information r times. The remaining problem is essentially to decide which information to choose for the given output each time, and then to retrieve it by a two step demodulation process: Tune to the right wavelength and then the right code. The necessary parameters must be obtained from the routing logic.

3.2 Routing logic. Each cell has its physical outlet(s) determined by the look-up process at the input side. This data can be used to establish a "request" matrix for each output that points to the cells to choose from for next transmission. The matrix will show cells delayed 0, 1, 2 ..., r cell periods. The output logic must decide which one to transmit and must keep track of the ones served. (In this article we do not detail possible physical implementation schemes for the control signals, but restrict ourselves to give a sketch of the logical scheme.) The request matrix for certain outputs can be visualized as having r words each with n bits as shown below, with $r = 12$ and $n = 8$ (Figure 9).

A "1" in the word bit position j corresponds to a request for transmission from input number j. When this request is met the corresponding bit is reset (to "0"). The bit pattern is shifted one position downwards per "cell period". The initial setting of he request word (for "delay = 0") has to be supplied from the input side. As only one request can be serviced per time slot, cell losses will occur whenever more than r (12 in this case) "1"-s can be found in the matrix. Current value (shown above) is 10, so there is still a possibility to avoid loss (if not more than 2 new requests are received during the next 12 periods). The matrix can be "refined" with priority markings.

A strategy of which request to serve next must also be established. A fair and straightforward version for a system with a common priority could be:

Find the oldest request(s). If more than one – pick one of them at random.

3.3 Demodulation. When the request for service is chosen, the wavelength (corresponding to the input) and the Hardamad code (corresponding to the delay line) is known. The signal information must be retrieved by a two step modulations process. It would probably be safest to demodulate according to the Hardamad code first. (If the two schemes are sufficiently independent it would not matter which demodulation was done first.) The switch to new parameters must be done within the "guard interval" in-between cell periods.

Conclusion

If a "broadband" two phase modulator/demodulator can be constructed for optical signals, the scheme is feasible. What is meant by this requirement is that the modulator should work with no special tuning to a given carrier (within a certain band). If this can be accomplished the two multiplexing systems are orthogonal, i.e. independent of each other. This should give the engineer great freedom in implementing cost efficient solutions.

References

1 Andresen, S, Harrison, S R. Toward a general class of time-division-multiplexed connection networks. In: *IEEE Transactions on communications,* vol COM-20, (5), 1972, 836–846.

2 Andresen, S. *Clos permutation networks : a study of combinatorial and structural properties.* Trondheim, ELAB, 1976. (ELAB report STF 44 A 76146.)

3 Ghafoor, A et al. Architecture of an all-optical circuit-switched multistage interconnection network. In: *IEEE journal on selected areas in communications,* vol 8, (8), 1990, 1595–1606.

4 Lindsey, W C, Simon, M K. *Telecommunication systems engineering.* Engelwood Cliffs, New Jersey, Prentice Hall, 1973.

Addendum

Some comments on an (ATM) implementation

An ATM cell with 53 bytes gives 424 bit periods. With a transmission rate of e.g. 2 Gb/s we get 0.5 ns per bit. To cater for synchronization and time delay variations it is proposed to use a "cell period" on the transmission side and to choose a cell period that is 54 · 8 bit periods long, i.e. 432 bit periods. Net frame rate per fibre (given a modulation rate of 2 Gb/s) will then become 4.6296 Mframes/s.

The propagation speed in an optical monomode fibre is given by $3 \cdot 10^8$ m/s divided by 1.45 (the diffraction index). This gives a speed of $2.07 \cdot 10^8$ m/s (about 200,000 km/second).

A delay of one ATM frame length + guard area (totalling 432 bit periods) corresponds to a fibre length of

$2.07 \cdot 10^8 \cdot 432 \cdot 0.5 \cdot 10^{-9} = 2.07 \cdot 43.2 = 44.685$ metres.

Cell period delay:	Fibre length:
1	44.685 metres
10	446.85 metres
100	468.5 metres

The wavelength of light is typically 0.5 micrometre ($0.5 \cdot 10^{-6}$ metres). This corresponds to a wavelength period of $0.25 \cdot 10^{-14}$ s (with light speed put to $2 \cdot 10^8$ m/s). This shows that a bit period corresponds to about $2 \cdot 10^5$ wavelengths of light at a modulation frequency of 2 Gb/s. Hence, a Hardamad scheme seems feasible.

The HARDAMAD codes

The presentation on Hardamad codes given is based on a description found in chapter 5 of (4): "Phase-Coherent Detection with perfect reference signals". The subsequent samples are by this author.

If +1 is represented by 1 and -1 by 0, the first 16 codes can be given as follows: (This corresponds to matrix H4, see below.)

0	0000 0000 0000 0000
1	0101 0101 0101 0101
2	0011 0011 0011 0011
3	0110 0110 0110 0110
4	0000 1111 0000 1111
5	0101 1010 0101 1010
6	0011 1100 0011 1100
7	0110 1001 0110 1001
8	0000 0000 1111 1111
9	0101 0101 1010 1010
10	0011 0011 1100 1100
11	0110 0110 1001 1001
12	0000 1111 1111 0000
13	0101 1010 1010 0101
14	0011 1100 1100 0011
15	0110 1001 1001 1001

(Reed-Muller is another name mentioned in the book on these codes.) The code values may be developed recursively:

Start the matrix H1:
0 0
0 1 = H1

Proceed with H2 which is established by "stacking" two copies of H1 vertically for the first two columns. The last two columns created by using a copy of H1 on the top and a bit inverted version of H1 (called H1) on the bottom.

00 00
01 01 = H1 H1 = H2
00 11 H1 H1
01 10

The same scheme can be used for recursive establishment of H3, H4, H5, etc. The resulting code words are found rowwise (or as columns).

Sample coding/decoding

A small coding sample, "a" and "b", with the codes 0101 (= Hx) and 0011 (= Hy), respectively, is shown below. Decoding is illustrated for "a", see Figure 10.

It is realized that the original signal can be retrieved if the transmission channel corresponds to a linear medium.

The sample shows "direct" modulation (or coding). If the original signal is on a carrier, this carrier could be keyed with 0° (corresponding to +1) and 180° (corresponding to -1) phase shifts to obtain essentially the same effect. (The author presumes this is easier to accomplish.)

Non-linearities in the medium may introduce noise and errors. Symbol rate for coding and decoding have to be exactly the same. But this does not represent a serious problem as both can probably be derived from the same source. Syncronization ("alignment" of code periods) should not impose big problems, referring to well known and short distances.

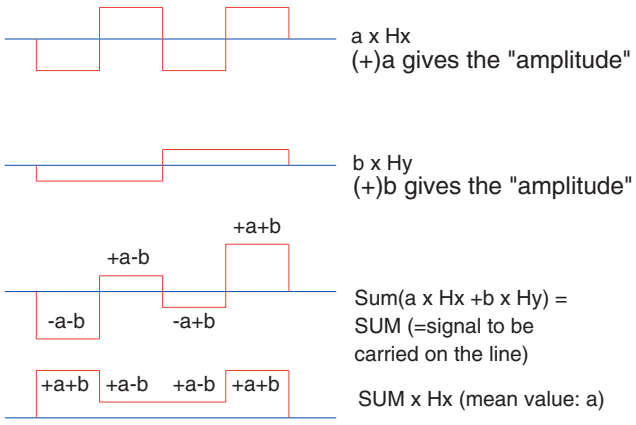


Figure 10 Sample coding and decoding using Hardamad codes

The Alcatel 1000 switching

BY KJETIL STENHAMMER

Introduction

ATM switching is attractive as the future switching principle for two main reasons: the ability to switch traffic from all kinds of services and the potential of gain in statistical multiplexing.

The ability to switch traffic from all kinds of services is made possible by segmenting information infocells, choosing a high bitrate and rely on a low bit-error rate. Hence, very few processor consuming algorithms need to be executed within the switches in order to support a proficient end-to-end high-quality service. In fact, all kinds of traffic, whether it represents data, voice or video, can be handled exactly the same simple way in the network.

Statistical multiplexing of cells allows traffic from different services to be mixed onto the same physical transmission lines without the restrictions of fixed time-slots and frames. This means better utilization of the transmission capacity. The ATM switches themselves are also in a sense like a packet switch, allowing the circuits within the switches to be shared the same way as the transmission lines.

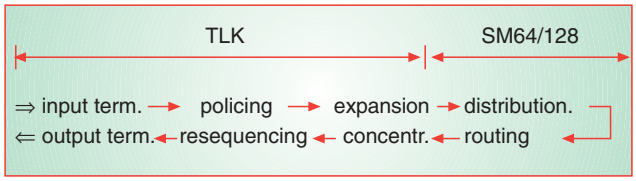


Figure 1 The A1000 MPSR main functions

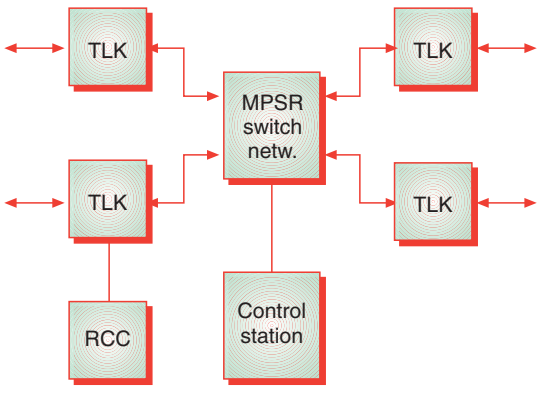


Figure 2 The transport part of A1000

Besides utilizing the ATM benefits, the A1000 MPSR switching concept also implements some further enhancements like integrated interworking with narrowband S-12 and direct SDH connection.

Basics for Alcatel MPSR switching concept

An ATM switch will generally always have to perform the following three functionalities: space switching/routing, i.e. from one input to a specific output, header translation (since the routing information in the header of the ATM cells is valid only for each separate transmission link) and buffering (due to the statistical behaviour of the ATM traffic). In addition, other functions like e.g. various solutions for internal distribution of the traffic are sometimes performed in order to enhance the traffic handling capabilities and increase the security in case of failures in the switch.

The Alcatel MPSR switching concept has the following main functionalities with respect to the transmission part of the switch (buffering and header translation is implicit), see also Figure 1:

- Input link termination
- Policing
- Expansion of the traffic
- Distribution of the traffic (MultiPath)
- Group routing (Self-Routing)
- Concentration of the traffic
- Re-sequencing of cells
- Output link termination.

