Optimization of Large-Scale Growable ATM Switching Fabrics[1]

Andrzej Jajszczyk and **Michal Roszkiewicz**[2]

Large switching fabrics are frequently obtained by arranging elementary switches in various ways. Growability of such structures is of the great importance. This paper presents results of our investigations on how switching elements should be initially arranged to minimize the fabric cost expressed by the number of elementary switches, and further how does the optimum structure behave in the growth process. In our investigations we have used the dynamic programming method, which is based on Bellman's optimality principle. We assumed the Clos-like architecture of the switching fabric with some modifications which are necessary due to the differences between arrangements of circuit-switched and ATM fabrics. As the result, the comparison of the cost of ATM switching fabrics composed of typical 8×8 and 16×16 elementary switches is given.

1   Introduction

Asynchronous Transfer Mode (ATM) has been adopted by CCITT as a worldwide standard for the broadband integrated services digital network. Among various ATM switching architectures described in the literature, the shared-memory architecture is one of most promising and has been implemented in several advanced field trials.

The maximum size of a single shared-buffer switch is limited by technology constraints. To obtain larger sizes more switches can be connected together in various ways. Many characteristics should be considered while designing or comparing various multi-switch fabrics. To select an optimum fabric structure it is necessary to find a compromise solution depending on a particular application, accessible technology, etc. The most important characteristics include: cost of switching elements, cost of control related to the fabric structure, delay of signals, size range, expandability, reliability, testability, traffic characteristics, mass producibility. In this paper we will investigate how the structure of the fabric has to be arranged to minimize the number of used ATM switching elements and to allow smooth fabric growability.

In Section 2 we present the considered fabric architecture and in Section 3 we discuss important implementation issues that influence the cost. Section 4 presents a discrete optimization method of growable switching fabrics based on the dynamic programming approach. Some numerical examples are shown and discussed in Section 5.

97

2 Switching Fabric Architectures

The conceptually simplest method to obtain a switching fabric of the size larger than that of a single switch is based on parallel connections between switches, as shown in Fig. 1. The number of individual switches in such an arrangement is $\lceil L/n \rceil^2$, where L is the total number of fabric links and n is the number of used ports per each side of a single switch. Implementation of branching points denoted by dots in Fig. 1 will be discussed in Section 3.
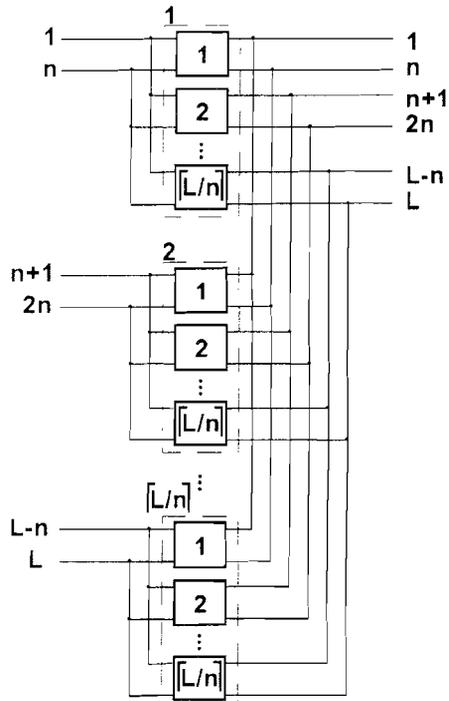
Figure 1 - Single-stage fabric

