

MILITARY TECHNICAL COLLEGE
CAIRO-EGYPT



FIRST INTERNATIONAL CONF. ON
ELECTRICAL ENGINEERING

MEMORY SIZE REDUCTION IN LARGE SCALE PRIVATE MULTICAST ATM SWITCH

Dr. Amani S. Amin* **Eng. Hanaa A. Ibrahim****

ABSTRACT

In this paper, a new architecture for a multicast large scale private multicast ATM switch is proposed. The proposed architecture is a two stage switch, which can be easily expanded from few tens to hundreds of input ports. The number of ports which must be served through the switch are divided into separate partitions. This results in reducing the memory capacity which required in the trunk number translators of the copy network modules. The memory reduction in the trunk number translator, is evaluated at the end of this paper.

KEY WORDS

ATM switch, Multicast network, Private switch, Trunk number translator

NOMENCLATURE

B: Number of simultaneous multicast connections
L: Maximum number of copies
M: Maximum storage capacity of table
N: Number of input ports
n: Number of input ports for each module

I- INTRODUCTION

Multicast capability is going to play a key role in the design of ATM broadband switches. In broadband ISDN, a large number of different services can be

* Head of the Switching Dep. National Telecommunication Inst. - Nasr City - Cairo - Egypt

** Teacher Assist. - Switching Dep. National Telecommunication Inst. - Nasr City - Cairo - Egypt

provided. In addition to the usual voice or data information transfer, typically point to point services, the same information can be simultaneously transferred from a single source to many users such as, i.e. point to multipoint services. The growing demand for point to multipoint services makes the introduction of multicast capability in ATM switching nodes, designed to operate in such networks, more and more essential.

Two approaches have been used to implement the multicast capability in ATM switches. i.e., the time and space approaches. In the time approach copies of a multicast cell are generated one by one using each input processor of the multicast switch [1]. This approach can lead to serious delay and heavy congestion if the numbers of copies requested by a multicast cell are large.

In the space approach, cell replication is accomplished by a copy network preceding a point to point switch, as indicated in fig. (1) two examples of this structure are proposed by Turner and Lee in Ref. [2] and [3] respectively. The copy network uses the parallel processing technique in generating the required number of copies. As a result, the cell delay can be reduced.

The main multicast cell entering the copy network is called the master cell. The header of the master cell contains two major values, the multicast channel number (MCN) and the number of copies (NC). The function of the copy network is to replicate the master cells from various sources concurrently and to provide the copy index (CI) labels to the copies. As the copy cells emerge from the output of the copy network, a table look up is performed, with reference to the MCN and CI fields of each copy cell at, a functional unit called the trunk number translator (TNT). TNT is used to identify the appropriate output address of each copy cell.

Most of the structures reported earlier cannot be adapted for very large multicast ATM switches. This is because the storage size required in the TNT's of these structures would become unwieldy as the size of a multicast switches grows [4, 5].

In this paper, a proposed architecture for nonblocking large scale multicast ATM switch is proposed. The storage size which is used in the multicast translation

tables can be reduced in the proposed architecture. The proposed structure uses the property that broadband Banyan network is nonblocking if the active inputs are cyclically concentrated and the corresponding output sets are monotone [3]. Section II, discusses the multicasting function in large scale private switches. In section III, the structure of the proposed large scale private ATM switch is presented. Finally, the memory size reduction of the proposed system is evaluated in section IV.

II- MULTICASTING IN PRIVATE LARGE SCALE SWITCHES:

While ATM is a definition of public network, it can also be used within private networking [6]. The switching node of the public network is the public ATM switch, while the switching node of the private network is a private ATM switch. The private ATM switch can be connected to the public ATM switch through public user network interface, the private ATM switch can be interconnected to the public ATM switch as a type of customer premises equipment.

To realize a public broadband network the required capacity of the public ATM switch must be not less than 10,000 high speed ports, while the required capacity of the private ATM switch may be in the range of hundreds of ATM lines.

Most of the proposed ATM switches can support only a limited number of ATM lines, i.e. a switch fabric size of about 400 high speed ports. Therefore, most of the proposed switches are private large scale ATM switches. For the private ATM switches, the switch modules can be organized in two vertical stages such that a system capacity ranging from one to nine hundreds of lines can be realized. To construct a large-scale switch in the range of thousands of lines, at least three vertical stages are required [7]. The rest of this section is dedicated for reviewing some of the large scale multicast switches (private switches).