With reference to Figures 1 and 2, the two types of printed circuit boards for the transmission part of the Alcatel A1000 ATM switch are the TLK (Terminal LinK) interface boards and the MPSR (Multi-Path Self-Routing) switching boards, called SM64 or SM128 depending on the capacity. We can also see from Figure 2 that the "spider" configuration known from S-12 is kept. TLK boards and other boards are connected through the switching network.

The TLK board first of all terminates broadband links, which in this case means SDH and ATM physical layer termination. A conversion between ATM and an internal format is also done and will be described later. Next, expansion and distribution of the traffic from the incoming links will be performed, as the ATM cells are sent into the switching network. Expansion of the traffic means to share the traffic from one external link to two internal links. Distribution of the traffic means to randomly spread the ATM cells from the TLK links to several internal links in order to harmonize the load.

When the cells enter the TLK board back from the switching network (i.e. the SM64/ SM128 switching boards), concentration of the traffic and re-sequencing of traffic cells is performed, and new headers are attached. Concentration of the traffic will in this case mean the opposite of expansion, and re-sequencing of cells is mandatory to avoid the cells coming out of order due to the distribution. Finally, SDH framing is added.

Similar to the distribution process between the TLK boards and the SM64/ SM128 boards, the SM64/SM128 boards also distribute the traffic all the way until the reflection plane is reached. (NB! The folded switching topology is not indicated in Figure 2, but will be discussed later when describing the plane switches.) After the reflection plane is reached switching/routing is performed back from the reflection plane with a self-routing technique. Buffering is, as will be seen later, implemented in each switching element.

The RCC module indicated in Figure 2 contains a cluster of printed circuit boards for rack control, clock distribution and supervision of power and cooling.

We will come back to the HW in the Alcatel ATM switches later, but generally, space switching and/or routing put stringent requirements in this sense. An ATM cell coming in on 155 Mbps or higher, given a specific header value, may appear at any instant at any port, and with a minimum of delay a way must be found through the switch.

Since the VPI- and VCI-fields are of restricted size, signalling must reserve new VPI/VCI values (based on numbering analyses) for every link between all the ATM switches that may be involved in a certain end-to-end connection. This information must be stored in registers in the switches. Hence, at the moment when a cell with a certain VPI/VCI combination arrives at the input of the switch, the VPI/VCI registers must be read and a new VPI/VCI combination attached before the cells enter the correct output and are transmitted along the next link.

In the Alcatel MPSR switching concept, the new header (including the new VPI/VCIs) is physically attached to the ATM cell in the TLK board in the outgoing direction.

Alcatel's MPSR switch has distributed buffering in the sense that there are equally dimensioned centralized buffers in *all* switching elements in the switching system. The switching element is an integrated circuit chip which is used generically in the switching modules and also on the TLK boards. SW is used to set up whether the switching elements shall distribute the traffic or route the traffic to a certain outlet. Output link buffers are additionally applied on the TLK boards on the outgoing direction.

Alcatel's internal format for switching

As ITU-T (formerly CCITT) only recommends the properties of the interfaces, and not the insides of the switches, the vendors are free to benefit from various solutions locally, as long as the boundaries of their switches conform to the recommendations.

Alcatel has chosen the principle of Multi-Slot Cells (MSC) as the internal switching format. As we will see in this chapter, this principle has many advantages to conventional ATM switching formats.

Figure 3 displays the format of the MSC utilized in the Alcatel MPSR switching concept. The MSC principle is based on segmentation of the ATM cells into smaller units, called slots. One of the advantages with this is that smaller packets are more efficiently switched than large packets when it comes to statistical multiplexing. Switching elements for smaller units will also require less internal buffer capacity, and besides, the MSC principle can be used for efficiently switching other formats as well, such as ISDN, MPEG video frames, SDH, etc.

There is no practical limitation on the number of slots in the MSC. Every slot consists of a Slot Control (SC) field of 4 bits and an information field of 8 bytes (64 bits). The Self Routing Tag, SRT, described below, is always the first slot of an MSC. The Alcatel SRT is not only a routing tag, but also contains other types of information like time-stamps, priority bits, etc. As can be seen from Figure 3, an ATM cell will occupy eight slots in the MSC.

In many ATM switch designs, information is added to the ATM cells for improving the switching performance by being able to simplify the internal architecture. This is also the case for the Alcatel 1000 ATM switch. An SRT added in front of the cell makes routing possible from a certain inlet to a certain outlet by HW, provided that the switching topology matches the self routing technique. After the MSC has reached the reflection plane seen in Alcatel's folded switching topology, HW reads the SRT for each switching element which is passed and makes the decision of which way to go based on a simple selection algorithm.

In some cases, as with the MPSR switch, the distribution network part followed by the self-routing network part may cause cells to arrive at the output out of sequence. Obviously, the internal format must take care of this by adding a sequence number or alternatively, a time-stamp for the entrance. Re-sequencing can hence be performed at the output buffers. Alcatel has chosen to put a time-stamp in the SRT field and let all cells have the same delay through the switching network. This has the advantage that the switch fabric will try to conserve the traffic profiles from the service sources.

Although ATM as a technology has only two levels of priority (at least if we look at the Cell Loss Priority, CLP, bit in the ATM header), internally in the switches there might be more levels. These levels can be negotiated in the signalling process or semi-permanently. Alcatel deals with priority by designated priority bits in the SRT field and dedicated priority on request on different inputs.

Of course, also different kinds of control mechanisms and checks are performed in relation to the internal format, but they will not be discussed in this paper.

Basic MPSR switch building blocks

The overall goal for Alcatel's MPSR switch designers was to implement a scalable and efficient ATM switch with respect to the number of interfaces and to the traffic load. The HW became compact and the switch could be built with only a few types of modules. If we consider the transport part of the switch, only two types of printed circuit boards are necessary: The link termination boards, TLK, and the MPSR switching module boards, called SM64/SM128. Both the access switches and the plane switches described later use the same types of boards.

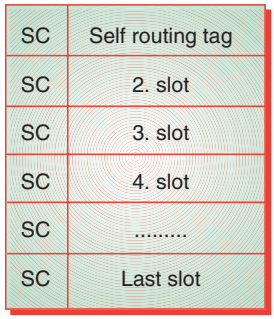


Figure 3 The MSC principle

The basic building block of the A1000 MPSR switching concept is the Integrated Switching Element (ISE) chip, see Figure 4. This chip is used both on the TLK board and on the switching board. Though the ISE chip for the moment is a 16 x 16 switching element chip (ISE16), it will later be upgraded to 32 x 32 (ISE32). As can be seen from the figure, a shared memory is controlled by the buffer management. The buffer management function collects routing information generated by storing and comparing the incoming SRT on the incoming MSCs with the routing registers. The shared memory can be implemented with dual port RAM, and in this case the access cycle is 27 ns (13 ns for ISE32). Such fast access times can be realized in 0.8 µm CMOS, which is commercially state-of-the-art now.

To obtain the goal of a Cell Loss Probability (CLP) better than 10^{-11} for an external traffic load of 0.8 E, it can be conservatively calculated that the internal traffic is 0.64 E (see (1)) and that the size of the shared centralized memory shall be 1.6 kbyte per ISE. The internal traffic is then calculated as if one of the four

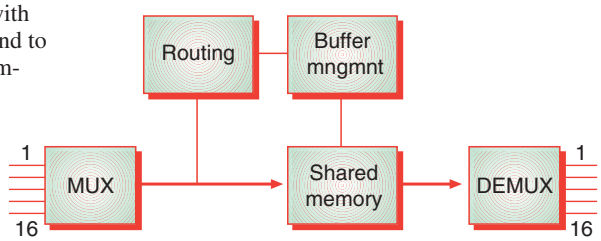


Figure 4 The ISE16 switching chip

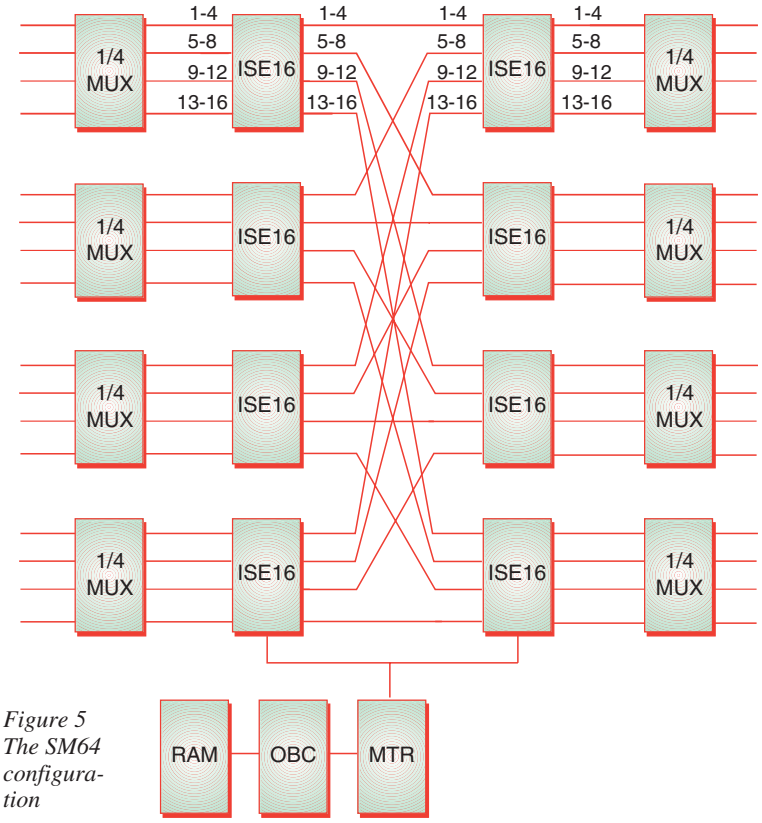


Figure 5 The SM64 configuration

access switch links or any other switching module has failed and the remaining three must take all the traffic. (1) shows that the MSC principle leads to a load increase of 21 %, as an ATM cell is packed in 8 slots, and that 0.8 E external traffic is expanded to 0.4 E internal traffic.

Base traffic load:
0.4 E x 4/3 x 1.21 = 0.64 E (1)

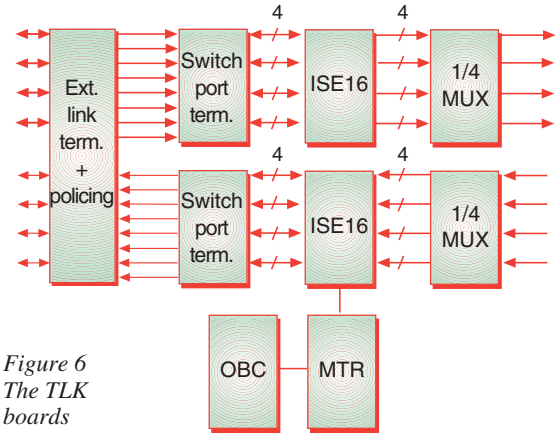


Figure 6 The TLK boards

The ISE chip is implemented as a VLSI, and will in addition to point-to-point services also offer multicast services. A controller interface towards the On-Board Controller, OBC, is provided, together with a test access interface.