98
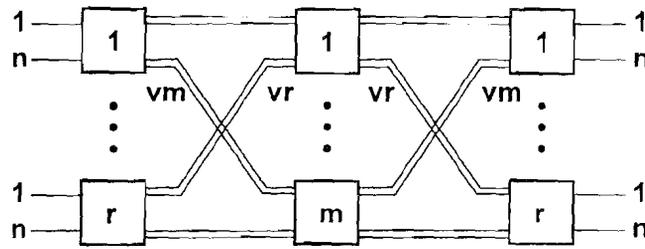
Figure 2 - Three-stage fabric
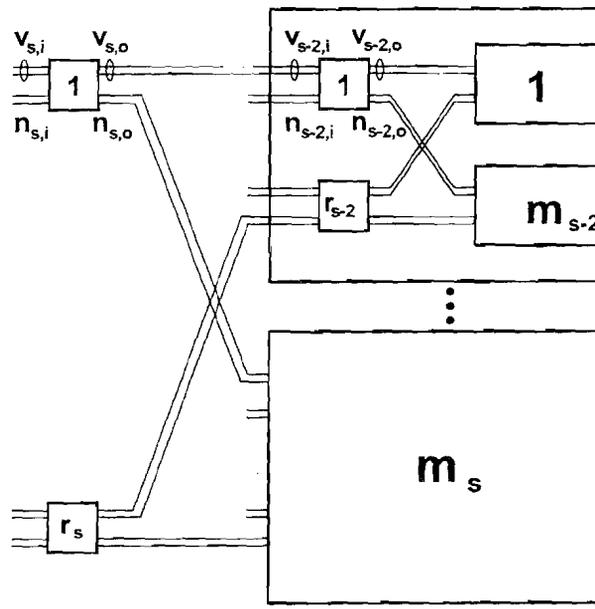
Figure 3 - Generic s-stage fabric

The rapid growth of the number of required switching elements in single-stage solutions as well as problems related to their interconnections, lead to multistage switching fabrics. Most common structures contain two or three stages composed of shared-memory switches. An example of a three-stage fabric is shown in Fig. 2. To reduce the probability of blocking, the number of parallel lines (v) between switches located in consecutive stages can be increased. Another possibility is to increase the ratio vm/n or to use the speed up of internal links.

To obtain larger fabric sizes, the middle-stage switches can be replaced by complete three-stage fabrics. In general, an s-stage fabric (see Fig. 3) is obtained by replacing each middle-stage switch in an (s-2)-stage fabric by fabrics of Fig. 2. The maximum size of such "regular" fabrics is closely related to the number of stages for a given switch size. Since the elementary switches are usually manufactured using an integrated technology, their number of ports is fixed. To obtain required sizes of the switching elements we can use a technique based on combining a number of elementary switches, as shown in Fig. 4 [1],[2]. For simplicity of presentation we have assumed that the multiplicity of the interstage links (v) is equal to one. The approach of Fig. 4a has been used in some ATM studies [3],[4],[5]. Both techniques of Fig. 4 can be applied to a switching fabric at the same time, also for fabrics containing more than three stages. We can note that by using multielement modules we reduce the number of required stages, thus reducing signal delay and simplifying path search procedures.
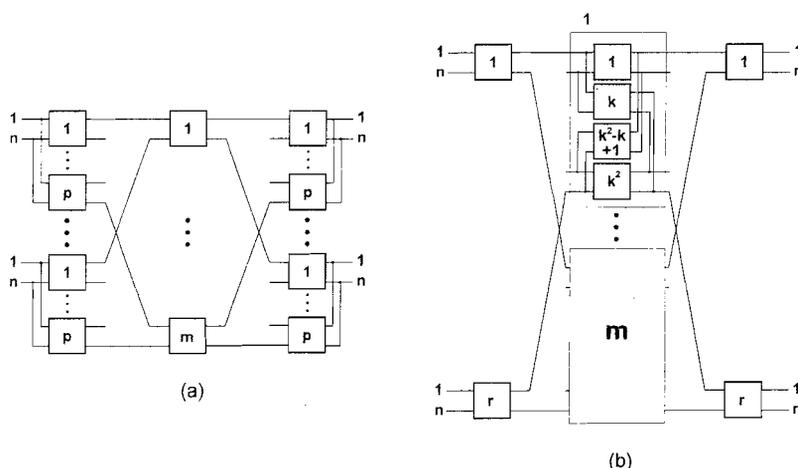
Figure 4 - Modified three-stage fabric

One of the important requirements for any ATM switching fabric is its growability. The fabric architecture should permit the modular growth of its size from a small number of ports to very large switch sizes. It is also desired that this growth does not result in a performance degradation as well as the reliability requirements are met at each stage of the expansion. A new, growable ATM architecture which makes it possible to expand the size of two-sided fabrics in a smooth way, without any changes of the existing cables, was proposed in [2].

