

A new architecture for ATM cell routing based on Banyan network

Carlos Antônio Ruggiero¹, Angelo Cesar Colombini²

¹ Universidade de São Paulo - USP
Instituto de Física de São Carlos
Av. Dr. Carlos Botelho, 1465 - São Carlos - CEP 13.560-300
Estado: São Paulo - Brasil

Telefone: (016) 273.9879 Universidade de São Paulo

Email: toto@if.sc.usp.br

² Universidade de São Paulo - USP
Instituto de Física de São Carlos
Universidade de Franca - UNIFRAN
Av. Dr. Armando Salles Oliveira, 201 - CxP. 082
Estado: São Paulo - Brasil

Parque Universitário - Franca - SP - CEP 14.404-600

Telefone: (016)724.2211 - Fax - (016)722.0086

Email: angelo@if.sc.usp.br

angelo@unifran.br

Keyword: Asynchronous Transfer Mode, ATM Switching, Banyan Network and Multicast

Abstract

Routing in ATM switches is very complex. The complexity increases when the applications require multicast capabilities. Both unicast and multicast traffic and also the Banyan network have only N^2 crossing points. Because of that, there is an inevitable contention for the same output link and a bottleneck point. It is important to mention that cost and speed of operation are essential requirements in ATM. The easy solution to this problem would be the buffer strategy. However, a better option is to modify the Banyan network in such a way that could keep its characteristics of auto routing, allowing a disassociation of Unicast and Multicast cell traffic inside the routing structure.

1. Introduction

Today, it is clear that the old analogic communication networks are being replaced by fast digital networks. This digital network is able to integrate many applications such as voice, video, data and images in a unique way. Many applications require point to multipoint and multipoint to multipoint connections. This is necessary to support multicast. There have been many proposals trying to design an ATM switch to efficiently route multicast traffic. Some of the proposals approach the problem with cell duplication techniques, where multicast cells are duplicated and routed as unicast cells. As an alternative,

there are some proposals that use buffer sharing output as a support to achieve efficient multicast [PHY 94]. It is relevant to mention that, based on the package replication method, the ATM switches architecture can be divided in two large groups: multicast tree and broadcast [CHA 98].

The multicast tree switches are composed of two modules. The first one is a copy network intended for replication of multicast cells. The second one is a routing network which aim to conduct the unicast cells to the desired outputs. The negative aspect of the copy network is that occupation of the first phase of the routing structure can add to the unicast traffic the copies done to achieve the multicast requests, increasing the probability of collision and the necessity of buffer usage across the routing structure. Another important aspect to consider is the restriction to the number of cells addressed to the same output. When many cells have identical output goals, contention can occur in any point of the Banyan network. Some tentative solutions for this problem have been proposed [OBA 90], [COP 93], [BYU 94], [TUR 98] e [URU 98].

The second module of the multicast tree group is called broadcast module. Each cell that enters the routing structure is sent in broadcast to other output, it assures the multicast achievement in a quietly form. One known project is MOBAS [CHA 95] which is a modified knockout switch.

The aim of this article is to propose a novel architecture for a high performance ATM switch that is able to solve the above problems. The routing structure used is highly

modular and suggests that the copy network occupies the posterior position of a routing network. Figure 1 shows a block diagram of the proposed architecture. It suggests that before the copy cells achievement destined to the multicast request, a unicast cells routing might be done. The simulation results show that the probability of cell loss can easily be reduced by the addition of a routing module number.

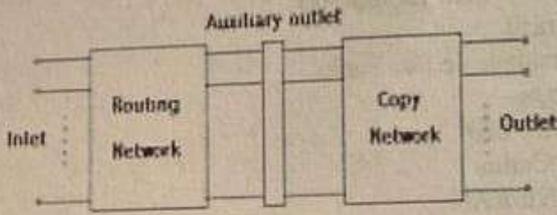


Figure 1: Block diagram of the proposed architecture.

2. Switch Architecture

The proposed architecture (figure 2) is composed by switching elements, routing links and expansion links. Cells with the same output address can follow parallel routes minimizing contention.