Consider the switching matrix in Figure 5. The switching boards will always be built this way, independent of whether the ISE has 16 x 16 ports or 32 x 32 ports. The SMs are constructed as a two-stage matrix of 2 x 4 ISEs. Phase one in the development of switching boards has 64 inputs and 64 outputs (2 x 4 ISE16), and the related MPSR switching board is called SM64. Later, the switching boards will have 128 inputs and 128 outputs (2 x 4 ISE32) and will be called SM128.

The traffic in Figure 5 is drawn to go from left to right. However, depending on the configuration of the SM boards, they can be considered to be either bi-directional or uni-directional. The links between the SM boards are 622 Mbps electrical (within racks) or optical (between racks) interfaces.

Other components on the SM boards are the OBC and the Maintenance Test Routiner (MTR) chip, memory, 4/1 multiplexers and a DC/DC converter. The OBCs on all boards are transputers, which are dedicated processors for multi-task high-speed operation. The MTR chip used between the OBCs and the ISEs is a custom designed VLSI for O&M (Operation and Maintenance) on board level.

The TLK board in Figure 6 provides the access to the external SDH links containing ATM cells. Besides providing the SDH transport system with OAM information, the TLK performers unpacking and packing of ATM cells from the STM-1 VC4 containers. Each TLK board has 8 external inputs and outputs of 155 Mbps according to ITU-T. TLK boards will also be able to interface STM-4, but then the 8 x 155 Mbps interfaces will be replaced by 2 x 622 Mbps.

Alcatel has also developed a line-termination board for 34 Mbps which will not be described here. Other interfaces are also considered.

All the external input links are split in two in the switch port termination on the TLK boards, i.e. the 8 interfaces from external links will be transformed to 16 internal links before entering the internal switching network via the ISE switching element. The ISE will require the internal switching format, so the ATM–MSC conversion is also done in the switch port termination seen in Figure 6.

As with the SM board, the OBC, the MTR, memory, 4/1 multiplexers and DC/DC converter are also present on the TLK board.

Policing is performed on the incoming links of the TLK boards via a dedicated chip. For the moment peak bandwidth is checked, but later more sophisticated policing functions will be implemented as the standards evolve.

In Alcatel terminology, the TLK boards together with a number of SM boards are called a Traffic Switching Unit, TSU. A TSU will consist of up to 8 TLK boards and 4 SM64 boards, or respectively, 16 TLK boards and 4 SM128 boards. The four SM64 or SM128 switching boards in the TSU act as access switches with full switching capability. For the sake of security, the output from each TLK towards the switching boards are distributed equally to all the four access switches. If the number of external links exceeds the capacity of the TSUs (64 respectively 128 interfaces on 155 Mbps), plane switches described below must be introduced.

After the ATM cells, or rather the MSCs, leave the TSUs, they enter the so-called Plane Switches, PS. The PSs are built of SMs. There are two stages within the PSs; PS1 and PS2. PS1 is the stage 1 plane switch, meaning that traffic can be considered to flow bi-directional. PS2 is the stage 2 plane switch, which always is used as the reflection stage and hence the traffic can be considered to flow uni-directionally. In a PS the number of PS1s is twice the number of PS2s (PS = n x PS1 + n/2 x PS2). The PSs are often drawn perpendicular to the TSUs, to

visualize the independence between the number of TSUs and the number of PSs and also the independence in the distribution of traffic from the TSUs to the PSs. See Figure 7.

For the sake of traffic distribution and security, the traffic from the access switches in the TSUs are spread equally to all the plane switches that are applied. The cells that are distributed to a certain plane switch will stay in the same plane until they leave the plane switch and enter a TSU again. User traffic is reflected in the last equipped stage (TSU, PS1 or PS2).

HW requirements for ATM switching

ATM puts new requirements on HW-equipment for switching. In fact, the entire deployment of the new switching technology is dependent on a successful choice of technology which is or will be cost-effective, upgradable in speed, and which will not limit the system design.

Several technologies have been considered, such as GaAs, ECL and CMOS. A combination of technologies that gains popularity is the so-called BiCMOS, which combines the benefits of ECL and CMOS. All the technologies have pros and cons in certain areas, which will be briefly mentioned.

Gallium Arsenide, GaAs, has been a promising technology for some years now. It has excellent high speed properties and has both electrical and optical properties, i.e. lasers and detectors can be integrated together with the electronic circuits. When GaAs is not the chosen technology for the Alcatel switch, it is mainly because the process of making GaAs circuits is expensive and not suited for VLSI design, since the crystal structure contains too many imperfections.

ECL is a technology using bipolar transistors. The high-speed properties are excellent. However, the technology suffers from high power consumption and low integration.

CMOS is very attractive when it comes to VLSI integration, since the crystal structure in Silicium is possible to grow on large "wafers" without imperfections. The MOSFET transistors are easily processed onto the wafers on a more or less two-dimensional way. Although the gate-width dimension (e.g. 0.8 µm) is often referred to as the criterion for how many transistors there may be on a wafer, this size is really a fraction of the size of the complete transistors. Nevertheless, there is a certain relation between the width and the length of the gate and the rest of the transistor.

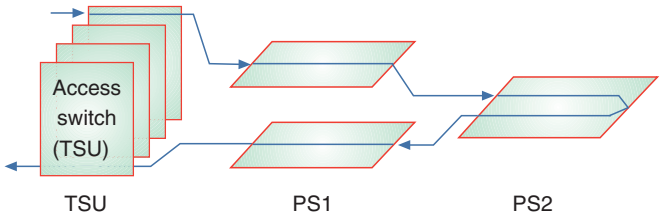


Figure 7 Access- and plane-switches

When the gate-width becomes smaller, the electrical field between the source and drain of a MOSFET transistor increases provided the power supply is kept on 5 V. At least two problems arise in this case: The process of being able to make such small dimensions is difficult since it is comparable with the wavelength of light and the supply voltage must be decreased (to around 3 V for 0.8 µm gate-widths) with the risk of ruining the signal to noise ratio of the system.

Apart from the fact that decreasing the gate-width means more transistors on the chips, another aspect is even more important for ATM switch designs, and that is the increased high-speed properties. These rely mainly on the fact that the charge is moved faster from source to drain, besides, the effect of smaller dimensions means reduced parasitic capacitances. With faster transistors one can build faster switches and buffers with lower access time.

Packaging techniques must cope with the heat-dissipation, which independent of technology will be quite high. To ensure junction temperatures low enough for long-term stable operation, a heat resistance in the order of between 5 and 10 K/W (Kelvin/Watt) must be obtained, provided a total dissipation per chip of 5 W. Alcatel has chosen the latest packaging technology to be able to fulfill this requirement.

The printed circuit board technology has been updated from S-12 to the new A1000 MPSR switch as well. Not only because of the expected explosion of heat dissipation per board but also because of other challenges such as the surface mounting technique where the pin-grid becomes smaller and smaller, and the electromagnetic radiation which increases exponentially with the clock frequency. Each board has thick copper layers for heat transport and electromagnetic shielding.

Software in the A1000 MPSR switching concept

The software has been and will be designed with the following external requirements in mind:

- Quick response to customer needs
- High level of availability, performance and capacity
- High Quality
- High Standards of Operations and Maintenance
- Overall reduced lead times.

A further objective has been to achieve minimum coupling between the control and transport functions, e.g. in order to increase the flexibility of interworking between separate control and transport vendors.

The relation between HW and SW is and must also be kept decoupled and the interfaces between the HW and SW will be generic and standardized. In addition, the SW as well as the HW, will be designed with separate and independent evolution in mind.

Of course, the concepts of Telecommunication Management Networks, TMN, Open Network Provisioning, ONP, and Intelligent Networks, IN, are taken into account when the SW development, maintenance and installation are planned.

Description and specification of the SW will be done by SDL. In some cases SDL will be used for automatic generated code. Most of the SW is and will be written in C and C++.

The SW can be divided into two main parts; the Kernel Software and the Application Software. Both parts will be described briefly.

Kernel SW

The Kernel SW basically covers the following items:

- The Operating System
- A data base
- An internal communication service
- Configuration Manager
- Error Handler
- Hardware Dependent Software
- Test support
- Means for initialization.

For different parts of the A1000 switching system there are different OSs. On the OBC level CHORUS is used. CHORUS is a small, highly portable, real-time, communication-oriented micro-kernel, and provides generic services used by subsystems to provide higher-level, network-transparent, OS interfaces. For the different Control Stations, CS, other OSs are used depending on the platform.

All database access is done through standard SQL. Various database systems can be used, as this is completely transparent for the application. On OBC level DMS (Data Management System) is used, whereas on CS ORACLE database is used.

The ICS (Internal Communication Service) is an Alcatel proprietary communication service, providing all the facilities necessary to make the SW modules able to send and receive messages and establish and release communication sessions.

Application SW

The Application SW is SW for signalling, connection handling and OAM functionality.

Signalling in ATM can be meta-signalling, broadcast signalling and point-to-point signalling. Implementation of SW for signalling is time critical, and it must be prepared for future upgradings and enhancements.

Connection handling contains several SW parts, such as means to set up semi-permanent connection, trunk resource manager and bearer connection control.

Application SW will be developed as the standards and market needs evolve.

Rack and interconnection practice

One Alcatel A1000 MPSR switch rack measures 220 x 90 x 60 cm. This rack can take 16 TLK boards with SM64 technology (32 TLK boards with SM128 technology) and 24 MPSR switching boards. In other words, 128 links (256 links) of 155 Mbps is the maximum that can be connected to one rack.

To exploit the maximum capacity with the SM64 technology, 2,000 links of 155 Mbps must be equipped. This means 16 full racks of 128 TLKs. Later, when SM128 is available, 16,000 links of 155 Mbps can be connected to the same switch. This means 64 racks filled up with TLK boards.

Interconnection between the racks will be performed with 622 Mbps optical links, i.e. 4 x 155 Mbps multiplexed into the same fibre. The optical connections will be terminated at the backplane, to avoid excessive power consumption on the boards.