Recently Zhong [5] has proposed a new copy network to be used in a modular structure of large scale multicast switches (Fig 2). With this copy network, the number of copies is not limited by the size of the copy network. So such a copy network can be used with large routing networks in a modular fashion. In order to reduce the size of the memory in TNT's, a connection is uniquely assigned to a copy network during the connection setup phase. Thus the TNT's at the outputs of a given copy network module just contain the table lookup data corresponding

only to the connections assigned to the module. However, because of the connection assignment blocking can occur.

The blocking problem is resolved in multipath self routing (MPSR) switching network [8]. In this three stage architecture, shown in fig 3, cells are evenly distributed to all the middle stage switches. However since the middle stage must simultaneously handle all the connections established between any input and any output, the memory size required in the TNT's at the output of the middle stage becomes very large. At the third stage, a connection level routing is used. In other words, a connection is uniquely assigned to a certain copy network at the connection setup phase. In this case, a copy module in the third stage has to handle the only connections routed to it, hence reducing the TNT's table size. In this architecture, two sets of trunk number translators are used. Each TNT of the first set must contain all the multicast routing information of all the system. As a result, the overall required storage size is very large and it is difficult to be used as a large scale multicast ATM switch.

An architecture for large scale multicast ATM switching is proposed in [9]. The architecture solves the blocking problem of Zhong and also reduces the storage size. As indicated in fig [4], it consists of four stages: connection assignment stage (CA), primary copy stage (PC), routing stage (R) and multicasting stage (MC). Cell replication is performed at the PC and MC stages. Thus two sets of TNT's are used. At the CA stage, each incoming connection is uniquely assigned to a certain copy network of the PC stage at the connection setup time. As a result, all the cells belonging to this connection will be routed to the selected PC (connection level routing). By using this routing technique, each TNT at the PC stage is assigned to a set of the multicast connections. As a result, its storage size can be reduced.

By using this connection level routing, blocking for any new connection may occur. The blocking probability is reduced by increasing the number of PC modules such that the number of alternative paths can be increased. By increasing the number of PC modules in the middle stage, the resulting system hardware cost and complexity increases compared to the other systems.

III- STRUCTURE OF THE PROPOSED PRIVATE MULTICAST LARGE SCALE ATM SWITCH:

In this section, a new architecture of a private multicast large scale ATM switch is proposed. The proposed architecture is designed in a way such that the storage size which is required for the multicast routing information is reduced. Fig [5] indicates the structure of the proposed switch. It is composed of two stages of switch module; the input switch modules and the output switch modules. The number of ports which must be served through the switch fabric is divided into n groups. Each group is connected to one input and one output switch module, such that it can transmit and receive traffic through the switch fabric simultaneously. The dimensions of each switch module (input switch module and output switch module) are $n \times n$. Each first stage switch module is connected to all the second stage switches such that any ATM input line can transmit traffic to any ATM output line. The first stage switch module is composed of a copy network, a group of trunk number translators and a routing network. The copy network is used to generate the required number of copies for each incoming multicast cell. The routing tag for each copy of the multicast cell is added through the TNT's. Finally, the routing network is used to organize the copied cells such that each copy can be routed to the appropriate output switch module on which the requested destination is connected. The output switch module is used to gather the incoming cells from all the first stage switch modules, then to route them to the appropriate ATM output ports according to the routing tag.

The advantage of this configuration is that the storage size which is required for the multicast routing information is smaller than all the other presented configurations. This reduction of the storage size results from the technique of dividing the number of ATM ports into partitions. Then each partition is connected to a specified copy network. As a result, each TNT must contain the multicast routing information which is needed for that partition only. Since the multicast routing information for each switch module is completely separated from each other, then the multicast channel number (MCN) value is unique per switch module and the same value can be repeated in other switch modules without conflict in the multicast routing information. As a result, the number of bits which is used to represent the MCN value are less than that used in the other techniques as will be indicated in the next section.

The overall storage size which is required for the multicast handling is also reduced in this configuration. This reduction results from the fact that only one set of TNT's is needed at the outputs of the copy networks existing in the first stage.

IV- EVALUATION OF THE MEMORY REDUCTION IN THE PROPOSED SYSTEM

In this section, the reduction of the storage size, which is used for storing the multicast routing information, is evaluated. The storage size of one TNT and the overall storage size which is used for storing all the multicast routing information are calculated and compared with the other presented systems.