A Banyan network with N inputs and N outputs ($N \times N$) is defined as a set of $n = \log_2 N$ stages of switching elements (SE's). Each SE has d inputs and d output ($d \times d$), within each stage. Figure 2 shows an architecture defined by $N=8$, $d=2$, $\epsilon=2$ (expansion). Observe that at the end of the routing network and at the beginning of the copy network there is a region called *intermediate output* which provides an exit to cells that reached the correct output. The outputs correspond to the last stage of the copy network. They are followed by two alternative output stages ($\epsilon=2$).

The proposed architecture can be determined through the following equations: $n_r = \log_2 N + \epsilon/2$ (n_r is the number of stages in the routing network) when ϵ is even and $n_r = \log_2 N + (\epsilon + 1)/2$ when ϵ is odd. In the copy network, the number of stages is given by $n_c = \log_2 N + \epsilon$. The total number of stages n is given by $n = n_r + n_c$. In the first block, the cells are routed through the Banyan network as if they were unicast cells. It is important to mention that this process happens without the addition of the multicast cell copies. In this block, the multicast cells compete in each link for their output. When a contention situation occurs, the first attempt is to follow the auxiliary

path (parallel to the desired path). In this case, a cell is deviated to an alternative path and it will only fight again for its output link in the second block of the routing structure.

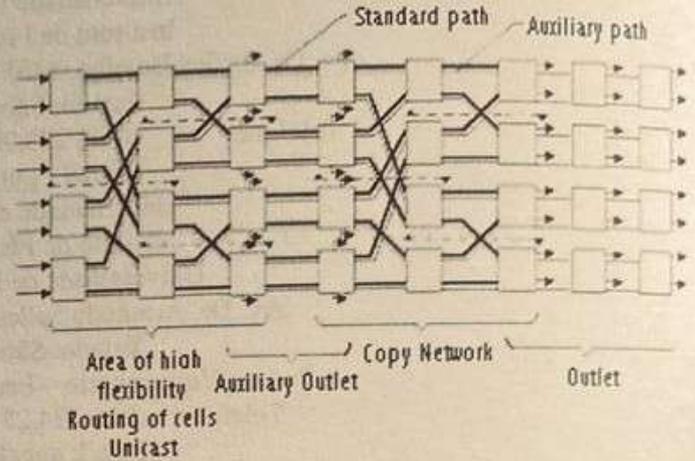


Figure 2: Diagram in architecture proposed blocks

The important thing here is the flexibility in the competition for an output link, which reduces the cell loss rate to very low levels. Another important aspect is the absence of multicast cell copies which could jam the routing structure. The second block has different characteristics compared to the first. In fact, the multicast cell copies take place here. In the second block the traffic is reduced because most unicast cells have been delivered in the first block. Here, the collisions are dealt with differently. In this case, the only option is the auxiliary path. This block is called the *convergent block*, because the cells have to converge to their output links. If this is not possible, they must be thrown away. The proposed structure is totally modular and does not use any buffers. The consequences are higher performance and lower costs.

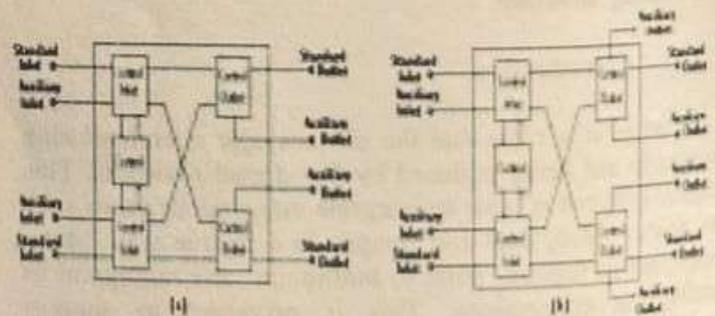


Figure 3: Diagram representation of standard routing module blocks (a) and module with the auxiliary output (b)

This work did not take into consideration the switch output module. Because the cells leave the routing structure at different points, the output ordering can be wrong. This problem has to be considered in the output module.