Conclusion

With the multipath, self-routing technique associated with the multislot cell transfer mode Alcatel has a concept for ATM switching which is unique. The "buzz-words" for the Alcatel A1000 broadband switch can be highlighted as follows:

- *MultiPath*
  is the concept of randomizing the incoming ATM cells to all possible internal links in order to avoid internal overflow in the buffers and create a harmonized traffic load.
- *Self-Routing*
  means that from the reflection plane in a folded switching topology, the ATM cells are routed to the correct output by dedicated bits in an overhead, whose properties are so that, independent of where the cells are coming from, they will be led to the correct output.
- *Multi-Slot-Cell*
  is Alcatel's internal format for switching ATM cells and possible other formats as well. An MSC requires less internal buffer capacity than ATM cells and will hence contribute to less internal delay.

The Alcatel switch will try to preserve the traffic profile by delaying all the cells equally through the switch. This will be of great importance for policing and traffic shaping.

The distributed control architecture with an on-board controller on each printed circuit board, controls and monitors all transport functions. A high-speed internal transport protocol provides control-communication between the different parts of the switching fabric.

From 8 to 16,000 links on 155 Mbps can be attached to the MPSR switching fabric, which is expandable both in the number of external links by adding TLK boards, and in traffic capacity by adding Plane Switches, i.e. SM boards.

ATM – the rails for "Infobahn"

BY KJELL GJÆRE

ATM (Asynchronous Transfer Mode) has passed the initial stages of a new technology. It now has to compete with established and evolving technologies such as FDDI, switched 100 Mb/s Ethernet, etc.

So what is so attractive about ATM? Among a number of potential benefits, I will point out two:

- ATM is extremely scalable. It can run on almost any known transmission system, adapting to any bitrate, or it can function as a transmission system itself.
- ATM is independent of any service and any protocol. The ATM switches can be made to perform very fast switching, currently running at 155 Mb/s per link, 622 Mb/s in field trials and even higher speeds to come.

How does ATM work?

To a large extent, an ATM connection is set up as in ordinary telephony. In a public network, each file server, video server, workstation or telephone will be accessed by a unique address as specified in E.164 by ITU (International Telecommunication Union). This will be a globally unique address similar to the country codes, area codes and local exchange numbers being used today.

Whether you are shipping data, voice or video across a private or public network, you must first establish a connection from your terminal to the desired destination. To do this, you must 'dial a number'. By signalling, you tell the network the identity of the desired destination. Your local switch will work with every node along the path to establish the connection. The one important difference between telephony and ATM is that during call set-up, ATM also has to negotiate the bandwidth and quality of service, QoS, for your connection. Each node will have to sign up to support this contract.

This traffic contract specifies characteristics such as peak cell rate, mean rate and maximum burst length. End-to-end the terminals will agree upon whether the connection is constant bitrate or variable bitrate. Based on the type of traffic agreed to, the switches will allocate a priority level to the connection. This priority will be used by the switches to decide which cell gets transmitted first if two cells happen to reach the same outlet at the same time. Priority also determines which cells get dropped when a link becomes overloaded.

Having introduced the 'cell', what is it?

In ATM all information is divided into packets being 53 octets long. These fixed length packets are denoted 'cells'. The fixed length and format of a cell is extremely important to make an efficient switching fabric in the ATM switch nodes. The cell is divided into two parts; the header and the payload. The header containing the connection identifiers and the payload containing the actual information conveyed from the initiating terminal to the receiving terminal.

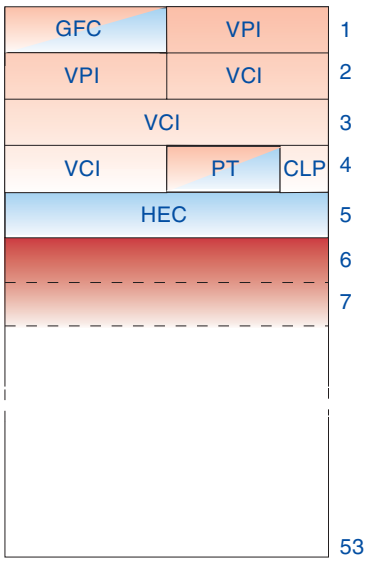


Figure 1 The ATM-cell

How do the ATM-cells find their way?

ATM-systems and switches have much in common with ordinary telephony switches as we see them in modern digitized telephony systems.

On transmission links, a time-slot in a TDM scheme is allocated to each telephony connections. A specific telephony connection can be identified on a specific link in a specific time-slot.

In a switch, the content of different time-slots are switched to different outlets (space switching). At the outlet, a number of connections may contend for the same time-slot. Thus, the outlet must have a short buffer to multiplex the different connections out on the link (time switching). In this scheme each telephony connection has its own dedicated path consisting of dedicated time-slots on the links and dedicated time- and space-links in the switches, a connection oriented line-switched system.

ATM-switching has lots of similarities to this scheme. On the links we find cell-slots and in the switches space-switch functions to convey a specific cell from any one inlet to any one outlet. So in an ATM-switch we find time-space-time switching and multi-staging as we do in current telephony switches.

So what is the difference?

On an ATM-link the different connections are not identified by the position of the actual cell-slot. A number of consecutive cell-slots may contain cells belonging to the same connection. Arriving at a switch, the switch must analyse the 'connection number' contained in the cell header to know which outlet the cell is to be switched to. At the outlet, the cell can be transmitted at the first available cell-slot, not waiting for a specific position. There is one restriction, the switch shall maintain the integrity of the cells. Cells being received in a specific order shall be transmitted in the same order, but not necessarily with the same inter-cell time delay.

On an ATM-link, the link capacity is divided into virtual paths, which may in turn contain several virtual channels.

A specific cell travelling down the link is identified as belonging to a specific path and channel by the Virtual Path Identifier, VPI, and the Virtual Channel Identifier, VCI, in the cell header. To save space in the header, an ATM-cell does not carry all the routing information to get it to its destination. The header only contains enough information for the next switch to identify the cell. The switch

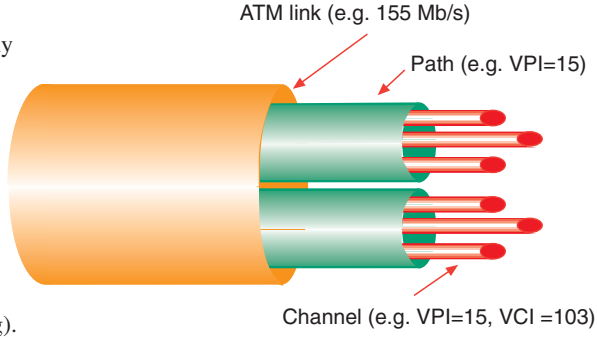


Figure 2 Paths and channels on an ATM-link

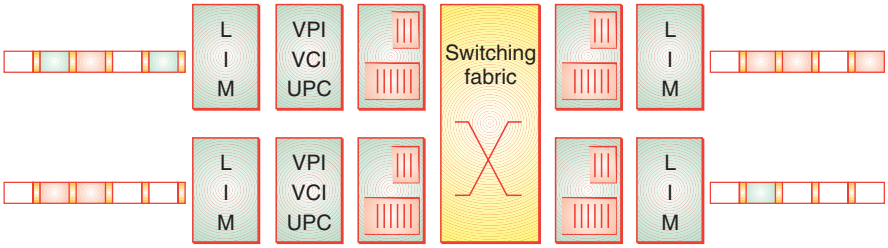


Figure 3 Main building blocks in an ATM-switch

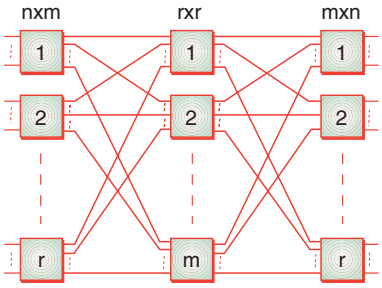


Figure 4 Multi-stage switching fabric

must maintain an address translation table to know where to send the cell and the new VPI/VCI values to be set in the cell header for the next hop.

Besides keeping track of where each connection is going, the switch should also know the grade of service which is allocated to the connection. This is important for two reasons: First, the switch must keep track of how much total bandwidth that has been allocated to a specific link. Second, the switch must monitor the traffic on each channel to ensure that no single channel causes problems by sending more traffic than has been agreed upon using more of the switch's resources than its allocated share. The switch must 'police' the traffic. Policing will normally take place at the User-Network-Interface, UNI.

Now, having passed the basic concepts of an ATM switch, let us look at an ATM-switch in more detail.

Any ATM switch must contain four essential building blocks: the transceiver, transmission deframing, address translation and the switching fabric itself.

In the transceiver, situated in the Line Interface Module, LIM, electrical or optical signals are received at the switch port. On the incoming side, a clock is recovered from the signal and the raw bit stream is recovered. Normally, it will be converted to 8 bit in parallel to take advantage of low power consuming CMOS technology.

If the bit stream was encapsulated in a transmission frame, such as in an SDH STM1 frame, it gets deframed. The LIM also verifies the integrity of each received cell's header. The 5th octet in the ATM-cell contains an 8 bit cyclic-redundancy-check (CRC) code or Header Error Control (HEC) as it is denoted in ATM terminology. Single-bit errors in the header are corrected, double bit and most multiple bit errors are detected. Cells with uncorrectable errors are discarded. At the transmitting side, the link was filled up to full capacity with empty, idle cells to keep the link synchronised. These idle cells together with cells with multiple bit errors, are now discarded as the cells enter the next switch.

The active cells will now be passed to the address analysis and translation function. The Usages-Parameter-Control, UPC, function, or 'policing' for short, will check that the actual Virtual-Connection, VC, keeps to its contract. The cell header gets its new VPI/VCI address for the next leg and the route through the switch fabric, to the actual physical outlet from which the cell shall leave the switch is determined.

The cell is passed trough the switch fabric to the actual outlet. The switch fabric itself may, in a small, say 8-by-8, switch, be as simple as a common back-plane bus. In larger, central office type, switches with 1 – 2000 outlets, the switch fabric will be a multistage switch consisting of a number of, say 32-by-32, switching blocks, arranged in a Banyan, Clos or similar manner.

In the outgoing part of the LIM, new HECs are computed, the cells are framed into transmission frames and idle cells are added to fill up the link. The transceiver sends it out on its next leg.

Congestion, the challenge of ATM

It quickly becomes apparent that cells arriving asynchronously to the switch may contend for a given cell-slot on one outgoing link. Also, variable bitrate traffic, bursty in nature, may overwhelm the maximum resources available in the switch fabric. These conditions are known as blocking and congestion.

Even if two cells are headed for different outlets, they may have to pass a common internal link or switching element at the same time. If the switch cannot accommodate both cells at the same time, one cell must be buffered or even discarded.

