Architecture for large dilated optical TDM switching networks

D.K. Hunter  
P.J. Legg  
I. Andonovic

Indexing terms: Cross-connect, Integrated optical switch arrays, Lithium niobate switch arrays

Abstract: A novel architecture for optical TDM switching is introduced, comprising 2 × 2 optical switches integrated onto large substrates, and fibre delay lines. The performance of the architecture is characterised in terms of switch count, crosstalk, frame delay, attenuation and control complexity. It is shown that in the near future a system with 16 inputs and outputs, handling a total of 4096 TDM channels, could be demonstrated experimentally with 448 switches on six substrates using state-of-the-art technology.

1 Introduction

This paper describes a new architecture for optical TDM switching applications and considers the feasibility of realising a TDM switch aimed at a possible future high-speed optical cross-connect application. These switching networks are likely to find application in telecommunications for three main reasons. First, optical switching offers virtually unlimited bandwidth, as well as transparency to bitrate and coding format, thus ensuring future-proof systems while at the same time reducing the need for optical-to-electronic and electronic-to-optical conversion. Secondly, TDM will be used both to carry high-speed-continuous bitrate services, such as high-definition television (HDTV) and high-quality video telephony, and as a transport mode for bursty traffic in ATM cells. Finally, optical switching implies a reduction in EMI problems compared to an electronic approach.

The architecture is constructed from multiple large substrates carrying guided-wave optical 2 × 2 switches interlinked by optical fibre delay lines, allowing implementation using lithium niobate directional coupler switch arrays and fibre delay lines, which represent the most mature technologies available at present. The use of large substrates minimises the fabrication (and probably interconnection) costs, and additionally improves the signal attenuation level. The switch networks are intended for use with block multiplexing [1], where the switches only change state between blocks of bits (i.e. every timeslot) rather than every bit. Thus the electronic control bitrate can be much lower than the optical data bitrate. Furthermore, since the crosspoint count varies logarithmically with the number of timeslots per frame, the hardware requirements are extremely economical.

Assuming a suitable resynchronisation scheme is employed with the switch network, the bandwidth available is ultimately limited by dispersion. Hence, these networks would be deployed in situations not considered by current standards. For this reason, standards issues have not been addressed in this paper, although it is anticipated that the networks could be interfaced to the complex frame structure of SDH/SONET [2] at the expense of greater hardware complexity.

2 Networks without frame integrity constructed from large substrates

The TDM switching networks are constructed from standard space switch networks — baseline [3], reverse baseline [3] and Beneš networks [4]. m × m networks of these types are represented here by B(m), R(m) and W(m), respectively. All consist of stages of m/2 switches, where Beneš networks have 2 log2 m − 1 such stages and the others have log2 m. Although Beneš networks are considered here, Waksman networks [5] are very similar and may be used instead.

The notation T(m, n) is used to denote a TDM switching network with m input links, m output links and n timeslots per frame. It is thus capable of switching a total of mn channels, there being (mn)! possible permutations. T(em, n) may be constructed from smaller networks T(m, n/f) of the same type by using theorem 5, Fig. 1 (see Fig. 2 notation). A proof of theorem 5 and a description of the bit-reversal function r(k, i) is presented in the Appendix. One starts the synthesis with T(m, 1) which is equivalent to W(m) (or any m × m space switch architecture, the Beneš network uses nearly the theoretical minimum number of switches). With one timeslot per frame, no interchange of blocks between timeslots is possible, hence the network functions as a space switch which changes state every timeslot. To build a larger network from a smaller one, the original network is substituted for each switch in the centre stage of Fig. 1, resulting in a new, larger network. This process can be repeated as many times as necessary to produce a network of the required size.

D.K. Hunter and P.J. Legg acknowledge the financial support of BT Laboratories during much of this work.

IEE PROCEEDINGS-J, Vol. 140, No. 5, OCTOBER 1993
These networks may be dilated \([6, 7]\) to improve their crosstalk performance at the expense of using just over twice as many switches; the improved crosstalk performance allows large networks to be constructed from relatively poor performance switches \([7]\).

3 Providing frame integrity

The method proposed here for providing frame integrity differs from those considered elsewhere \([7-9]\), and is particularly applicable to large dilated networks.