Suppose that the switch fabric size is $N \times N$, and the size of each copy module is $n \times n$. The maximum allowable number of multicast connections within the system is B . To calculate the storage size of each TNT, the routing information which is stored in it must be identified : Table 1 indicates the contents of the TNT.

Table 1. The trunk number translation table.

MCN	CI	RA
1	1	1100...1
	2	1010...1
	.	0101...0

	n
2	1

	n
B		

The first column contains the MCN parameter which is used to identify the cells belonging to the same multicast connection. The copy index (CI), is the second parameter which is used to identify each copy of the multicast cell from all other, such that it can take different destination address. The routing address (RA) is added to each copy of the multicast cell such that it can be directed to the final destination.

Thus the storage size of each TNT is given by multiplying the width of each entry by the number of existing entries. If the width of each entry is W , and each multicast connection can generate up to L copies, then each table requires :

$$M = BLW \text{ bits of storage} \quad (1)$$

By using this formula, the storage size of the proposed architecture can be calculated and compared with the storage size of the other presented techniques.

In Lee's multicast switch [3], all the multicast connections within the system are processed through each TNT. Thus each multicast translation table must contain all the routing information of the system. Since each multicast connection can generate up to L copies, then the storage size which is required to store the CI value is given by $\log_2 L$. The number of bits which is used to store the MCN parameter is determined according to if it is unique within the switch fabric or not.

In this architecture, the MCN value is unique within the switch fabric, thus its required storage size is $\log_2 N$. Since the fabric size is $N \times N$, then the number of bits which is required for storing the routing address is given by $\log_2 N$.

The storage size which is required for the width of each entry is:

$$W = \log_2 N + \log_2 L + \log_2 N \quad (2)$$

The number of the required memory bits which is required for one TNT is represented by :

$$M = BL [\log_2 N + \log_2 L + \log_2 N] \quad (3)$$

Since there are TNT in each output of all the system, then;

The overall storage size required for all the system is :

$$M_{Lee's} = BLN (2\log_2 N + \log_2 L) \quad (4)$$

In the Zhong's multicast switch, the number of multicast connections routed to each copy module is reduced on the average to B/n . For this architecture, the MCN is unique within the switch fabric, and the number of generated copies for each multicast connection are L . Then the required storage size for one TNT is :

$$M = \frac{B}{n} L [2\log_2 N + \log_2 L] \quad (5)$$

The required storage size for all the system is :

$$M_{zhong} = \frac{B}{n} LN (2\log_2 N + \log_2 L) \quad (6)$$

In MPSR switching network [], there are two sets of TNT's each copy module need to generate and translate up to n copies for each connection. Since all B connections pass through every copy module in the middle stage, then the first set of TNT's must process all the B connections. The memory size for each TNT in the middle stage is given by :

$$M_1 = BL [\log_2 N + \log_2 L + \log_2 n] \quad (7)$$

The second set of TNT's process only B/n connections. Thus the size of each TNT is given by :

$$M_2 = \frac{B}{n} L [\log_2 N + \log_2 L + 2\log_2 n] \quad (8)$$

The storage size for all the system is :

$$M = BLN [\log_2 N + \log_2 L + \log_2 n] + \frac{B}{n} LN [\log_2 N + \log_2 L + \log_2 n] \quad (9)$$

$$M_{MPSR} = BNL [\log_2 N + \log_2 L + \log_2 n] [1 + 1/n] \quad (10)$$

In the large scale multicast switch, which is proposed in [9], there are two sets of translation tables for the PC and MC stages.

For the PC stage, the TNT's process only B/m connections. Thus the storage size of each TNT is given by :

$$M_{PC} = \frac{B}{m} L [\log_2 N + \log_2 L + \log_2 n] \quad (11)$$

For the MC stage, each TNT process only $\frac{B}{n}$ connections, the storage size of each one is given by :

$$M_{MC} = \frac{B}{n} L [\log_2 N + \log_2 L + \log_2 n] \quad (12)$$

Thus storage size of all the system is given by :

$$M = \frac{B}{m} L m n [\log_2 N + \log_2 L + \log_2 n] + \frac{B}{n} LN [\log_2 N + \log_2 L + \log_2 n] \quad (13)$$

$$M = BL [\log_2 N + \log_2 L + \log_2 n] [n + \frac{N}{n}] \quad (14)$$

In the new proposed system, the MCN is unique per switch module. In other words, the same value can be repeated in each switch module without causing any confusion through the translation table. Thus, the storage size which is used for the MCN is less than the other presented systems.