Internal blocking can be avoided by proper design of the switch. Output blocking, however, is unavoidable due to the nature of ATM. This happens every time two cells compete for the same outlet in the same cell-slot. Only one cell can pass at a time. To avoid loss of cells, there must be buffers in the switch fabric.

These buffers, wherever they reside inside the switch, may be arranged to serve the cell-stream on a first come, first serve basis (FIFO) or as a number of parallel queues, all with a different priority. In this way, time bound traffic like voice and video may pass data traffic which can stand a few milliseconds delay.

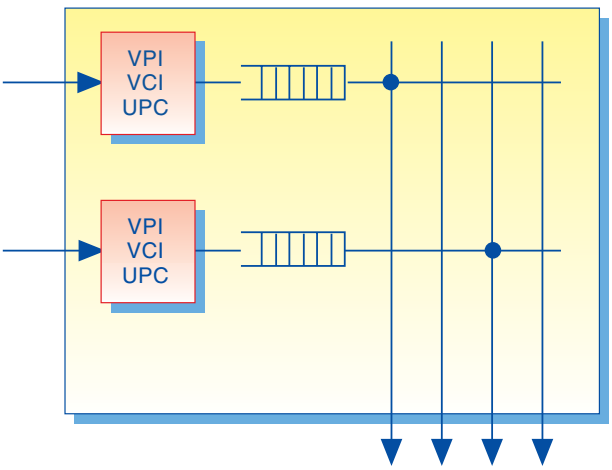


Figure 5 Input buffer architecture

Due to the nature of ATM, even if all connections where set up as Constant-Bit-Rate (CBR) traffic, collisions may still occur on a given outlet at one point of time. As the buffers must have a finite length, probability of buffer overflow and eventually cell-loss occur. In situations with traffic surges in a Variable-Bit-Rate (VBR) traffic scenario, this problem is even grater. These problems are solved in a number of ways.

First, traffic surges may be smoothed out by the transmitter. The transmitter 'shapes' the surge to be more like CBR traffic over a period of time. Second, the internal buffers queue up shorter bursts, and third, the network can inform the transmitter to speed up or slow down according to available resources in the network

It becomes evident that buffer-size and buffer position in the switch fabric is one of the key parameters in ATM switch designs.

Buffer layouts in switching fabrics

Buffers are necessary in ATM switching fabrics to ensure a low cell-loss rate. Deeper buffers are needed than in other time-multiplexing systems due to the asynchronous nature of ATM and bursty data traffic. On the other hand, if too deep, they cause latency into a connection and buffers cost money!

Buffers may be placed at different places in the fabric: input buffering, output buffering, cross-point buffering and a shared memory scheme.

By using input buffering, the switch fabric itself can be made very simple.

There is one major problem, though. For example, if the next cells to leave from two different buffers are bound for the same one output, one of them must wait. This waiting cell will in turn block its subsequent cells in its buffer from being serviced. This Head-Of-Line (HOL) blocking will slow down the throughput of the switch. Some designers have used RAM as buffers. In this case, they can pick the next cell from a buffer at random. By using a sophisticated scheduler, they have dramatically reduced the HOL problem. On the other hand, this scheduler, keeping track of all traffic coming in on all ports and the status of all buffers, is not easily scalable into larger switches.

By using output buffering the HOL problem is gone.

In this solution the internal space-switch must run at a much higher speed. In the same one time-slot there may be cells all bound for the same output. To prevent cell-loss, in worst case the fabric must have the ability to service all inputs and move one cell from each to the same one output buffer in one cell-slot.

Another disadvantage is in a case where all traffic is headed for one or two outputs, a lot of buffer-space is unused at the other output. The ATM switch cost, therefore, is driven to a large extent by the cost of memory. The memory requirement can be reduced in a shared memory solution.

In a switch with shared memory the buffers are common to all outputs.

In most cases a TDM bus on input and output sides of the buffer will transport all cells from the inputs to the buffer and accordingly from the buffer to the different outputs.

Studies have shown [1] that the total buffer capacity of shared memory based switches is approximately 5 – 7 times less for a central buffered switch than an output buffered one with 16 – 32 ports. Another advantage is that it is easy to divide the common memory into buffers with different priority, letting time dependent voice and video get priority over data traffic.

Small versus large central office type switches

Up till now we have discussed the basic principles of an ATM-switch with its input, output or shared memory buffer architecture. For smaller switches, say 16-by-16 ports, this discussion is somewhat exhaustive.

Telenor AS is currently installing 4 switches from FORE Systems Inc. of these 'suit-case' sized switches for the Bynett service in Oslo. These switches have 16-by-16 ports. The ports are grouped 4-by-4 on 4 interface cards. By choosing interface cards, the switch may run with 155 Mb/s SDH, 34 Mb/s E3 or 100 Mb/s TAXI interfaces. At this point in time the switch cannot be expanded beyond 16 ports. (Or 24 if all ports are TAXI-type.)

For larger, central office type switches, switching blocks are connected in a multistage configuration to form a switch with thousands of ports.

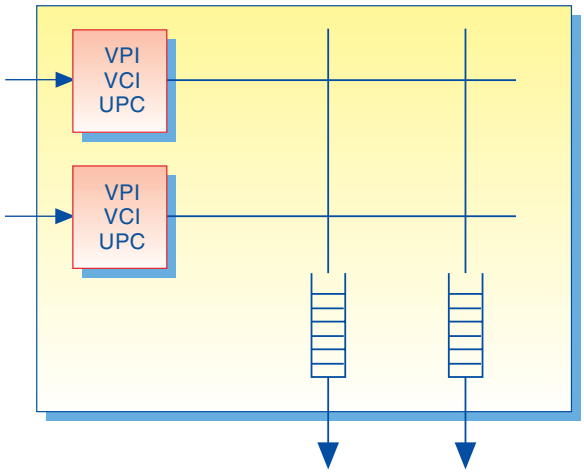


Figure 6 Output buffer architecture

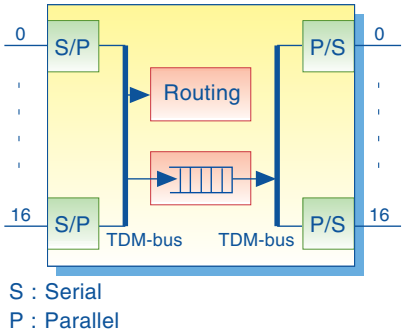


Figure 7 Central buffer architecture

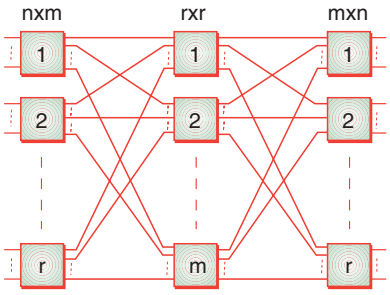


Figure 8 Multistage switching fabric

As the Norwegian node in the European ATM-pilot network, Telenor is using a switch from Alcatel, the A1000.

This switch is scalable up to 2048 ports running on 155 Mb/s SDH links.

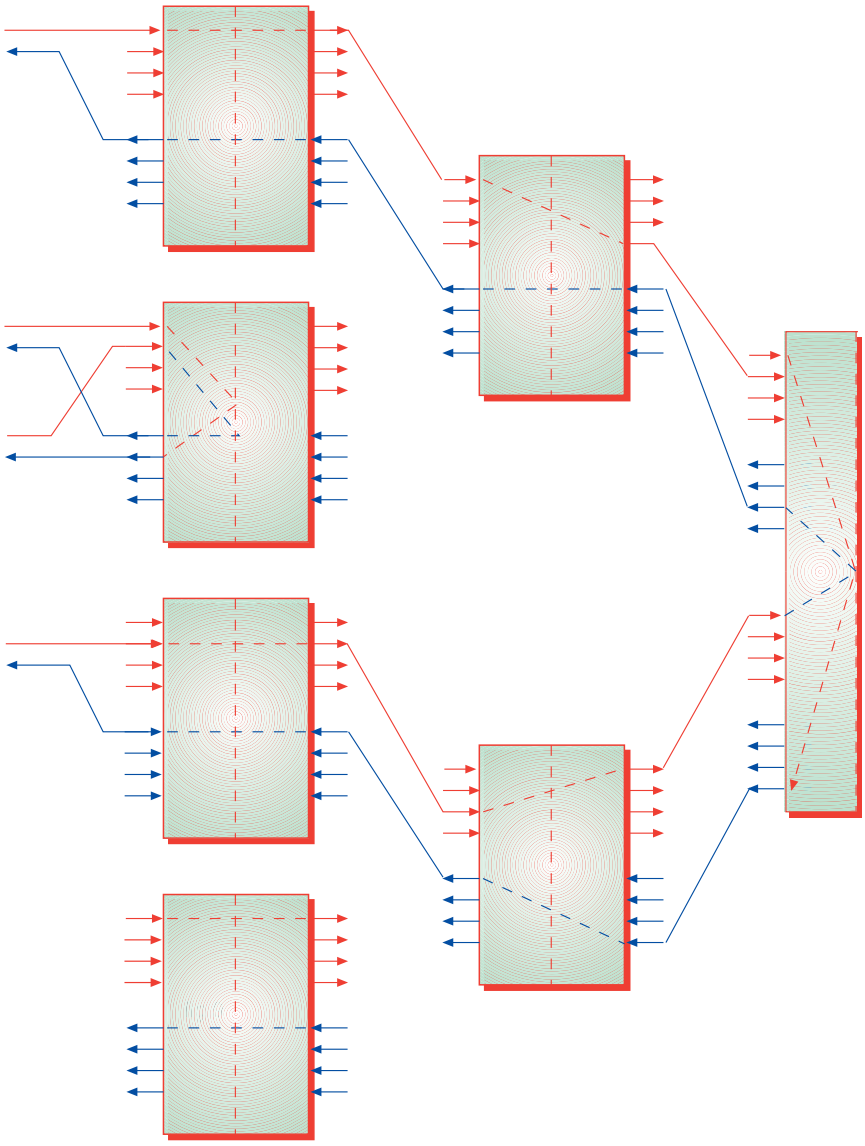


Figure 9 Folded multistage switching fabric

This switch is arranged as a multistage, folded network built from 32-by-32 port switching blocks.

The switching blocks are 'split' into two halves; one half carrying traffic from left to right, the other half carrying traffic from right to left. By this arrangement, most cells need not traverse the whole switching fabric. As soon as the cell has moved deep enough to a point from where the output can be 'seen', it is reflected and is switched back to the actual output.