Figure 3 shows the *standard module blocks* and the *auxiliary output module blocks* used in cell routing.

3. Simulator Description

Figure 4 presents the main classes used in the simulator design and the form they are related to each other. The C_STATUS, C_CELULA and C_ATM classes deal with the ATM cell; the class GERAENTRADA is responsible for the production of input data and the classes MODULO_HARDWARE, MODULO_HARD_DEF_ENTRA_SAI and MODULO_HARD_MONTA_TAB_ROT represent the hardware internal structure. Finally, the TAB_ROTAMENTO class brings together the three classes mentioned previously to form the proposed routing structure.

3.1) The ATM cell

The ATM cell is represented by the C_STATUS, C_CELULA and C_ATM classes, as shown in figure 5. The GERAENTRADA class is responsible for the production of input data. The relationship between these classes is described below:

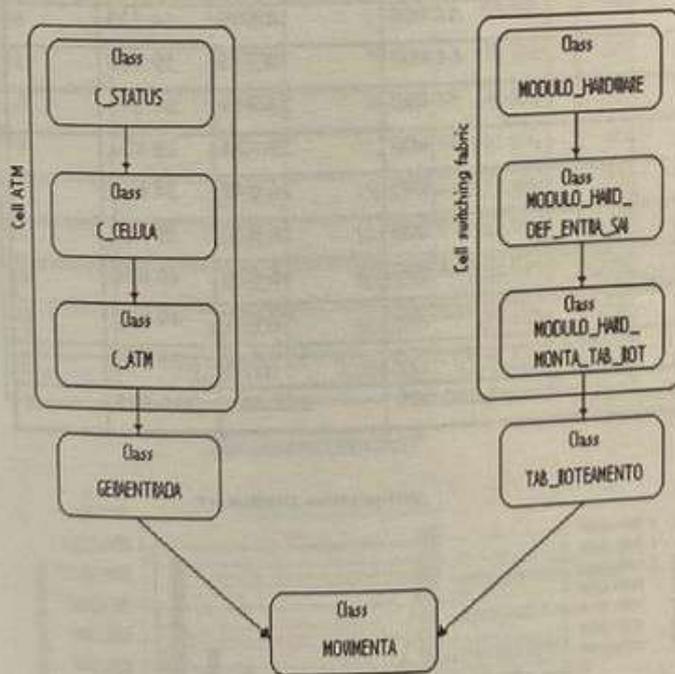


Figure 4: Diagram of simulator blocks

- a) C_STATUS: is responsible for assembling the ATM cell header and also for connection control.
- b) C_CELULA: together with the class C_STATUS, assembles the ATM cell header.
- c) CELULA_ATM: is derived from the C_STATUS and C_CELULA classes. It is responsible for the ATM cell final composition and for the controlling devices used in the simulation.
- d) GERAENTRADA: brings together all the previous classes and is responsible for the generation of input data.



Figure 5: Representation of the ATM cells used for simulation tests

3.2. Hardware

The hardware that forms the routing structure is represented by the MODULO_HARDWARE, MODULO_HARD_DEF_ENTRA_SAI and MODULO_HARD_MONTA_TAB_ROT classes. The function of the TAB_ROTAMENTO class is to implement the physical structure of hardware. All the classes mentioned previously are brought together through this class. The relationship between the classes is described below:

- a) MODULO_HARDWARE class: defines the external aspect of each module inside the routing structure.
- b) MODULO_HARD_DEF_ENTRA_SAI class: through the position and the number of inputs/outputs

established in the MODULO HARDWARE class, defines who these outputs are. Figure 6 shows the switching module, seen as a black box, and emphasizes the inputs and the standard and auxiliary output.

c) Class MODULO_HARD_MONTA_TAB_ROT: this sets up a table 4x5 (internal routing table), in which the possibilities of available routes for a cell inside the switching module are mapped.

d) TAB_ROTAMENTO class: this class brings together all the previous ones. It represents the routing structure which can be seen in figure 2.