Through the new architecture, a one set of TNT's is needed. Each TNT processes only $\frac{B}{n}$ multicast connections. The storage size for the MCN is given by $\log_2 n$.

The storage size which is required for the copy index parameter is $\log_2 L$. Since the proposed architecture including two sets of routing networks, then the storage size for the routing address is given by :

$$2 \log_2 n \quad (15)$$

Therefore, the storage size for one TNT is :given by

$$M = \frac{B}{n} L [3\log_2 n + \log_2 L] \quad (16)$$

The storage size for all the system is :

$$M_{New} = \frac{B}{n} LN [3\log_2 n + \log_2 L] \quad (17)$$

For the purpose of comparison among various architectures, assume that $N = n^2 = 1024$ And the number of allowable multicast connections within the system is $B = 2N$ and the maximum number of copies which can be generated by each copy are equal to 8. The memory size required in each TNT of the switches based on Lee, Zhong's, MPSR, Tiarawut, and the proposed structures are summarized in table 2.

The overall storage size which is required for each system is indicated in table 3. As indicated in table 2, the memory size required for one TNT in the new proposed architecture and in the Tiarawut is less than that required for the systems based on Lee, Zhong and MPSR. Table 3 indicates that the overall storage size which is required for the new architecture is less than all other systems.

Table 2 The Memory size for one TNT.

Architecture	The Required memory size for N = 1024	
	Lee's	$BL (2\log_2 N + \log_2 L)$
Zhong's	$\frac{B}{n} L (2\log_2 N + \log_2 L)$	1177.6
MPSR	$BL (\log_2 N + \log_2 L + \log_2 n)$	294814
Tiaraut	$\frac{B}{n} L (\log_2 N + \log_2 L + \log_2 n)$	921.6
New	$\frac{B}{n} L (3\log_2 n + \log_2 L)$	921.6

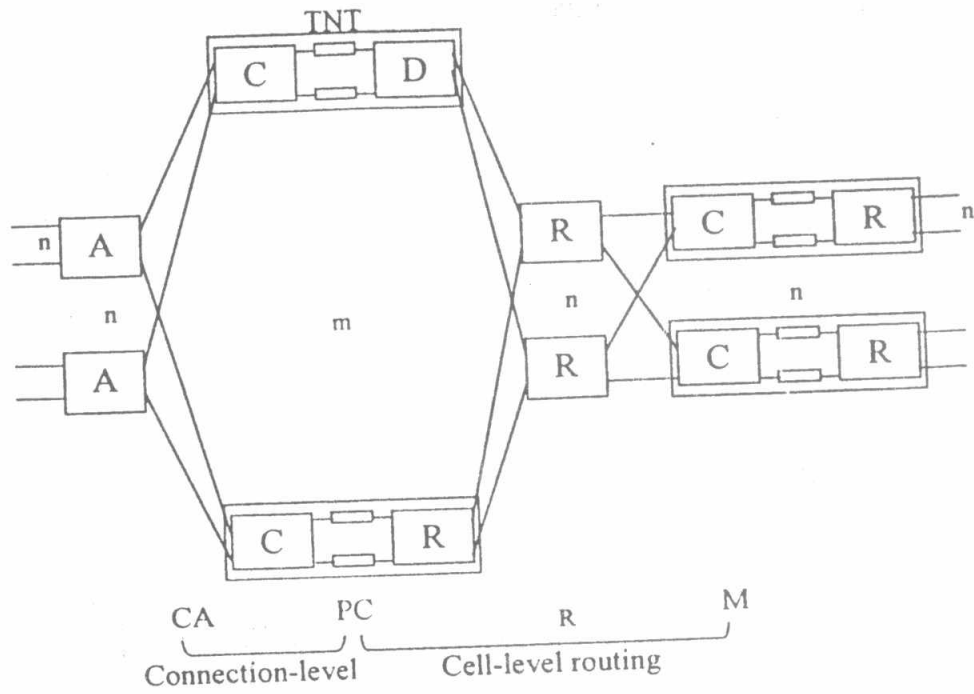
Table 3 The Memory size for all the system.