Consider Figs. 1 and 2. This type of network does not have frame integrity because some blocks may cross a frame boundary when passing through a delay line, while others may not. To avoid the problem, extra delay lines can be added so that all the blocks cross over the frame boundary into the same frame, resulting in the structure shown in Fig. 3, which is intended for dilated networks.

\[
\begin{align*}
\text{Fig. 1} & \quad \text{Theorem 5, creating } T(m, n, f) \text{ from } T(m, n, f), \text{ can be used to create networks of any size from large substrates.} \\
\end{align*}
\]

\[
\begin{align*}
\text{Fig. 2} & \quad \text{Shorthand notation for delay lines in Fig. 1, with } j = \log f \\
\end{align*}
\]

Here, each switch corresponds to one input or output link in Fig. 2 and only one block passes through a switch at once. Hence each block may be routed through either the original length of delay line or the original length plus \(f\) timeslots. Blocks which would not normally cross the boundary are forced to do so by being passed through this additional delay. This yields frame integrity without using any extra switches.

Suppose that the frame at the right hand side of Fig. 3 is \(f - 1\) timeslots behind the left-hand side. The blocks directed to the lowest left stage switch would never pass through the corresponding \(f + r(f - 1, j)\) delay line, since a delay of \(f - 1 = r(f - 1, j)\) will always transfer the block into the correct frame. Therefore, the \(f + r(f - 1, j)\) delay line is omitted.

4 Performance of networks

4.1 Formulas characterising performance

In this Section, certain assumptions are made about the network to make the calculations tractable. Let \(d = ef\) be the number of inputs or outputs on each substrate. Then the procedure used for creating a TDM switching network, after starting with a Beneš network \(N(d)\) is

(a) Fig. 1 was used \(\log n\) times with \(e = 1\) and \(f = d\) to create \(T(d, n)\).

(b) Then Fig. 1 was used \(\log m\) times with \(e = d\) and \(f = 1\) to produce the desired fabric \(T(m, n)\).

Crosstalk is always assumed to add to the signal incoherently, and the path dependence of waveguide bend radii and intersection losses will be ignored in the formulas. Thus they should be viewed as providing only an approximate indication of network performance.

Table 1 summarises the performance of \(T(m, n)\) networks. The frame delay is the delay between an input frame entering the network and the corresponding output frame leaving. \(W\) is the loss due to a single fibre/substrate coupling, and \(L\) is the mean insertion loss per crosstalk (excluding coupling loss). \(X\) is the extinction ratio of each switch in dB.

\[
\begin{align*}
\text{Table 1: Summary of dilated network performance} \\
\begin{array}{|c|c|}
\hline
\text{Parameter} & \text{Formula} \\
\hline
\text{Frame delay} & \mu = 2m \\
\text{Number of switches} & S = 2m(\log mn + \log mn - 1) \\
\text{Attenuation} & A - 2L(\log mn + \log mn - 1) \\
\text{Signal-to-crosstalk ratio} & SXR = 2X \\
& - \frac{20 \log_{10}(2 \log mn + 2 \log mn - 3)}{} + 3 \\
\hline
\end{array}
\end{align*}
\]

4.2 Control complexity

These networks can be controlled by adapting standard algorithms for controlling Beneš and Waksman networks; essentially, the problem of controlling a \(T(m, n)\) network is the same as controlling an \(mn \times mn\) Beneš or Waksman network \([7]\). The control complexity depends on the number of processors used. With one processor the complexity is \(O(mn \log mn)\) \([10]\), but with \(mn\) processors it becomes \(O((\log mn)^2)\) \([11]\).

To reduce the control complexity, it is possible to stack \(\log mn\) networks in parallel (using multiplexers and demultiplexers to interconnect them), to form a network which is strict-sense nonblocking \([12, 13]\); such a network also has the advantage of exhibiting fault tolerance. Control algorithms are currently being investigated that exhibit a more favourable trade-off between complexity and hardware cost, and the results will be reported in a future publication.

4.3 Example network

Fig. 4 is an example of the TDM networks discussed in this paper. It has 16 inputs and outputs, is dilated, has frame integrity, and switches 256 timeslots per frame. The total number of TDM channels is thus 4096. The fabric consists of six 16 × 16 substrates which, as discussed in Section 6, have actually been implemented [14]. An experimental realisation of a multisubstrate space switch network of the same size, 448 switches, has recently been reported [15].

5 Performance graphs for dilated frame integrity networks

In this Section, graphs are presented that illustrate the attenuation and crosstalk performance of different sizes of dilated frame integrity networks. These allow limits on network size to be determined, depending on the desired system performance.