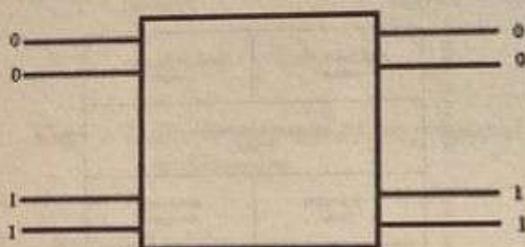


Figure 6: Visualisation of the path duplicity in the commutation module

	E 1 SAIDA_0	E 2 SAIDA_01	E 3-1 SAIDA_1	E 3-1 SAIDA_01	B DESCARTES
E 1 ENTRADA_0					
E 2 ENTRADA_01					
E 3-1 ENTRADA_1					
E 3-1 ENTRADA_01					

Figure 7: Routing table, internal to each object of MODULO_HARD_DEF_ENTRA_SA.

3.3. The cell movements in the routing structure

Cell movement and result collection along the routing structure is MOVIMENTA class responsibility. This class integrates the two class hierarchies described previously, composing the simulator parts.

4. Simulation and result analysis

The simulation has been divided in two parts. The first accepts only the unicast requests and the second deals with unicast and multicast requests. Each simulation unit is composed by 11 files with a total of 1.280.000 ATM cells. The routing structure used in the simulation has 16 input/output ports and $E = 2$. The simulation results are presented through the use of tables and graphics.

4.1. Simulation of unicast traffic : in this simulation all the cells that go through the routing structure are unicast cells. There was no control data produced, except for the conversion from multicast to unicast. The lost cells in this simulation come from the excessive number of cells with the same destination output in the same cycle.

In the second unicast simulation group a condition was imposed. All the 16 routing structure inputs were submitted to a unicast request.

The simulation aim is to submit the routing structure to extreme conditions. Therefore, all the inputs and outputs receive cells all the time. Also, there were no restrictions on the destination output. This allowed a high number of cells to be destined to the same output in the same cycle. An increase in the collision number was caused by contention for the same link inside the routing structure.

Table 1: Unicast simulation: in this simulation all the multicast requests were replaced by unicast requests.

Simulation	Generated cells	Useful cells	Success	Losses
1	16 000	4 979	4 979	0
2	32 000	9 865	9 865	0
3	48 000	14 809	14 809	0
4	64 000	19 771	19 770	1
5	80 000	24 749	24 749	0
6	96 000	29 624	29 624	0
7	112 000	34 648	34 648	0
8	128 000	39 470	39 469	1
9	144 000	44 456	44 456	0
10	160 000	49 452	49 452	0
11	400 000	123 561	123 561	0
Total	1.280.000	395.384	395.382	2

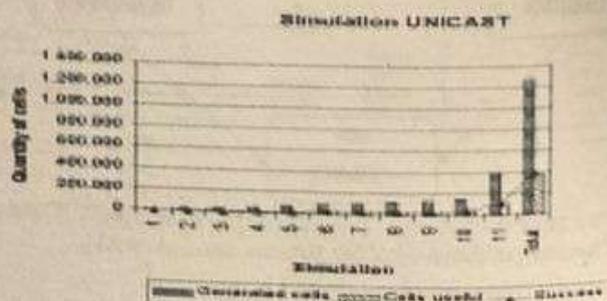


Figure 8: No restriction on input data in the unicast simulation

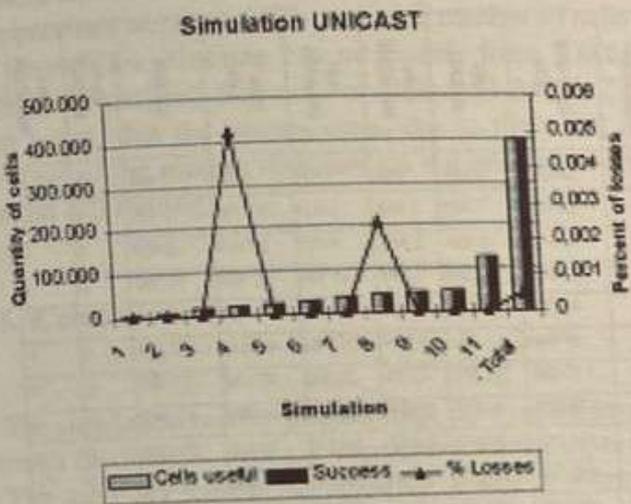


Figure 9: Lost cells related to the total number of useful cells

Table 2: Unicast simulation considering a constant traffic in the 16 switch ports