Architecture	The required memory size for N = 1024	
	Lee's	$BLN (2\log_2 N + \log_2 L)$
Zhong's	$\frac{B}{n} LN (2\log_2 N + \log_2 L)$	1.2×10^6
MPSR	$BNL (\log_2 N + \log_2 L + \log_2 n) (1 + 1/n)$	1.5×10^9
Tiaraut	$B L (\log_2 N + \log_2 L + \log_2 n) (n + \frac{N}{n})$	1.8×10^6
New	$\frac{B}{n} LN (3\log_2 n + \log_2 L)$	9.4×10^5

CONCLUSION

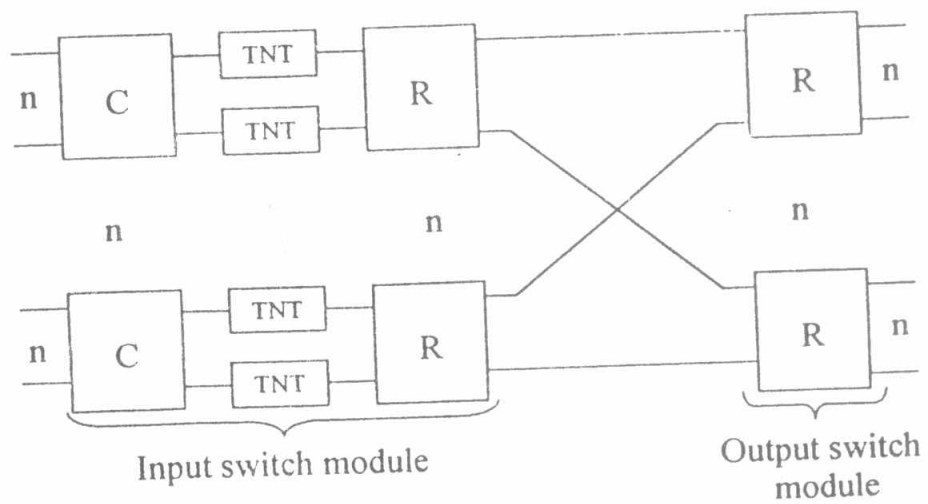
In this paper, a new architecture for a large scale private multicast ATM switch is proposed. The main aim of this architecture is reducing the storage size which is required for storing the multicast routing information.

According to the results, it is shown that the smaller size of the TNT is used by the new architecture and the architecture proposed by tiaraut.



CA : Connection Assignment R : Routing
 PC : Primary Copy MC : Multicasting

Fig. 4 large scale multicast switch with hybrid routing



C : Copy network
 R : Routing network

Fig [5] structure of the proposed large scale switch

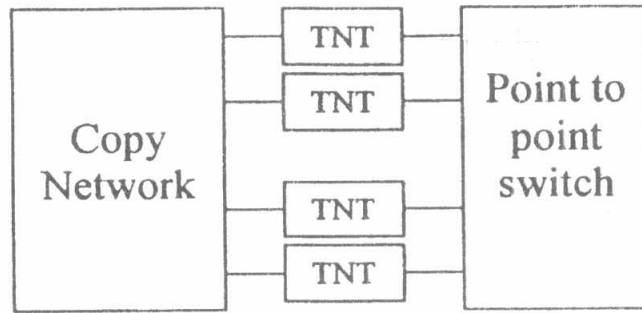
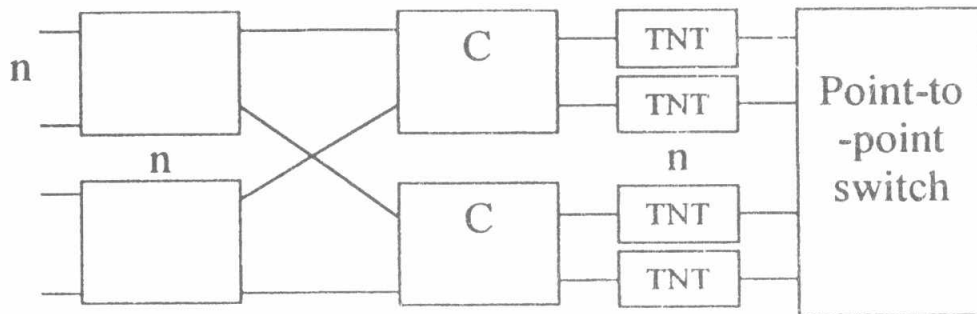


Fig 1 common structure of multicast switches



Distribution C : Copy Network

Fig. 2 Zhong's modular multicast switch structure [5]

