An Optical Contention Resolution and Buffering Module for ATM Networks

David K. Hunter and Ivan Andonovic,
Optoelectronics Group,
Dept. Electrical and Electronic Engineering,
University of Strathclyde,
204 George Street,
Glasgow G1 1XW,
United Kingdom

ABSTRACT

A new type of buffering module is introduced, suitable for incorporation into optical switch fabrics handling ATM cells. It comprises 2x2 optical switches and optical delay lines, and is equivalent to two interlinked buffers. Performance is characterised in terms of attenuation, crosstalk and cell loss.

INTRODUCTION

A new architecture for contention resolution in future optical ATM switching fabrics is introduced. Using ATM [1] in optical networks has the advantage over TDM techniques [2] of allowing extreme flexibility in the type of traffic that is carried; any bitrate can be supported, and bursty traffic can be carried without wasting capacity. In addition, optical switching is inherently transparent to bitrate, coding format and wavelength. Due to the unscheduled nature of the transmission of ATM cells, some means of resolving contention within the switch fabric is required. A module which performs this function is described here.

The architecture consists of a chain of 2x2 optical switches and optical delay lines, and is equivalent to two interlinked buffers. Implementation using lithium niobate switches and fibre delay lines is assumed since this represents the most mature technology, although other technologies could be used in the future without marked modification of the architecture. The geometry exhibits a low component count which is logarithmically related to the buffer depth; to double the size of the buffers, only one extra switch is required. Furthermore, it may be controlled electronically, using a simple control architecture consisting of a counter and associated random logic. Feed-forward delays are used throughout, implying uniform attenuation and superior crosstalk performance [2].

The switches only need change state at the cell rate. Thus, with the inclusion of a suitable guard band, the electronic control rate is effectively much lower than the optical data rate. Due to the transparent nature of optical switches, the data rate is limitless, and in principle, data rates of tens of gigabits per second should be possible.

As an example, the unit could be incorporated into a buffered banyan architecture, with each optical buffered switch module representing one buffered switch in the banyan network. This approach is expandable to any size, limited only by attenuation and crosstalk, and would require simple header recognition since only one bit of the address would need to be decoded in each module.

MODULE ARCHITECTURE

The architecture is referred to as *B(n)* where n is the combined capacity of the two buffers within it. Each module consists of a chain of $\log_2 n + 1$ switches and $\log_2 n$ delay lines, and may be defined recursively (Fig. 1). An example of such a network, *B(8)*, is shown in Fig. 2. To simplify the discussion, it will be assumed that every timeslot is full; this is easily adapted to the more general case [3].

Each switch module has two inputs and outputs; each cell entering the network is designated "U" or "L", depending on whether it is destined for the upper or lower output respectively. The objective is to ensure that on each timeslot, it is impossible for two "U" cells or two "L" cells to reach the outputs simultaneously. The module is therefore effectively buffering the cells, since each cell is stored until a free output is available to transmit it. There are three possibilities for the cells entering the input - two "L" cells, two "U" cells, or one "L" cell and one "U" cell.

Header detection and recognition would be required at the inputs to each module. Here only one bit of the header need be detected if the module were incorporated within a buffered banyan architecture. The header processors would determine whether each incoming cell was a "U" or "L" cell, this information being passed on to the control electronics. The control logic would operate at the optical cell rate rather than the optical bit rate, and would consist of a $\log_2 n$ - bit up/down counter to keep track of the network state, plus associated random logic. Control signals for the switches could be derived by using a parallel electronic network which would contain electronic switches and delays corresponding to those in the optical network. The electronic network would carry only small packets representing the delay of each cell; these small packets would self-route through the electronic network, while the switch settings from the electronic switches, once determined, would be transferred to the optical switches.

Suppose that the network state is represented by an integer s which can take on values from 0 to $n-1$. If the network is in state s and one "L" and one "U" cell arrive then the "L" cell will be delayed by s timeslots before emerging on the lower output, and the "U" cell will be delayed by $n-1-s$ timeslots. The network remains in state s.

Again, suppose that the network is in state s, but two "L" cells arrive at the input. Then one will be delayed by s timeslots and the other will be delayed by $s+1$ timeslots. The network goes into state $s+1$. Similarly, if the network is in state s and two "U" cells arrive, one cell will be delayed by $n-1-s$ timeslots and the other will be delayed by $n-s$. The network goes into state $s-1$.

This ensures that one "U" cell and one "L" cell arrive at the output on every timeslot, preserving the ordering of the cells, and effectively being equivalent to a buffering operation. If two "L" cells arrive in state $n-1$, or two "U" cells arrive in state 0, one cell will be wrongly routed since the buffer has overflowed. For example, in the former case, one will be delayed by $n-1$ timeslots and sent to the lower output, while the other will be delayed by 0 timeslots and sent to the upper output. To make this as unlikely as possible, the network should initially be in state $n/2 - 1$ or $n/2$. To cope with the case of an input timeslot being empty, the empty timeslot can be treated as a "U" or "L" cell in such a way as to make the state *s* move towards $n/2 - 1$ or $n/2$ if possible [3].