Simulation	Generated cells	Useful cells	Success	Losses
1	16.000	16.000	15.976	24
2	32.000	32.000	31.975	25
3	48.000	48.000	47.952	48
4	64.000	64.000	63.925	75
5	80.000	80.000	79.950	50
6	96.000	96.000	95.927	73
7	112.000	112.000	111.901	99
8	128.000	128.000	127.927	73
9	144.000	144.000	143.755	245
10	160.000	160.000	159.777	223
11	400.000	400.000	399.628	372
Total	1.280.000	1.280.000	1.278.693	1.307

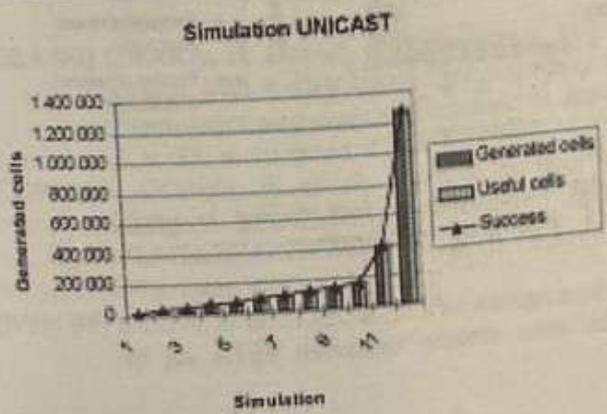


Figure 10: Constant unicast traffic simulation in all switch ports.

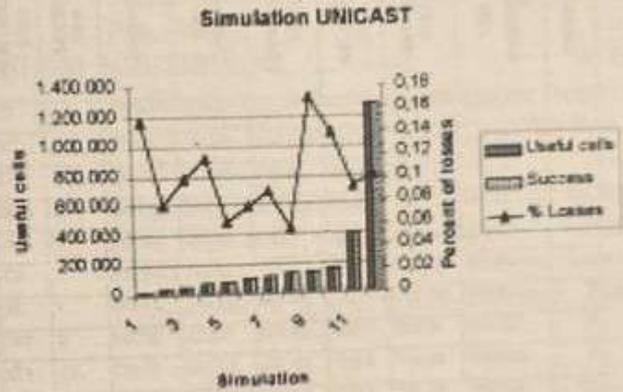


Figure 11: Percentage of lost cells related to the quantity of useful cells

Although the amount of lost cells has increased from 0,0005% to 0,1021%, the reason for this was the same: excessive number of cells with the same destination output in the same cycle.

4.2 Multicast traffic simulation: in this simulation, cells which go through the routing structure the follow unicast and multicast requests. There is no previous control of the generated data. This allowed multiple requests to be serviced in the same cycle. It is interesting to observe the high amount of lost cells caused by the excessive increase of cells going through the second part of the routing structure (conversion region). The high number of cells is consequence of the high quantity of multicast cell copies produced.

Despite the increase in the amount of lost cells, it was not a direct consequence of multicast requests. The real cause was the excessive quantity of multicast request in the same cycle. In this simulation worst case, 8 out of 12 active inputs turned out to be multicast requests.

In the second multicast simulation, the condition imposed was the acceptance of only one multicast request per cycle. Another condition was that each valid simulation cycle should contain one multicast request. There was no restriction on output ports. It is interesting to note the increase of the number of cells going through the routing structure and the unicast and multicast collision reduction. This characterizes the unicast/multicast traffic separation.