The basic fabric, the size of which is then expanded, is formed by an s-stage fabric having L ATM lines on each side. This fabric can have the structure either that of Fig.

2, Fig. 4a, Fig. 4b, or a combination of these structures (including fabrics with $s > 3$). To expand the number of lines two times we add a second identical fabric to the existing basic structure and connect them by additional modules, as shown in Fig. 5. These modules have the same structure as the middle s_2-stage part of the basic fabric.

To increase the size of the fabric by additional L lines we add the next $L \times L$ fabric and connect it to the existing structure by using the s_2-stage connecting modules. In general, by expanding the size of the fabric k times we obtain the structure, which contains k identical basic fabrics and $2\binom{k}{2} = k(k - 1)$ modules interconnecting these basic fabrics. To increase the fabric's size from $(k - 1)L$ to kL ATM lines one basic $L \times L$ fabric and $2(k - 1)$ intermediate modules are added. Note that each line in Fig. 5 represents a group of lines incoming to or outgoing from switching modules.
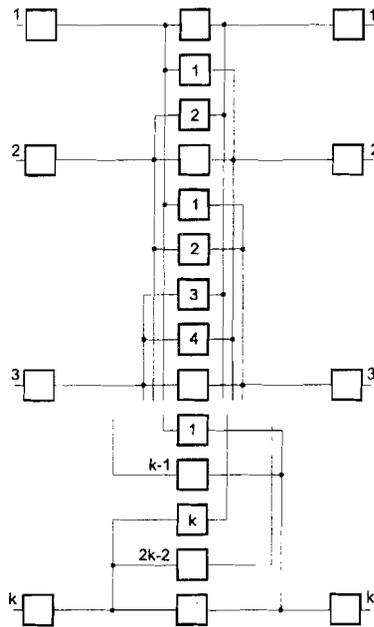
Figure 5 - Principle of the fabric growth

We should note that a cell crosses the same number of switching elements independently of the fabric's size. Since such a cell always "sees" each path as a path in a basic fabric, the fabric growth does not have an important impact on the throughput. We can also see that the throughput of the fabric can be increased at any stage of its growth by adding switching elements in the first and the third stages (see Fig. 4a) and increasing the number of switching elements in the middle-stage modules, accordingly. Such a procedure does not involve any disruptions of the existing cabling, either.

3 Implementation Issues

One of the important implementation issues associated with the discussed fabrics concerns branching points denoted by dots in Figs 1 and 4. Such points, along with complicated wiring and high-speed ATM signals, may lead to great technological difficulties due to electrical signal reflections. In some systems, simple passive branching at the splitting points gives satisfactory results [6]. In the case of passive branching, the switching elements should have the capability of accepting only cells with the appropriate values in the self-routing headers. Another option is to apply demultiplexers and multiplexers in splitting and combining points, respectively. Demultiplexers can be controlled by bits in a cell header.

More complex functions have to be performed by multiplexers. These functions can include the prevention of the internal congestion when two or more cells compete for the access to the same multiplexer output. This can be accomplished by appropriate arbitration and buffering. The multiplexers can be kept simple if an appropriate mechanism is provided between switching elements in consecutive stages. Kozaki et al. used standard 32×32 ATM switching elements as 16 two input-one output multiplexers [4]. In another solution, the function of the input and the output branching points (as those shown in Fig. 4a) was performed by header translators [5].

4 Optimization

The synthesis problem of selecting a network having the minimum cost can be solved using discrete optimization. Here, we adapt the method proposed in [7] for circuit-switched fabrics. Because of the iterative structure of multistage networks, optimization will be treated as a multistage decision problem and will be solved by using the dynamic programming method. Dynamic programming is based on Bellman's optimality principle: an optimal decision policy has the property that whatever the initial state of the system and the initial decision are, the rest of the decisions must form an optimal policy according to the state produced by the first decision.