5.1 Attenuation against network size

Fig. 5 characterises the dependence of the attenuation of networks fabricated from 4 × 4 and 16 × 16 substrates on the number of TDM channels supported. Values of $L = 1.25$ dB (crosspoint loss) and $W = 1$ dB (fibre/substrate coupling loss) were used. Fabrics of a given total capacity made from small substrates have higher attenuation than those made from large substrates. This may be understood by noting that smaller substrates imply a greater number of both fibre/waveguide couplings and switch stages (on dilation each substrate has one extra stage added to it). Thus besides probably exhibiting lower interconnection cost, fabrics made from larger substrates have lower attenuation too. Networks of differing spatial dimension but having equal total channel capacity and equally sized substrates exhibit equal attenuation; the number of stages is dependent only on the total number of channels.

5.2 Effect of crosstalk on bit error ratio

To estimate the system power penalty due to crosstalk, a simple model of the detection process was developed [16, 17] (Fig. 6). The presence of equal independent and
uncorrelated crosstalk contributions, and the absence of loss variations was assumed. (The effect of crosstalk-induced interferometric noise would have to be taken into account in a more thorough evaluation [18].) A system power penalty of <0.1 dB is readily achieved at a bit error rate (BER) of 10⁻⁵, even for large networks fabricated from crosspoints of moderate performance (isolation < -14 dB).

### 6 Feasibility study on example network

This Section presents a possible implementation of the network T(16, 256) (Fig. 4) using lithium niobate space switches of order 16 x 16 interlinked by polarisation-maintaining (PM) fibre delay lines. Preliminary calculations suggest that the cost of the switch fabric would be dominated by the cost of the substrates.

#### 6.1 16 x 16 lithium niobate space switches

Integrated optics employing titanium indiffused lithium niobate waveguides has been extensively researched and represents the most mature technology for optical space switch realisation [19]. Space switches up to dimension 16 x 16 have been demonstrated, their size being limited by the lower bound of the crosspoint length (to ensure low switching voltage and therefore rapid <1 ns rise time), the minimum bend radius (to minimise propagation losses) and the length of the substrate itself (<100 mm). Only arrays of polarisation-dependent crosspoints demonstrate sufficiently small drive voltages to permit rapid switching.

The 16 x 16 dilated Beneš switch at the centre of T(16, 256) has been demonstrated in lithium niobate [14]. The array was fabricated on two butt-coupled substrates and employed polarisation-dependent directional coupler crosspoint of switching voltage 12.4 ± 0.3 V, switching time ~1 ns and crosstalk < -15 dB at 1.3 μm. The fully packaged switch had a loss of 13.4 dB. Variations in the insertion loss of differing input to output connections arise principally from the path-dependent waveguide intersection count, and lead to a system power penalty by increasing the effective interchannel crosstalk. Insertion loss variations of ~1.5 and ~4.0 dB for a dilated 16 x 16 baseline/reverse baseline and Beneš network, respectively, have been estimated from a reported value of ±1.5 dB for an 8 x 8 dilated Beneš structure [20]. The baseline and reverse baseline arrays may be realised on single 4 in substrates by adding an extra switch stage to half a dilated Beneš switch.

#### 6.2 Fibre optic delay lines

The polarisation dependence of the space switches necessitates polarisation control [21] for each single-mode input fibre to the fabric and PM fibre interlinks cut to the correct length to achieve the time delay required. For a temporal delay error <1/16 of the bit period, t, seconds, the cut tolerance is given by ΔL < 2.05 x 10⁻⁶ t/16 metres. For example, at a line rate of 2.488 Gbit/s, ΔL < 5.2 mm. Furthermore, the fibre delay lines are sensitive to temperature, as characterised by the delay of vitreous silica which has been measured as 40 ps/(km °C) [22] at 1.3 and 1.55 μm. Therefore to maintain a delay error of <1/16, the temperature fluctuations is ΔT < 1/(16 x 8.2 x 10⁻⁹ x de) °C, where d is the delay length in time slot units, and e is the number of bits per timeslot. If ΔT = 1 °C, de < 7610. For T(16, 256), the maximum value of d is 480 implying e < 16. A delay of one timeslot is consequently approximately one metre in length given a data rate of 2.488 Gbit/s. These restrictions on the block format could be eliminated by employing a suitable resynchronisation scheme.