PROOF OF OPERATION

The routing of cells described above can always be accomplished without any contention within the fabric. A *connection* is either a delay line or a link. A *virtual connection* is an ordered pair of a connection and a timeslot. As discussed earlier, the delay to be experienced by a cell is decided before it enters the network. In addition, for any input/output pair and any given delay, there is only one possible path through the module. The controller can be thought on as reserving virtual connections for the use of each new cell, where each virtual connection represents a connection and the timeslot on which the cell will enter that connection. The *reserved set* in a particular state is the set of virtual connections reserved for all the cells currently in the module. When two new cells enter the module, the appropriate virtual connections are added to the reserved set. To demonstrate the validity of the modules, it will be necessary to show that no conflict exists between the reserved set and the virtual connections for each two new cells entering the module.

The proof is by induction. Consider Fig. 1, and suppose that the network $B(n/2)$ operates correctly and that there is a particular reserved set which is always associated with each state. This is obviously true for the trivial case of $B(1)$, and to complete the induction, it will be necessary to show that it is true for $B(n)$ also. Supposed that the right-hand switch in the upper half of Fig. 1 is in the bar-state. Then it can be seen that if the $B(n/2)$ network is in a state *s* where $s = 0 \ldots n/2 - 1$, the whole

$B(n)$ network will also be instate s. Also, if $B(n/2)$ moves between these states, so will $B(n)$, and the reserved set associated with any state of $B(n)$ will always be dependent only on s. Likewise, if the switch is in the cross-state, it can be shown that if $B(n/2)$ is in state s then B(n) will be in state $s + n/2$. The "U" and "L" cells emerging from $B(n/2)$ must be exchanged, and this can be done by inverting the state of the rightmost switch in $B(n/2)$.

It now remains to show that the transition from state $n/2 - 1$ to $n/2$ is possible; the proof for the converse is similar. It is easily shown that the reserved sets for states s and $n - 1 - s$ are the same. So a module in state $n/2 - 1$ may be put into state $n/2$ without fear of any conflict arising. Suppose it is in state $n/2 -1$, but two "L" cells arrive. Normally, if an "L" and a "U" cell arrived, "L" would be delayed by $n/2 - 1$ timeslots and "U": would be delayed by $n/2$. The two "L" cells may be delayed by $n/2$ and $n/2 - 1$ without altering the reserved set. The reserved set is the same as for state $n/2$, so the network may be driven into that state.

NETWORK PERFORMANCE

Markov chain analysis can be used to evaluate the probability of a given cell being lost to the wrong output [3]:

$$W = \frac{p}{8} \frac{a - 1}{a^{n/2} - 1} a^{n/2-1}$$

where

$$a = \frac{p^2}{p^2 - 4p + 4}$$

p is the load (i.e. the probability that a timeslot is full), and the distribution of traffic is assumed to be uniform. The assumption of uniform traffic makes the network easy to analyse, but gives an over-optimistic estimate of the cell loss. Nevertheless, it is a useful first approximation. Fig. 3 is a graph of the cell loss W against load p for various sizes of network. As expected, the cell loss becomes lower as the switch network becomes larger; for a large $B(128)$ network, the loss is as low as 3.9×10^{-13} for a high channel load of 0.9. Even for $B(16)$, the cell loss is 1.2×10^{-8} for a load of 0.5. To evaluate the cell loss of a buffered banyan network using this architecture, it is necessary to multiply the loss by the number of stages.

The attenuation of a $B(n)$ network is $A = L(\log_2 n + 1)$ where L is the insertion loss of a single switch, including coupling loss. If L = 5dB, then $B(16)$ has 25dB attenuation; few systems could accommodate greater loss without optical amplification. This figure was calculated assuming that fibre delay lines were used - the use of a delay line between each pair of consecutive switches prevents the use of large substrates. This could be overcome by using a technology, such as silica waveguides, where the delays can be integrated with the switches.

As a simple estimate of crosstalk performace, the signal to crosstalk ratio (SXR) is given by $SXR = | X | - 10\log_{10}(\log_2 n + 1)$, where X is the extinction ratio of one switch. Using a more sophisticated model [4], it can be shown that $B(16)$ i.e. 5 switches, each having an extinction ratio of 20dB, results in a degradation of BER from 10^{-9} to 1.42×10^{-9}. With a three stage switch module, as would be found in a 8x8 banyan network, the BER is degraded to 2.61×10^{-9}; this effectively results in each cell passing through 15 switches. In this case, optical amplifiers would have to be included, and their noise performance would have to be taken into account in a complete characterisation of the network.