A simplified s-stage regular fabric (see Fig. 3) with $v_{k,i} = v_{k,o} = 1$ ($k = 3,5, ..,s$) will be used to illustrate the application of dynamic programming to optimization of ATM switching fabrics. We also assume in this example that there is no expansion and no concentration at any switching stage, that is $n_{k,i} = n_{k,o} = n_k$ for $k = 3,5,...,s$, and that the fabric contains L ATM links at each side. An elementary ATM switch contains n incoming and n outgoing ports. Some ports may not be used, so n represents only the upper limit. The optimization problem may be formulated as follows:

$$C = \min\left\{2r_s + n_s\left[2r_{s-2} + n_{s-2}\left[...[2r_3 + n_3]...\right]\right]\right\} \tag{1}$$

where C is the cost of the fabric expressed by the number of elementary ATM switches.

$n_s r_s \geq L$;

$n_k r_k \geq r_{k+2}$, for $k = 3,5,...,s-2$;

$n_k \leq n$, $r_k \leq n^{\frac{k-1}{2}}$, for $k = 3,5,...s$;

n_k, r_k are expressed by integer numbers, for $k = 3, 5,...s$.

Using the optimality principle, the problem may be reformulated. Let us denote by x_t the number of inlets for a q-stage fabric. $T(x)$ is the set of decisions (n_i, r) permitted by x, where $x_q = r_q$, for $q = 3,5,...,s-2$;

$$T(x) = \left\{(n_i, r)|0 < n_i \leq n, 0 < r \leq n, n_i r \geq x\right\} \tag{2}$$

The optimization problem is now embedded in a set of problems:

$$\min\left\{2r_q + n_q\left[2r_{q-2} + n_{q-2}\left[...[2r_3 + n_3]...\right]\right]\right\} \tag{3}$$

for $q = 3,5,...,s$

$n_s r_s \geq L$

$n_k r_k \geq x_k$

$n_k \leq n$, $r_k \leq n^{\frac{k-1}{2}}$

n_k, r_k are expressed by integer numbers

$x_k = r_{k+2}$ for $k = 3,5,...,q$

The optimum solution of this problem is denoted by

$$f_q(x) \text{ for } 0 < x \leq L, q = 3,5,...,s$$

The function $f_3(x)$ can be easily found for all $0 < x \leq L$ from the definition

$$f_3(x) = \min\left\{(2r + n_i) \mid (n_i, r) \in T(x)\right\} \tag{4}$$

To determine the function $f_5(x)$ describing the minimum cost of the five-stage fabric, the principle of optimality is used. It should be remembered that no matter what the first decision (n_i, r) is, the continuation must be optimal, according to the value of r obtained during the optimization of three-stage fabrics. The value r denotes the number

of inlets for each three-stage fabric which forms a "major" middle switch nested inside a five -stage fabric. The best solution for the five-stage fabric, therefore, can be found by comparing the total cost resulting from the different decisions (n_i, r) and from their optimal continuations:

$$f_5(x) = \min\{[2r + n_i f_3(r)] \mid (n_i, r) \in T(x)\}, \text{ for } 0 < x \le L \tag{5}$$

Similarly, after determining $f_{q-2}(x)$, we obtain

$$f_q(x) = \min\{[2r + n_i f_{q-2}(r)] \mid (n_i, r) \in T(x)\} \tag{6}$$

The optimum solution for a given q is found by using enumeration methods. The essence of these methods is to enumerate all solutions, i.e., all integer vectors in the given range, and to choose according to the objective function the best one from among those satisfying the constraints of the problem. Enumeration can be used for our problem because a high proportion of solutions is implicitly examined, i.e., excluded in large groups without direct checking, as they are shown to violate the set of constraints. It should be noted that we find the parameters of a q-stage fabric using a table of the optimal solutions for $(q-2)$-stage fabrics.

An analogous approach can be used for more complex fabric structures, allowing single-stage multielement switches (as those shown in Fig. 4) as well as concentration and expansion.