Tabela 3: Multicast simulation with no restrictions per cycle

Simulation	Generated cells	Cells UNICAST	Cells MULTICAST	Cells daughters	Total of cells	Success	Lost UNICAST	Lost MULTICAST
1	16.000	4.432	465	3.429	6.322	6.319	3	12
2	32.000	8.887	955	2.907	12.741	12.712	1	28
3	48.000	13.452	1.395	4.231	19.078	19.045	1	72
4	64.000	17.829	1.885	3.652	25.347	25.303	2	60
5	80.000	21.628	2.989	3.995	33.612	33.461	7	144
6	96.000	26.812	2.831	8.477	38.120	38.051	2	67
7	112.000	31.339	3.299	9.294	44.362	44.273	3	36
8	128.000	35.748	3.760	11.344	50.832	50.723	8	141
9	144.000	40.331	4.205	12.702	57.238	57.055	11	172
10	160.000	44.597	4.882	14.058	63.487	63.294	7	146
11	490.000	111.778	11.707	34.960	158.445	158.024	13	390
Total	1.280.000	356.934	38.173	114.545	509.612	508.270	56	1.286

Simulation MULTICAST

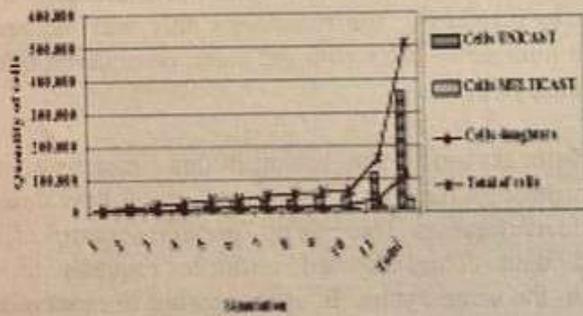


Figure 12: Multicast simulation with restriction

Table 4: Multicast simulation imposing that just one multicast request can be present per cycle

Simulation	Generated cells	Cells UNICAST	Cells MULTICAST	Cells daughters	Total of cells	Success	Lost UNICAST	Lost MULTICAST
1	16.000	3.867	550	3.027	7.444	7.436	0	8
2	32.000	7.728	1.394	6.054	15.736	15.736	0	8
3	48.000	11.458	2.998	8.945	23.428	23.428	0	8
4	64.000	15.399	3.993	11.938	31.590	31.590	0	8
5	80.000	19.455	4.983	14.938	39.412	39.412	0	8
6	96.000	23.281	5.986	17.952	47.219	47.219	0	8
7	112.000	27.179	6.974	21.042	55.194	55.194	0	8
8	128.000	30.958	7.979	23.969	62.897	62.897	0	8
9	144.000	34.942	8.972	26.874	70.329	70.329	0	8
10	160.000	38.732	9.974	29.872	78.678	78.678	0	8
11	480.000	96.793	24.921	74.923	196.637	196.635	2	8
Total	1.280.000	356.931	39.351	239.673	629.919	629.017	2	8

Simulation MULTICAST

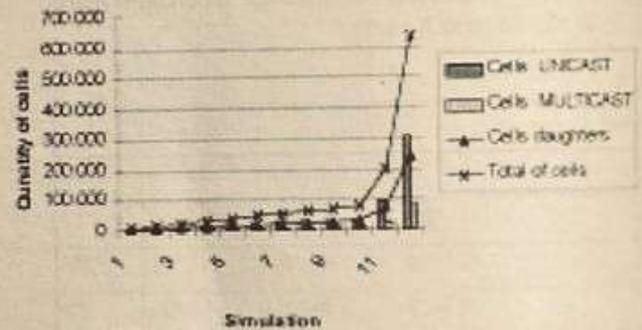


Figure 14: Multicast simulation servicing one request for cycle

Simulation MULTICAST

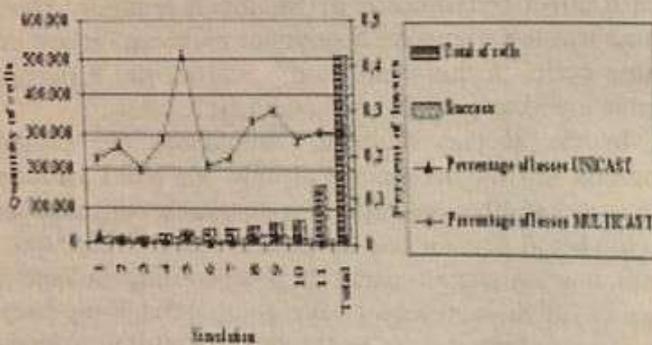


Figure 13: Percentage of lost cells related to the total number of useful cells.