CONCLUSIONS

A new architecure for buffering in optical ATM switching systems has been introduced. The architecture is equivalent to two interlinked buffers, and is realised with a low component count. It is suitable for incorporation within larger architectures such as buffered banyan netoworks. The network is easy to control using electronics; for a data rate of 40Gb/s, the control rate is less than 100Mb/s. It can be implemented using devices which exist now, and an experimental verification of this architecture will be reported shortly. Finally, it may have other applications, such as, for example, a universal delay and buffering unit for optical computing.

REFERENCES

1. Fox, A. L., Joy A. K.: "ATM-Based Switching for the Integrated Broadband Network", *Electronics and Communication Engineering Journal*, August 1990, pp137-146

2. Hunter, D. K., Andonovic, I., Smith, D. G., and Culshaw, B.: "Architecture for Optical TDM Switching", *Electronics Letters*, vol. 28, no. 18, 27th August 1992, pp1746-1747

3. Hunter, D. K.: "An Optical ATM Switching Node Based on the Buffered Banyan Architecture", *University of Strathclyde,* 6th May 1992

4. Saxtoft, C.: "Photonic Switching Networks - Component Characteristics versus Network Requirements", *Ph.D. Thesis,* Technical University of Denmark, March 1992

FIGURE CAPTIONS

Fig. 1: Definition of the buffered switching module $B(n)$

Fig. 2: An example of a 2x2 buffered switching module, $B(8)$

Fig. 3: Log of cell loss probability $\log_{10}(W)$ against load p for various sizes of buffering module

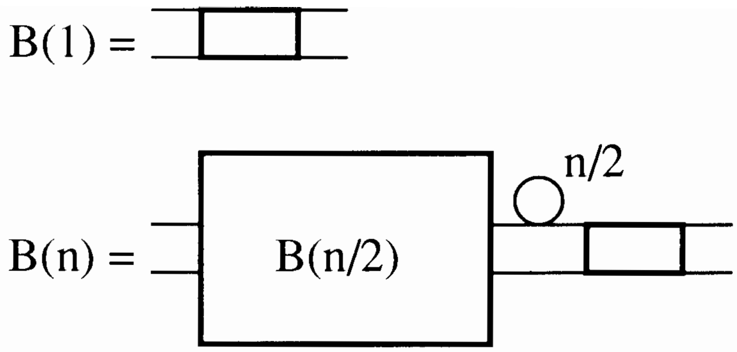

Fig. 1

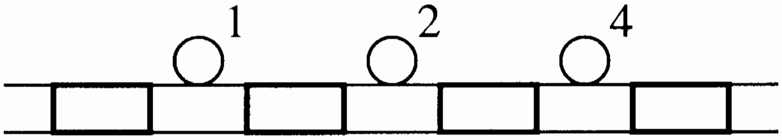

Fig. 2

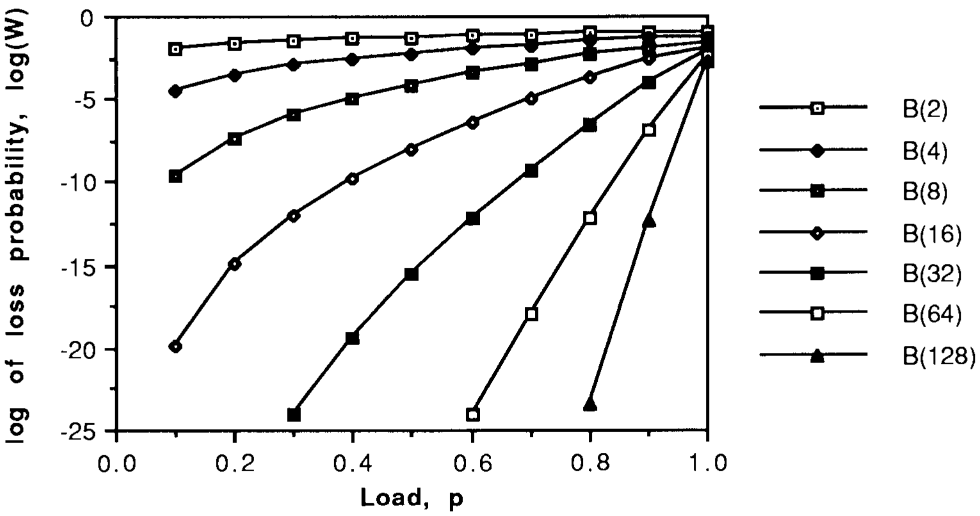
log of loss probability, log(W)
0
-5
-10
-15
-20
-25
0.0
0.2
0.4
0.6
0.8
1.0
Load, p
B(2)
B(4)
B(8)
B(16)
B(32)
B(64)
B(128)

Fig. 3