5 Numerical Examples

Dynamic programming formulas illustrated by Section 4 were used to calculate extensive tables of optimum ATM switching fabrics. Here, we discuss some results of the optimization. The cost was expressed by the number of used elementary ATM switches. In all calculations we assumed that splitting points (as those in Fig. 4) are passive and their cost is negligible. The combining points were implemented with standard elementary switches and they are added to the total switch count.

For example in Fig. 6 a 16 × 16 switching fabric composed of 4 × 4 switches is shown with expansion ratio in the first stage equal to 2: 1. As we can see the implementation of the combining points requires 8 additional ATM switches forming the last stage (each 4 × 4 switch serves as two 2 × 1 multiplexers).
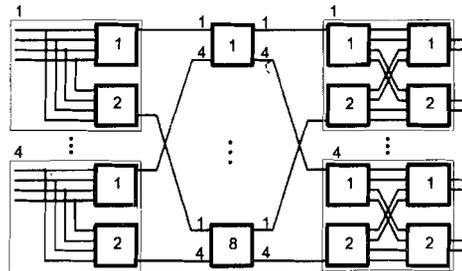
Figure 6 - 16 × 16 switching fabric

Let us consider the growth of an ATM fabric composed of 8×8 shared-buffer switching elements. We assume the expansion ratio of the first stage to be $n : 2n - 1$ that would result in the strict sense nonblocking fabric, provided circuit-switched environment (although an ATM fabric can incur some loss). Fig. 7 shows three types of basic elements used in our architecture. They can be implemented as separate boards. Fig. 8 illustrates growth from 64 to 128, and then to 192 ATM ports. We can note that if boards of Fig. 7 are used, the fabric can grow up to 256 ports without any changes of the internal board structures and without any changes of existing cabling (some cables are only added). However, we can note that the application of 4: 1 multiplexers (two of them are implemented by using a single standard 8×8 ATM switch) increases initial cost of a growable fabric.
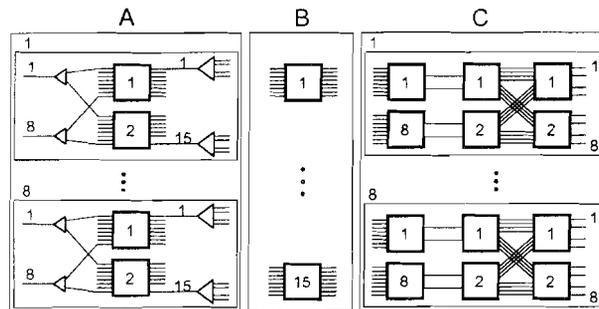
Figure 7 - Basic boards

Table 1 illustrates the differences between the numbers of applied ATM switching elements and the number of switching stages for different architectures. The first column shows optimum structures (in the sense of the minimum number of ATM elements). Such an architecture requires extensive recabling at each growth stage. The second column represents regular Clos-like fabrics, i.e., such fabrics that do not contain single-stage arrangements such as those shown in Fig. 4, and as the result, no ATM

105

switches are spent for combining points. The third column gives the number of ATM elements for our growable architecture. We can see that the initial cost is relatively high, due to ATM switches serving as 4: 1 multiplexers.
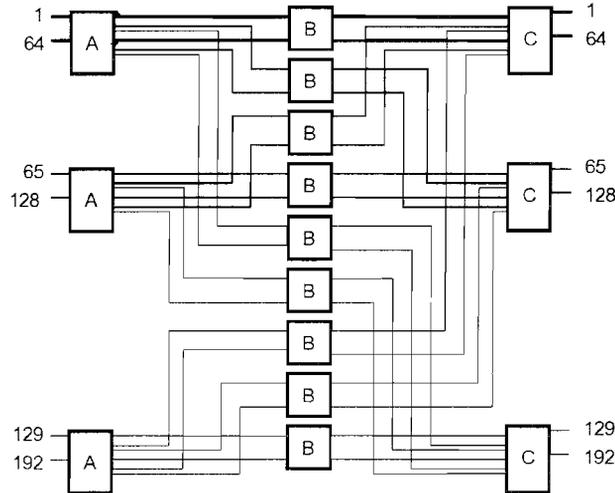
Figure 8 - Growth form 64 to 192 ATM ports

L	Optimal	Stages	Regular	Stages	Growable	Stages
64	63	4	116	5	127	5
128	216	5	225	5	284	5
192	429	6	708	7	471	5
256	537	6	931	7	688	5

Table 1 - Number of ATM switches for different fabric architectures

However, the number of switching stages (including those containing multiplexers) is kept constant that reduces delay problems.