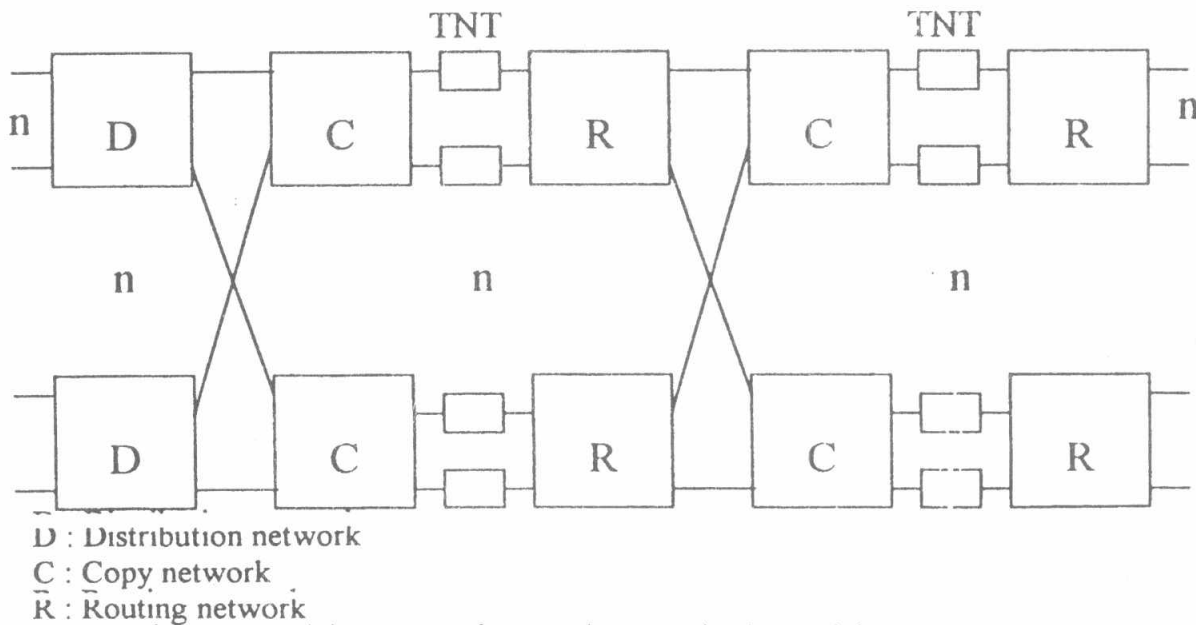


Fig. 3 Multicast self-Routing switch architecture with multicast space-division switch modules

As a result, the reduction of the overall cell delay through the switch fabric can be achieved.

But the proposed system can reduce the overall storage size required for the multicast routing information.

References

- [1] **Lea C.T.**, "A multicast broadband packet switch," IEEE Trans. Commun., vol. 41, no 4 pp. 621-630, April 1993.
- [2] **Turner J.S.**, "Design of a broadcast packet network," IEEE Trans. Commun., vol. 36, pp. 734-743, June 1988.
- [3] **Lee T.T.**, "Nonblocking copy networks for multicast packet switching," IEEE J. Select. Areas. Commun, vol. 6, no. 9, pp. 1455-1467, Des. 1988.
- [4] **Turner J.S.**, "A practical version of Lee's multicast switch architecture," IEEE Trans. Commun., vol. 41, no. 8, pp. 1166-1169, Aug. 1993.
- [5] **Zhong W.D.**, **Onozato Y.**, and **Kaniyil J.**, "A copy network with shared buffers for large-scale multicast ATM switching," IEEE /ACM Trans. Networking, vol. 1, no. 2, pp. 157-165, April 1993.
- [6] **ATM forum**, "ATM user network interface specifications", the version 2.2, June 21, 1993.
- [7] **John Bellamy**, "Digital Telephony", John Wiley & Sons, Inc. 1982.
- [8] **Henrion M.**, **Eilenberger G.**, **Petit G.**, and **Parmentier P.**, "Technology, distributed control and performance of a multipath self-routing switch." ISS'92, A5-1 Oct. 1992.
- [9] **Tiarawut S.**, **Satio T.** and **Aida H.**, "An architecture for large scale multicast ATM switching with Hybrid routing", IEICE Trans. Commun., vol E 78 - B No 10, October 1995.