Fibre attachment to the space switches involves the alignment of silicon v-groove fibre arrays to the waveguides followed by epoxy bonding [23]. Lateral misalignment typically adds a loss of 0.5 dB [23], while angular misalignment of <1.5° has been reported for PM fibre arrays [20]. If both ends of a PM fibre delay line are angularly misaligned by θ with respect to each waveguide, the (worst case) fractional power coupled into the incorrect polarisation (TE say) is given by ~1 - cos²θ [17]. Within the space switch the TM and TE components are routed differently owing to the inherent polarisation sensitivity of the crosspoints, the TE being widely distributed among several TM message channels. This phenomenon is repeated at each stage of the fabric; the crosstalk accumulates largely as a TE-polarised waveform. Consequently a polarising element at the end of the switch fabric, either a dielectric or metal-loaded polariser integrated at the end of the final substrate, or a polarising fibre tail attached to each output, should significantly reduce this polarisation crosstalk.

![Fig. 7 Dependency of worst-case polarisation crosstalk on fibre-waveguide angular misalignment](image)

(Fig. 7). It may be concluded that tight alignment tolerances are maintained (θ < 1.5°), the system power penalty due to polarisation crosstalk level (< ~ 22 dB) is insignificant (<0.1 dB). Alternatively, tolerances may be relaxed if simple polarisers are employed (a small power penalty due to signal loss is incurred, 0.2 dB for θ = 4.5°).

#### 6.3 Assembled system performance

The tolerance of the fibre delay lines, as discussed, limits the block capacity to 15 bits at a data transmission rate of ~3 Gbit/s. To allow switch reconfiguration between successive blocks, a guard band of 3 bit periods or ~1 ns may be included at the end of each block, preceded by 12 bits at an instantaneous bitrate of 3.06 Gbit/s to give a net line rate of 2.488 Gbit/s.

Evaluation of the attenuation formula (Table 1) with W = 1 dB and L = 1.25 dB gives a total loss of 45 dB. If the input power to the fabric is 0 dBm, and the receiver sensitivity is ~ -35 dBm for a BER = 10⁻⁹, amplification of >15 dB is required for a 5 dB power margin. However, to limit effective crosstalk growth, the loss variations in the space fabrics must be compensated, and therefore automatic power control (APC) amplifiers are proposed [24]. In a study of a space network of similar dimensions to T(16, 256), the optimum location for APC
semiconductor amplifiers was found to be either side of the central space switch [24]. Fibre amplifiers are unsuited for insertion in the fibre delay lines because of the necessary modification for single-polarisation operation [25], the tight tolerance on the optical path length and the slow gain control. Semiconductor amplifiers are more suited to this application particularly if integrated as an array on a single substrate with associated monitor photodiode and bias control electronics. However, the length tolerance on the fibre links will be halved in this case.

The SXR (Table 1) when evaluated for a crosspoint isolation of $-15$ and $-20$ dB equals 4.4 and 14.4 dB, respectively. However, more realistic calculations for architecture $T(16, 256)$ (Section 5.2, Fig. 6) gave a power penalty of $<0.1$ dB for a crosspoint isolation of $<14$ dB. If APC amplification is employed a power penalty close to this value should be realisable.

7 Conclusions

A new type of dilated optical TDM switching network with frame integrity has been proposed, suitable for realisation with large lithium niobate switch arrays. Its performance with regard to such factors as crosstalk, attenuation and control complexity has been considered. The network is suited to high-capacity cross-connect applications and exhibits economy in use of hardware. Modifications to the architecture to reduce the control complexity (and introduce fault tolerance) are possible which would make this type of architecture suitable for use as a switching node rather than as a cross-connection.

The feasibility of building a network with 4096 TDM channels, 16 in the space domain and 256 in the time domain, has been considered. It should be possible to demonstrate such a network experimentally in the laboratory using state-of-the-art integrated optical switch arrays interlinked by polarisation-maintaining fibre delay lines incorporating automatic power control amplification. Careful control of the ambient temperature and the delay line fabrication should enable transmission at 2.488 Gbit/s.

8 References

2 NEWALL, C. (Ed.): 'Synchronous transmission systems' (Northern Telecom, 1991), issue 3
17 LEGG, P.J., and HUNTER, D.K.: 'Feasibility study on realisation of large dilated OTDM switching networks'. Strathclyde University, December 1992

9 Appendix: Proof of theorem 5

This Appendix contains the mathematical proof of theorem 5 from Section 2, which illustrates the steps required to realise TDM switching networks without frame integrity and of any sizes.