Another quality of the A1000 switch is its ability to distribute the traffic over the whole switching fabric. As the cells arrive at one input, the VPI/VCI values are substituted for the next transmission leg. In addition, a routing tag is added to the cell. For short, the routing tag is the physical output number for the port from where the cell shall leave the A1000. The switch will transfer the cells in a connectionless manner through the switching fabric, always trying to avoid contention and links with instantaneous high loads. In this way the instantaneous load on the switch is distributed over its whole switching fabric. As the cells from one connection may take different routes through the switch, the cells may arrive at the output port out of order. The routing tag also contains a time-stamp, the time at which the cell arrived at the input port. This time-stamp is used at the output to transmit the cells after a fixed delay regardless of instantaneous load and delay in the switching fabric itself. Under normal conditions this will take care of the cell integrity, transmitting the cells out in proper order.

Conclusion

ATM has been selected the technology to carry Broadband-ISDN, B-ISDN. ATM is currently entering the market at a much higher speed than was expected a few years ago. Even if there is a lot more work to be done, especially on the signalling part, to get B-ISDN up and running, basic switching technology is mature enough to carry and switch ATM cells at speeds of 155 Mb/s.

Let's go and look for the services that can take advantage of this new technology!

Reference

1 Prycker, M D. *Asynchronous transfer mode : solution for broadband ISDN*. New York, Ellis Horwood, 1991. ISBN 0130535133.

Photonic switching techniques for broadband applications – current status and perspectives

BY JEAN-MICHEL GABRIAGUES AND JEAN-BAPTISTE JACOB

The development of interactive services requiring a large bandwidth at each user access depends on both the growth of the demand and on the availability of transport and switching systems capable of handling gigabit-per-second data streams. In this paper, early photonic ATM switching experiments are reported. The emphasis is put on the description of the OASIS switching matrix concept, implementation and performance. Two other concepts, the fibre loop memory switch and the multidimensional switch, investigated within the RACE ATMOS project, are also introduced.

1 Introduction

The very high speed requirements that characterise interactive and real-time high-performance applications like parallel processing, video services and high-quality imaging, initiated a considerable interest in networks that could provide a large bandwidth to the users. The operation of such networks, supporting various services at different bit rates, will preferably rely on the Asynchronous Transfer Mode (ATM) technique, a key multiplexing mode characterised by standard format fixed-length cells.

In the future Broadband Integrated Services Digital Network (B-ISDN) the photonic technologies will play an essential role, since the transmission bandwidth of optical fibres virtually offers unlimited transport capacities. A future broadband switching node will require a switching capacity of several Tbit/s, fed by input links with rates ranging from 155 Mbit/s to 2.5 Gbit/s. The switching node bandwidth and speed could therefore become a bottleneck and photonics appears attractive to ensure high performance, increasing the switching node throughout, operating speed and flexibility.

In fact, as the transmission will be performed through an optical infrastructure, it could be an advantage to keep the information in an optical form during the switching process. This is a long term goal, since photonic techniques now offer modest performance in terms of signal processing capability. Nevertheless, recent successful experiments carried out in various laboratories suggest that the development of photonic switching matrices could become realistic within a 10-year perspective [1–3].

The cornerstone of this paper will be the feasibility demonstration of the so-called OASIS (Optical ATM Switching Integrated System) high-speed switching matrix [4–6], performed by Alcatel. We shall start by a brief review of the key optoelectronic components required to perform the basic switching functions, then recall the principle of operation of OASIS and the recorded performance, and finally compare our approach to ATM switching to other approaches studied within the RACE project ATMOS.

2 Key optoelectronic components suitable for broadband switching

Today, the range of integrated optoelectronic components suitable for high-speed switching, i.e. having nanosecond response times, is rather limited. Basically, what is needed consists of fast active crosspoints capable of steering optical beams from one input port to several output ports; unfortunately, most photonic integrated space switches are either too slow (e.g. digital optical switches), or exhibit performance highly influenced by the state of polarisation of light (e.g. directional couplers in III-V semiconductor technology), or need high control voltage for proper operation (e.g. directional couplers in $LiNbO_3$ technology. In the practical sense, active components capable of switching in the wavelength domain are more attractive: wavelength tuneable semiconductor lasers, now available in the laboratory, are at the basis of development of the wavelength routing technique, whereby optical signals to be switched are first assigned one wavelength dedicated to one destination throughout the switching network, then mixed together and finally discriminated by means of a wavelength demultiplexer.

The basic "wavelength switchable" component is the wavelength converter. From a functional standpoint it is a device featuring one optical input, one optical output and one electrical control input; it should be capable of transforming an incoming modulated optical signal into an image optical signal, the wavelength of which is determined by the state of the electrical control signal. Beside the obvious solution involving optical-to-electric conversion of the input data, followed by the re-emission of the signal by using a tuneable laser source, considerable efforts have been devoted to the design of "all-optical" Indium Phosphide based monolithic devices: one of the most attractive solutions exploits the modulation of the internal gain of a semiconductor optical amplifier (SOA). The SOA-based wavelength converter, shown in Figure 1, is a monolithic amplifying element operated in the saturation regime, in conjunction with a tuneable source acting as a generator of an optical carrier at an adjustable wavelength. When two incoming signals, the input intensity-modulated signal carrying the information, and the optical carrier from the tuneable source, are launched in the

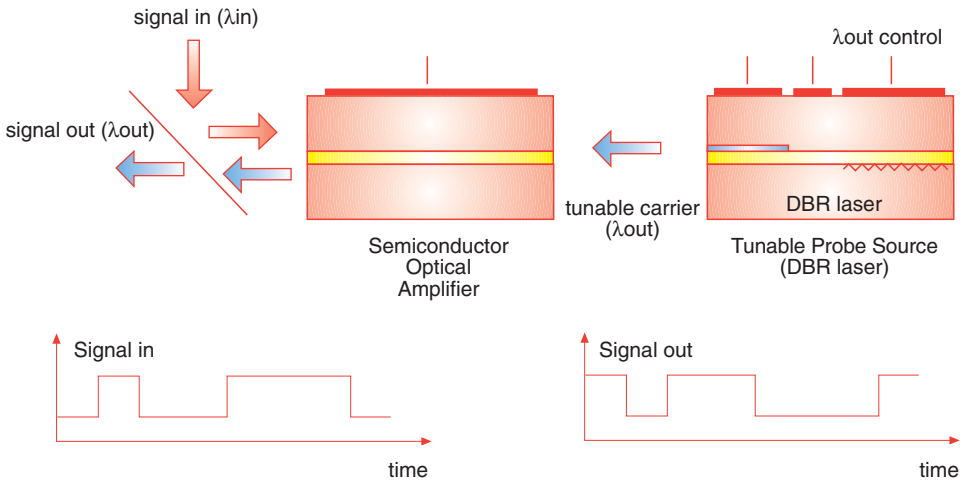


Figure 1 Selected key components for high-speed optical switching applications.
Top: Wavelength converter, based on a semiconductor optical amplifier and a tuneable DBR source
Bottom: Optical gate, based on a semiconductor optical amplifier

SOA, the gain experienced by the carrier is modulated by the intensity of the input signal; therefore, the output signal is an amplified combination of the two input beams, from which the signal at the wavelength of the tuneable carrier, carrying the "inverted" information, can be extracted. SOA-based wavelength converters are usable up to 10 Gbit/s data rate, over a broad range of wavelength, determined by the gain bandwidth of the SOA [7].

Besides using the wavelength routing technique, one can switch optical signals in the space domain by means of the "split-and-select" technique, whereas the waveguide carrying the signal is split into several branches, each being equipped with one optical gate: the routing function is accomplished by activating the gate located at the desired output branch. Today, the best solution for gating is to make use of gain-switched Semiconductor Optical Amplifiers (SOA-Gates). SOA-Gates benefit from the tremendous recent progress in the SOA technology [8], which makes them capable of simultaneously offering high gain (up to 20 dB fibre-to-fibre gain), low polarisation sensitivity (±1 dB using almost square-shaped waveguide section) and turn on / turn off times in the range of 200 ps under electrical control. Recently, monolithically-integrated SOA linear arrays have successfully been fabricated [9], and the perspectives for (hybrid) integration of SOA-G arrays with passive splitters in planar waveguide technology are quite good.

3 The OASIS ATM photonic switch

Alcatel Alsthom Recherche and Alcatel-CIT launched in 1990 a research project aimed at investigating the feasibility of key functions required in any ATM switching system by means of photonic techniques: the outcome of this project, called OASIS, was the public demonstration of a 4 x 4 photonic ATM switch operating at 2.5 Gb/s in September 1993 (ECOC'93 exhibition). The principles of operation and implementation issues have been addressed extensively in reference [5]. In this section we shall briefly recall the principles and characteristics of the OASIS switch, prior to discussing the advantages and drawbacks of the technical approach we have adopted.

The N-inputs N-outputs Switching Matrix, illustrated in Figure 2, is composed of four different functional units:

- The cell encoders, located one at each matrix inlet and based on a wavelength conversion tuneable laser, which converts the incoming cells at the wavelength associated to the target outlet
- The cell buffer, based on a set of *K* optic fibre delay lines, accessible via an *N* x *K* optical network composed of N diffusers, *N* * *K* optical gates and *K* combiners; the purpose is to delay the cells by an appropriate number of time slots in order to avoid output collisions between cells with the same destination
- The demultiplexer, based on a passive optical star coupler and a set of *N* fixed-wavelength optical filters, which enable each outlet to select only the cells addressed to it
- The routing and control logic, essentially based on an electronic command memory, which drives the wavelength converters (routing function) and the *N* * *K* optical gates (output queuing function) on the basis of the address and function information contained in