Fig. 9 compares fabrics composed of 8×8 with $vm = 2n - 1$, i.e., fabrics that would represent nonblocking circuit-switched structures. The numbers of ATM switches resulting from the following strategies are presented.

Strategy 1: The initial growable fabric has 128 ATM ports and is optimal in the sense of the minimum number of switches (15×8 first-stage modules).

106

Strategy 2:   The initial growable fabric has 128 ATM ports and the first stage
contains 31 × 16 modules.

Strategy 3:   The initial growable fabric has 64 ATM ports and the first stage contains
15 × 8 modules.

Strategy 4:   For each number of ports the fabric contains the minimum number of
ATM switches.

Strategy 5:   For each number of ports the Clos-like fabric (i.e., without single-stage
multi-switch modules) contains the minimum number of ATM switches.
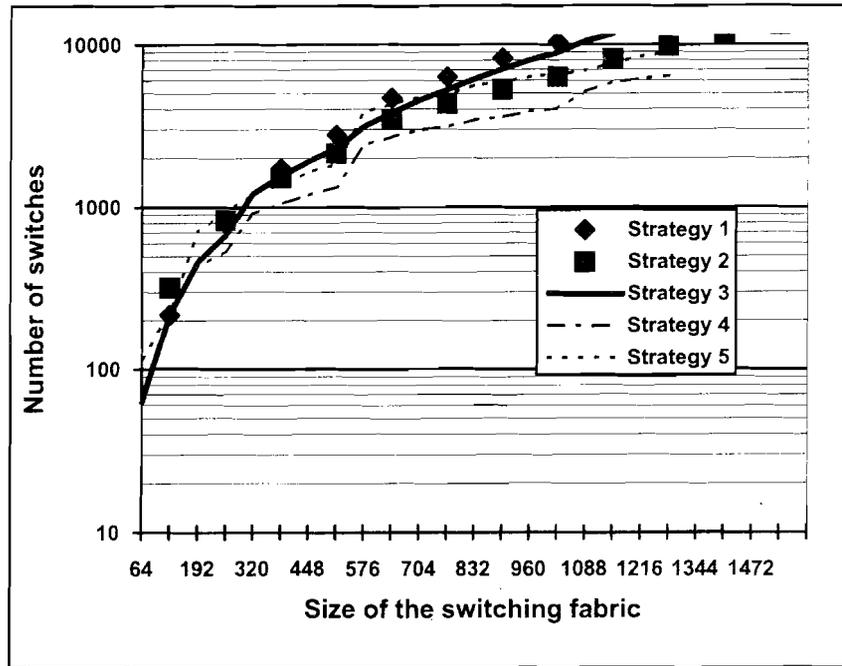
Figure 9 - Comparison of the cost of growable and conventional fabrics

We can see that selection of the optimum structure as the initial fabric does not
necessarily result in the minimization of the number of ATM switches as the fabric
grows. An important design decision involves the size of the initial fabric. If we
compare Strategies 2 and 3, we can see that Strategy 2 gives lower-cost fabrics starting
from 384 ports while Strategy 3 is superior for smaller fabrics. Strategy 3 allows also
smoother growth (in smaller steps) than Strategy 2. Although Strategy 4 always selects

107

optimum fabrics, it would require extensive recabling at each expansion step. Clos-like fabrics contain even more switches than some growable structures (despite the fact that they do not require ATM switches serving as the multiplexers).

6 Conclusion

The appropriate selection of structural parameters of a ATM shared-buffer-based fabrics has a significant influence on the fabric's cost. We can note that introduction of single-stage submodules in multistage fabrics can significantly reduce the cost, provided there is appropriate expansion in some stages and concentration in others.