Figs. 8 and 9 show how larger rearrangeable networks can be built out of smaller networks [7]; these embody lemmas 1 and 2, respectively. Although these could be used to build large networks, they do not explicitly show how to construct the networks from large substrates.

Fig. 8: Lemma 1, producing $T(n,m)$ from $T(m/2,n)$
Instead, they are used as an intermediate step to producing networks from large substrates. Their validity can be proved [7] by using a theorem due to Kruskal [26].

![Diagram](image_url)

**Fig. 9** Lemma 2, producing $T(m, n)$ from $T(m, n/2)$

Lemmas 1 and 2 are used to prove a more general theorem which shows how to build large networks up from Beneš networks, baseline networks and reverse baseline networks. Before proving this theorem, some preliminary notation and definitions must be introduced. $r(k, i)$ is the modified bit-reversal function. It takes the binary representation of $k$ and reverses the order of the least significant $i$ bits. If $b_u \cdots b_2 b_1 b_0$ is the binary representation of $k$ then

$$r(b_u \cdots b_{i+1} b_i b_{i-1} \cdots b_1 b_0, i) = b_u \cdots b_{i+1} b_i b_0 b_1 \cdots b_{i-1}$$

For example,

$$r(13_{10}, 4) = r(1101_{12}, 4) = 1011_2 = 11_{10}$$

From this definition it is easy to prove lemma 3:

$r(x, i) = 2r(x, i - 1)$, $0 \leq x \leq 2^{i-1} - 1$

$r(x + 2^{i-1}, i) = 2r(x, i - 1) + 1$, $0 \leq x \leq 2^{i-1} - 1$

As a first step to proving the generalised network definition, the validity of Fig. 10 (theorem 4) needs to be proved, where $f$ is an integral power of 2, and $j = \log_2 f$. The proof is by induction on $f$.

First, observe that with $f = 1$, the diagram reduces to a single $T(m, n)$ switch. It then remains to show that Fig. 10 implies the validity of Fig. 11, which can be regarded as equivalent to Fig. 10 with $2n$ and $2f$ substituted for $n$ and $f$, respectively. Note also that on substituting $n$ for $2n$ in Fig. 11, one arrives back at Fig. 10 where the delay lines in the centre stage are multiplied by twice the factor, and the baseline and reverse baseline networks are twice as large.

This proof is carried out by considering two adjacent $B(f)$ networks on the input stage; the proof for the $R(f)$ networks on the right-hand side is a mirror image. Consider two adjacent $B(f)$ networks from the original network of Fig. 10. After lemma 2 has been applied to double the frame size, Fig. 12 is obtained. The order of connection to the centre stage in Fig. 9 does not matter, and this fact is used to advantage here; the way in which the connections have been made in Fig. 12 is crucial to the next part of the proof. To produce Fig. 13, the delay lines of length 1 are first slipped over the lower $B(f)$ network and added onto the delays on the baseline network's output. This would involve advancing the control signals to the lower baseline network by one timeslot. By lemma 3, the delay lines may be redesignated as shown. Then the two $B(f)$ networks and the switches have been combined to form a $B(2f)$ network [3]. Fig. 13 corresponds to Fig. 11, thus completing the proof.

Fig. 1 shows theorem 5; this is a general result that shows how to construct large TDM switching networks by using large substrates. Again, the proof is by induction, with the induction on $e$. With $e = 1$, it is equivalent to theorem 4 (Fig. 10). To satisfy induction, it is necessary to prove that Fig. 1 implies Fig. 14, which is equivalent
to Fig. 1 with $e$ replaced by $2e$. For convenience, we only consider the left-hand side of the network, since the right-hand side is symmetrical. From Fig. 1, and by using lemma 1, half of $T(2em, n)$ appears as shown in Fig. 15.

![Diagram](image)

**Fig. 15** One side of Fig. 14, after using Lemma 1 to double number of inputs and outputs

There are $2m/n$ $B(ef)$ networks in Fig. 15. Pairs of these networks which are separated from each other by $m/n - 1$ other networks form $B(2ef)$ networks along with $ef$ of the $2 \times 2$ switches; this can be verified from the definition of the baseline network [3]. By performing a perfect shuffle on these networks, these pairs of $B(ef)$ networks can be brought together to form $B(2ef)$ networks (Fig. 16). Also, it is easily seen that carrying out the perfect shuffle means that one set of connections go from each new $B(2ef)$ network to each centre stage $T(m, n/f)$ network, making Fig. 16 equivalent to the left-hand side of Fig. 14. Q.E.D.