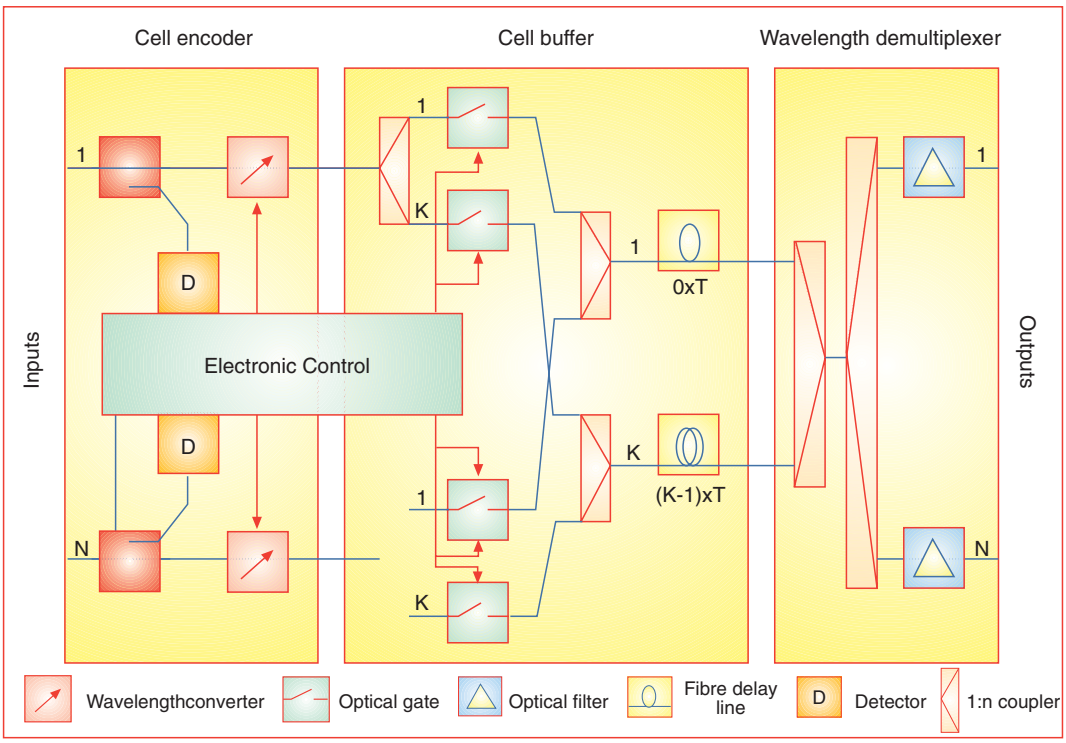


Figure 2 Basic OASIS architecture

the cell internal tags, all processed simultaneously at each time slot.

The space switching principle is the wavelength routing scheme, whereby each incoming cell is assigned the wavelength corresponding to the desired filtered output. The time switching scheme is based on the use of variable delays to solve the contention problem: K fibre delay lines are implemented, the delay introduced by the k^{th} fibre being equal to K time-slot. This realises a space switch architecture with pure output queuing, where the parameter K is determined on the basis of quality of service considerations, given the traffic load and the maximum value of the cell loss probability. The function of the matrix is to switch the received ATM cells at a bit rate of 2.5 Gbit/s, on the basis of the information contained in their header. In our demonstration, we have chosen to substitute a "routing tag" for the original header section of the cell. The routing tag consists of a series of $n + 1$ bits, aimed at identifying the output address (with $N = 2^n$) and whether or not the cell is an idle one (1 bit).

The principle of operation is the following. It is assumed that the incoming cells arrive in phase at the N inlets together with their associated n-bit long tags. While the tag is selected, the cell is delayed in a suitable length of fibre, with a transit delay duration of less than a time slot, and then launched into the wavelength converter laser chip. The routing information deduced from the electronic analysis of the cell tag is used to address a control parameter table, which, in turn, delivers the binary value corresponding to the required drive current to be applied to the converter, in order to tune it on the appropriate wavelength just at the time the first optical bit of the cell is launched into the chip. Thus, at the output of each converter, the cells have been wavelength encoded.

Then the problem is to avoid collisions between synchronous cells with identical wavelengths: this is the role of the time switch, which is an optical partially shared buffer. It consists of a set of K optical fibre delay lines, with increasing lengths corresponding to relative propagation times of 0 x T up to (K - 1) x T, where T is the time-slot duration. The output of each wavelength converter is connected to all the K fibre delay lines via a 1-to-K diffuser and a set of K optical gates, since each fibre can support up to N cells at N different wavelengths in the same time slot. On the other hand, one wavelength-encoded cell only must enter one delay line (later is described the case of multipoint connections), and for this purpose a contention resolution algorithm is implemented to randomly, but fairly, determine the queuing order between the contending cells. Out-of-sequence is impossible, because only cells belonging to different connections enter the switch in the same time slot.

Finally, the time switched cells can propagate towards the different matrix outlets throughout the space switch. The outputs of the K optical fibre delay lines are combined and connected to all outlets, by means of an optical star coupler. Each outlet is a port equipped with a fixed-wavelength optical filter, the peak transmission of which is tuned to one of the N possible wavelengths the cell encoders can accommodate. Thus, each filter receives all the cells, but selects only the appropriate traffic stream at each outlet.

4 Switching matrix performance

For our demonstration (Figure 3), we have chosen to build up a fully-equipped demonstrator, operated at 2.5 Gb/s, featuring four inputs, four outputs and one four-element optical buffer, considering a realistic trade-off between the need for a critical size of the system and the level of effort to be devoted to the fabrication of a large number of packaged advanced components.

The aim of the demonstration was to show both wavelength routing at high speed and contention resolution using all optical buffering technique. The inputs of the switching matrix were connected to a traffic generator, supplying ATM-type binary patterns, via an E/O interface board, and the outputs of the switching matrix were connected, via an O/E interface board, to a transmission testing equipment capable of recording Bit Error Rates, BER, and of displaying the outgoing signals.

The traffic generator was basically a set of 4 programmable binary pattern generators, operating synchronously at 2.488 Gb/s. For each cell, one five-bit routing tag was programmed in the header section, in order to provide the matrix control with proper information: in order to reduce the operation speed of the control circuitry, the tag was encoded at a bit rate of 622 Mb/s in the original "header" section of the cell; one bit was used to distinguish between idle and real cells, the four other bits to identify the destination, the control having been designed in order to ensure up to 16-wavelength operation. Moreover, a three-octet guard band was used between cells to accommodate the tuning delay of lasers in the wavelength switching operation.

BER measurements and eye-diagram displays were performed at one output at a time. For that purpose, the input traffic was programmed so as to provide a continuous cell stream at the output under test (i.e. without any idle cell). There was no clock recovery, in order to avoid disturbances resulting from non-ideal synchronisation of output cells.

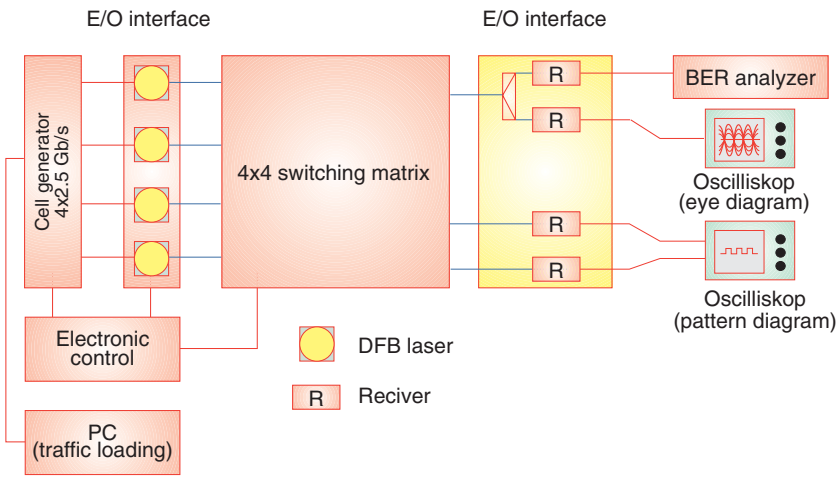


Figure 3 Schematic of the OASIS demonstration

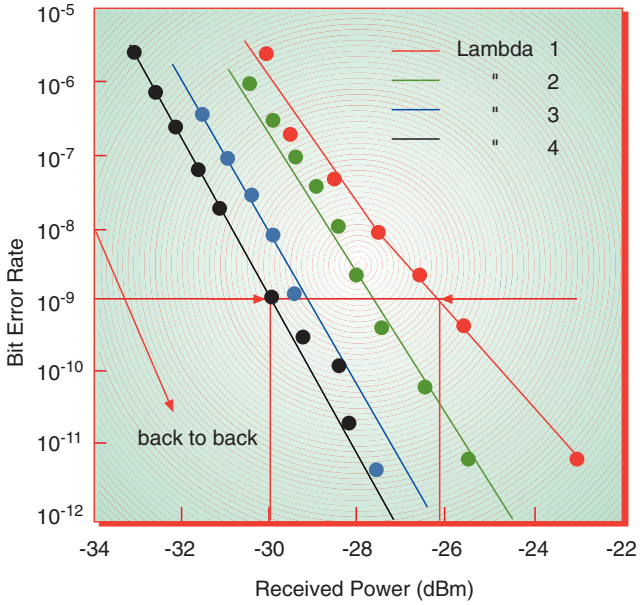


Figure 4 Switch matrix performance at 2.5 Gb/s

Figure 4 shows the switching matrix performance, characterised in terms of bit error rate figures, and as such presents the combined effects of all the sources of degradation that affect the optical signals propagating throughout the switching network. The ATM traffic configuration was settled so as to feed the inputs of the switching matrix equipped with all-optical wavelength converters with sequences of cells aiming at outputs 1 (wavelength lambda 1) to 4 (wavelength lambda 4), including some degree of contention compatible with the size of our fibre buffer. The BER record of Figure 4 shows the penalties experienced by the optical signals after the switching process: this average penalty is in the order of 3 dB, with noticeable difference between the four outputs, mainly resulting from imperfect simultaneous optimisation of the operation at all wavelengths. The average penalty results from the small reduction of the extinction ratio of the signals after wavelength conversion and gating, the spurious loss unbalance between the different paths throughout the switching network, and a small polarisation dependence of the SOA-Gates. The limitations brought by the cell buffering technique, relying on calibrated delays, arise from the unbalance between the various optical paths that signals are experiencing throughout the switch; this can be accounted for by considering the differential losses and the differential delays between various paths.

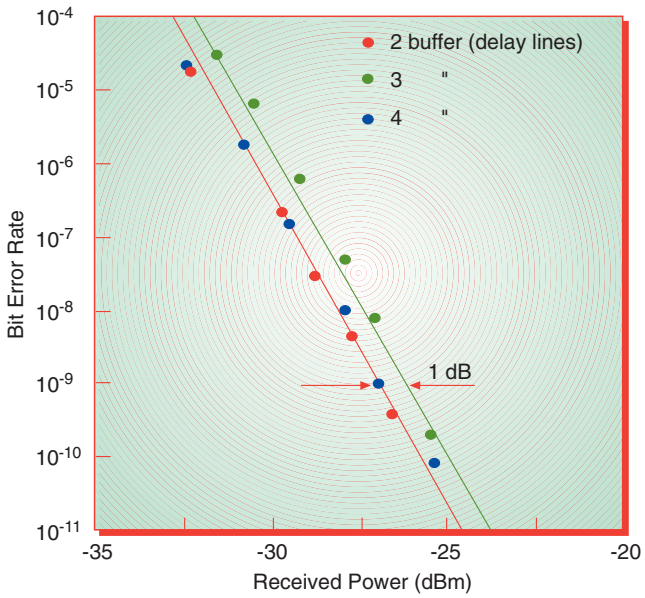


Figure 5 Penalties recorded as a function of the size of the fibre buffer

Figure 5 illustrates the impact of loss and delay unbalances, by presenting BER plots obtained for different traffic configurations leading to cases where propagation occurs throughout 2 (red dots), 3 (green dots) and 4 (blue dots) delay lines. As can be seen the recorded sensitivity penalties do not exceed 1 dB, and are not increasing with the length of the delay, and hence the size of the buffer.