Simulation MULTICAST

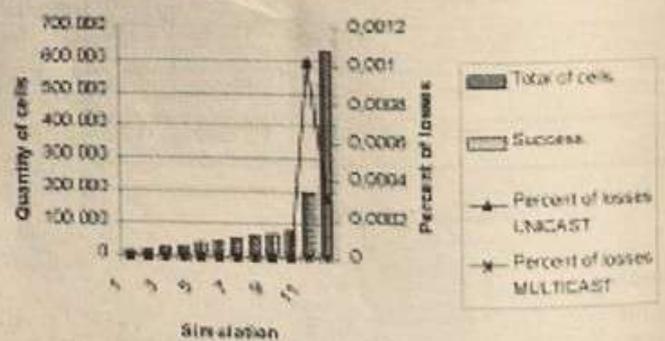


Figure 15: Percentage of loss related to total number of useful cells.

These simulation results proves the conclusions drawn in the previous simulation. Though the number of cells that goes through the structure has increased from 509.612 to 629.019, the total cell loss decreased from 1.342 to 2 cells. It was noted that the unicast cells, due to their exit in the first stage of the routing structure do not interfere with the multicast cell routing.

5. Conclusion

The simulation results indicate that the routing structure, although flexible, is strongly dependent on the input data.

The auxiliary output in columns 3 and 4 allowed an intensive drainage of unicast cells. The copy of multicast cells in the second stage of the routing structure warranted a reduced lost cell rate, with almost no effect on unicast traffic.

The random generator used in the input data production does not reflect the traffic to which the ATM switch is submitted. In addition, the generation of output ports and cell types seemed too biased. In a way, this was good because the routing structure was tested in extreme conditions.

One interesting point to consider in the routing structure is that it is built with simple elements. The absence of buffers, besides reducing the costs, also decrease the management and control complexity increasing the throughput.

Thus, the results show a robust structure that can reduce the lost cell rate to values close to zero.

6. References

- [PRY 94] PRYCKER, M. de *Asynchronous transfer mode solution for broadband ISDN*. 2. Ed. New York: Ellis Horwood, 1994.
- [CHA 98] CHAN, S. K. et al. Clos-Knockout: a large-scale modular multicast ATM switch, *IEICE Trans. Commun.*, v. E81-B, n. 2, p. 120-137, Feb., 1998.
- [OBA 90] OBARA, H. Design of a multistage self-routing switch with a distributed cell sequence control, *Electronics and Communications in Japan*, part 1, v. 73, n. 10, p. 14-27, Sep., 1990.
- [COP 93] COPPO, P. et al. Optimal cost/performance design of ATM switches, *IEEE/ACM Transactions on Networking*, v. 1, n. 5, p. 566-575, Oct., 1993.
- [BYN 94] BYUN, J. W., Lee, T. T. The design and analysis of an ATM multicast switch with adaptative traffic controller, *IEEE/ACM Transactions on Networking*, v. 2, n. 3, p. 288-298, Jun., 1994.
- [TUR 98] TURNER, J., Yamanaka, N. Architectural choices in large scale ATM switches, *IEICE Trans. Commun.*, v. E81-B, n. 2, p. 120-137, Feb., 1998.
- [URU 98] URUSHIDANI, S., et al. A high-performance multicast switch and its feasibility study, *IEICE Trans. Commun.*, v. E81-B, n. 2, p. 284-296, Feb., 1998.
- [CHA 95] CHAO, H. J., Choe, B. S. Design and analysis of a large-scale multicast output buffered ATM switch, *IEEE/ACM Transactions on Networking*, v. 3, n. 2, p. 126-138, Apr., 1995.