The following procedure can be used in the design process of growable ATM fabrics:

Step 1.   Select the fabric size range.

Step 2.   Select the growth step.

Step 3.   Select the expansion ratio.

Step 4.   Find the optimum set of fabric structures within the size range.

Step 5   Check the traffic performance. If not satisfactory, modify the expansion ratio and go to Step 4.

Step 6.   Design modules A, B, C (see Fig. 7).

The optimum set of fabric structures, that form the optimum trajectory, is found by using the dynamic programming method described in Section 4. Since different trajectories, depending on initial structures, can cross each other, the designer has to take the final selection decision, depending for example, on the expected number of fabrics of each particular size to be manufactured (therefore minimizing the total number of ATM switching elements used in the production process). Since the selection of the expansion ratio in Step 3 was rather arbitrary (accurate analytical methods for performance evaluation are yet to be developed), the traffic performance of the selected fabrics has to be checked by using simulation methods. If the obtained results are significantly different from the expected performance, the procedure has to be repeated for the new expansion ratio.

7 References

[1]   A. Jajszczyk, "Novel architecture for a digital switching network", *Electronics Letters*, vol. 20, p. 683, Aug 1984.

[2]   A. Jajszczyk, W. Kabacinski, "A growable ATM switching fabric architecture", *IEEE Trans. Commun.*, vol. 43, No. 2, Feb. 1995.

[3] Y. Sakurai, N. Ido, S. Gohara, and N. Endo, "Large scale ATM multi-stage switching network with shared buffer memory switches", in *Proc. XIII Int. Switching Symp.*, Stockholm, Sweden, May/June 1990, vol. IV, pp. 121-126.

[4] T. Kozaki et al., "32 x 32 shared buffer type ATM switch VLSI's for B-ISDN", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1239-1247, Oct. 1991

[5] A. Itoh et al., "Practical implementation and packaging technologies for a large-scale ATM switching", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1280-1288, Oct. 1991.

[6] W. Fisher, O. Fundneider, E.-H. Foeldner, and K. A. Lutz, "A scalable ATM switching system architecture", *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1299-1307, Oct. 1991.

[7] A. Jajszczyk, "A dynamic programming approach to optimization of switching networks composed of digital switching matrices", *IEEE Trans. Commun.*, vol COM-35, pp. 1342-1346, Dec. 1987.

Andrzej Jajszczyk received the M.S. degree in electrical engineering in 1974, and Ph. D. and Dr. Hab. degrees in communications from Poznan University of Technology, Poznan, Poland, in 1979 and 1986, respectively. From 1974 to 1992 he was with Poznan University of Technology. In 1993 he joined EFP - The Franco-Polish School of New Information and Communication Technologies in Poznan, where he is presently a Professor and the Head of the Telecommunications Switching and Networks Group. In 1989/90 he spent 12 months as a visiting scientist at the Teletraffic Research Center, the University of Adelaide in Australia. In 1991 and 1992 he was a visiting scientist at the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada. He has been engaged in research and teaching in the areas of telecommunications switching and computer communications. He is the author of 6 books and over 100 papers, as well as 19 patents in these areas. He has led several projects for industry concerning the design and performance evaluation of digital switching systems. His current research interests include bradband networks and switching, both electronic and photonic. He has served as a Consultant to telecommunications industry and operators as well as government agencies in Poland, Australia, and Canada. He is the Editor of the IEEE Global Communications Newsletter and the Editor for Switching Theory and Fabrics for the IEEE Transactions on Communications. Dr. Jajszczyk is a member of the IEEE and the Assosiation of Polish Electrical Engineers.

Michal Roszkiewicz received the M.S. degree in electronic engineering from Poznan University of Technology, Poland, in 1993. Shortly after graduation he joined the EFP - The Franco-Polish School of New Information and Communication Tecnologies in Poznan, Poland, where he is an assistant at the Telecommunications Switching and Networks Group. He is currently working towards his Ph. D. degree. His research interests are ATM switching fabric architectures and large-scale ATM switching networks.