Using these experimental data, we have modelled the physical limitations of the proposed switching architecture and tried to evaluate the sizeability of our switching architecture. From Figure 6 it is clear that a maximum of 16 inputs/outputs could be envisaged, given the present status of the technology [5].

5 Alternative approaches for ATM switching

In this section we shall examine alternative solutions for optical ATM switching, with emphasis on work carried out within the RACE project ATMOS (ATM Optical Switching).

The Fibre Loop Memory Switch Concept, developed by CSELT [2] aims at performing the same functions as the OASIS one, but relies on a completely different architecture. As shown in Figure 7, the switching function only exploits the wavelength domain to encode the incoming cells, to buffer them in a wavelength-controlled recirculating fibre loop memory (if necessary) and to finally route them to the final destination. The basic idea is to make use of the possibility of simultaneously storing several cells, identified by different wavelengths, inside one single fibre loop. The cell encoding is performed only on the basis

of vacant memory positions and not for the purpose of identifying the actual path throughout the switch. The handling of optical cells is to be performed by dynamically tuneable filters. A 2 x 2 switching module has recently been demonstrated [2], showing practicality of the fibre loop buffer; the main limitations arose from the accumulation of optical noise within the loop, as a result of the need to amplify signals circulating in the loop. The sizeability of the concept is now being evaluated and first results show figures close to the ones previously described for OASIS.

The multi-dimensional Switch Concept, developed by Alcatel SEL [10] is based on the exploitation of space and wavelength dimensions for overcoming the lack of large optical memories as well as for the demonstration of high throughput switches with contention resolution at the output ports. The idea is that the collision of contending cells can be avoided simply by assigning them different wavelengths: therefore, they can simultaneously propagate throughout the switch following the same physical path. Contention resolution is to be solved only at the output of the switch, either by optical means, as shown above, or by means of electronic buffers. Once more, in this scheme, the wavelength is not identifying one physical destination address, but is used as a resource for the switching process: hence a number of wavelengths (depending on the traffic load and expected quality of service) have to be generated within each switching module to cope with contention at the different inlets. A 2 x 2 switching module has recently been demonstrated, showing feasibility of the proposed concept; the main limitations arose from the losses of the space switching stage (see figure) whose size has to be equal to the number of outputs multiplied by the number of wavelengths.

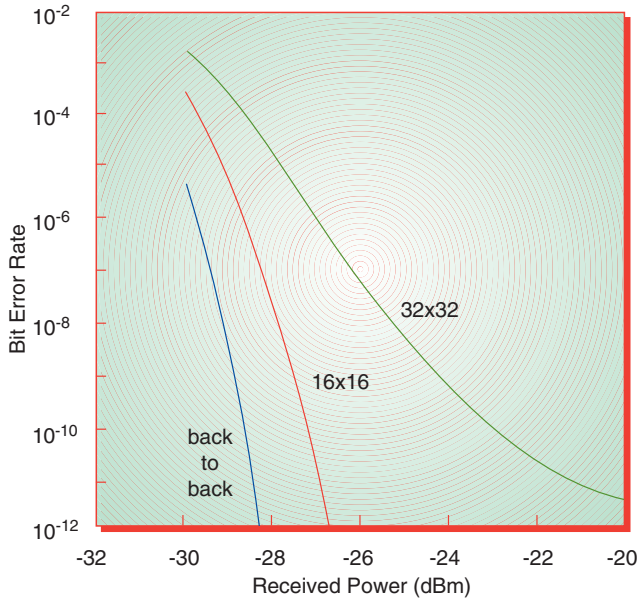


Figure 6 Calculated sensitivity penalty as a function of the number of inputs/outputs of the switching matrix

6 Conclusions

In this paper, we have presented advanced solutions for photonic ATM switching. The OASIS concept has been described and analysed in terms of performance: a 4 x 4 fully functional ATM switching matrix has been assembled and tested. The operation principle relies on wavelength routing and cell buffering using optical fibre delay lines for con-

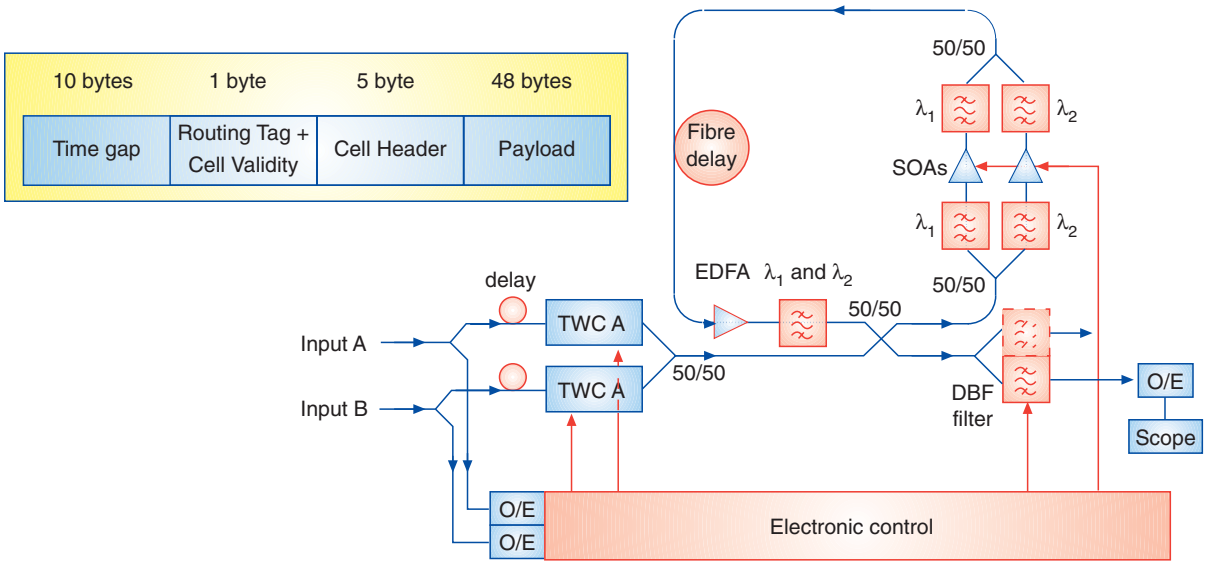


Figure 7 The fibre loop memory approach for cell buffering (TWC: Tuneable Wavelength Converter)

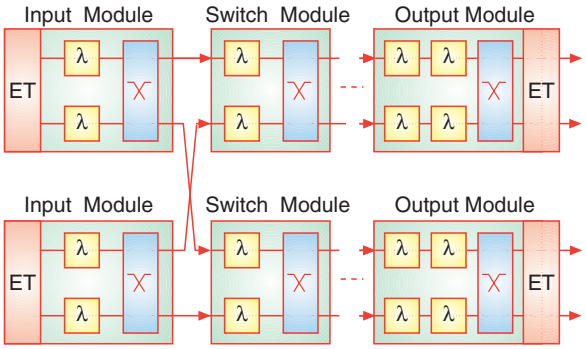


Figure 8 The multidimensional switching approach for contention resolution (ET: exchange termination, l: wavelength switching stage, T: time switching stage, X: space switching stage)

tention resolution. The present status of technology of wavelength converters and semiconductor optical gates allows for the implementation of switching modules having 16 inputs and 16 outputs. Alternative approaches, investigated within the RACE ATMOS project, have been described: it was shown that contention resolution could be solved in different ways, by storing cells in fibre loop memories or by exploiting the wavelength domain for collision avoidance in the input stages of a switch fabric.

There is no doubt that optical switching techniques still are in their infancy and that further improvement of both the technology and the switching architectures is needed prior to fully demonstrating mature photonic switching systems. However, the first experimental demonstrations reported here have given quite encouraging results which, in our opinion, suggest that photonics can play a major role in the design and engineering of future broadband switching systems.

Acknowledgements

The work described in this paper has been partly funded by the European Community, within the frame of the RACE 2039 project.

References

1 Cisneros, A, Brackett, C A. A large ATM switch based on memory switches and optical star couplers. *IEEE Journal of sel. areas on commun.,* vol 9, (8), 1991, 1348–1360.

2 Calvazara, M et al. Resolution of ATM cell contention by multiwavelength fiber loop memory. In: *Proceedings topical meeting on photonics in switching,* 25–29 Sept. 1994, Firenze, vol 2, 567–570.

3 Blumenthal, D J, Prucnal, P R, Sauer, J R. Photonic packet switches : architectures and experimental implementations. *Proceedings of the IEEE,* vol 82, (11), Nov. 1994, 1650–1667.

4 Gabriagues, J M, Jacob, J B. Exploitation of the wavelength domain for photonic switching in the IBCN. *Proceedings European conference on optical communications ECOC'91,* Sept. 1991, 59–66.

5 Gabriagues, J M et al. Design, modelling and implementation of the ATMOS project fibre delay line photonic switching matrix. *Opt. & quant. electron.,* vol 26, 1994, 497–516.

6 Gabriagues, J M, Jacob, J B. OASIS, a high-speed photonic ATM switch : results and perspectives. *Proceedings XV international symposium on switching ISS'95,* 23–28 April 1995, Berlin.

7 Stubkjær, K E et al. Optical wavelength converters. In: *Proceedings European conference on optical communications ECOC'94,* 25–29 Sept. 1994, Firenze, vol 2, 635–642.

8 Grard, E et al. Polarization independent compact semiconductor amplifier module with high gain and low consumption. *Tech. digest European fibre optic communications EFOC/LAN'92,* Paris, June 1992, 18.

9 Leclerc, D et al. High performance semiconductor amplifier array for self-aligned packaging using Si V-groove flip-chip technique. In: *Proceedings European conference on optical communications ECOC'94,* 25–29 Sept. 1994, Firenze, vol 4, 117–120.

10 Wünstel, K et al. Multidimensional optical switching with advanced key components. In: *Proceedings European conference on optical communications ECOC'93,* 12–16 Sept. 1993, Montreux, vol 2, 89–92.