Optical memory for communications switching

David K. Hunter, University of Essex
Stuart Walker, University of Essex
Michael Parker, Fujitsu Labs

21st September 2005

After a brief overview of optical resonant filter technology, and three-dimensional waveguide technology, which are key candidates for optical memory, this report discusses potential applications in detail, with a clear indication of the trade-offs involved. We assume that in order to find practical applications, optical memory must be organised into delay-lines. Otherwise, for example, if optical memory elements were connected in parallel, addressing would be an insoluble problem. Finally, we outline how work could proceed – assuming future developments in technology – to develop potentially useful optical switching systems for future communications networks.

Optical delay-line and memory technologies

There are two main ways of storing information optically – via resonance effects in optical filters and via conventional delay-lines made from optical fibre or integrated waveguides such as silica. Optical fibre delay-lines are not considered here because their bulk and temperature instability make all but the shortest delays impractical. The main technologies may be categorised as follows:

- **Photonic crystal filters** [1] have been used to implement delays; however there are problems with bandwidth, dispersion and attenuation.
- **Fibre Bragg Gratings** (FBG) [2] are a good example of coupled cavity optical waveguides.
- **Ring resonators** [3] are not suited to being cascaded to produce long delays, however, they can implement small delays of a few bits.
- **Electromagnetically Induced Transparency** (EIT [4]) is presently not mature enough to be used, even for the construction of systems demonstrators in the lab.
- **Three-dimensional optical structures** – their use for improvements in data storage is well-known e.g. [5]. In the context of optical delay lines, spirals [6] and their possible stacking offer an opportunity. However, the use of photonic crystal three-dimensional structures remains an interesting option e.g. [7]. Assuming a simple serpentine meander feed forward fill, 5 micron waveguide and a 50 m path length (2500 bits @ 10 Gb/s) then a guideline cubic dimension of 1 mm is indicated.

Applications

The following suggest themselves as likely application areas for the subsystems we will describe in this report:

- **Optical interconnects for computer systems** – the most likely near-term use of optical interconnects is for board-to-board interconnection within a cabinet or
box, forming part of a commercial digital computer system [8]. Although there is some doubt about such board-to-board interconnects [9], on-board or on-chip interconnects are even less likely to be deployed in the foreseeable future, and are also not such likely applications of optical switching and memory. Reliability and cost are both important concerns, which have so far been impediments to the deployment of this technology.

- **Local Area Networks (LANs)** – Optical LANs usually have a star topology, and are based on a central node consisting of a passive star coupler, and AWG (Arrayed Waveguide Grating) or optical switching. A good example of the latter is being investigated in the SOAPS project [23].

- **Wide Area Networks (WANs)** – WANs are probably the most widely investigated application for fast optical switching, usually based on either OTDM (Optical Time Division Multiplexing) or OPS (Optical Packet Switching).

### Device tradeoffs in a switching subsystem

The switching subsystems considered in this report mostly consist of $2 \times 2$ switches and optical delay-lines, such as the ladder structures shown in Figure 1. Ladder structures can be used for a wide variety of applications, including most of those detailed below. Integration of such repeating sub-system components is advantageous in integrated optical design, such that optical cross-points and delay devices of the type mentioned above fall well within current technology.

A ladder network has two inputs, although, for some applications, only one is used. Likewise, either one or two outputs may be used. The application that is chosen determines the delay line lengths $d_1, d_2, d_3$ etc.

Other more complex structures have been proposed, but ladder networks serve as a good example of these techniques. In terms of device operation, the main factors affecting device operation within these subsystems are:

- **Bitrate in bits per second, $b$**
- **Switching rate in operations per second, $s$** – it is questionable whether switching at the bit rate is desirable, or in many cases, even practical; this is discussed below. Hence ideally we require $s << b$. 

![Figure 1: Ladder networks](image-url)
• **Maximum length of delay in seconds**, \( d \) – for each technology, there is a practical limit on the length of delay-line that is possible, due to factors such as attenuation and dispersion. For the technologies discussed here, this depends on the bandwidth and hence the bitrate.

By introducing two further variables, it will be possible to relate these quantities to each other:

- **Size of unit being switched**, \( u \), in bits. For units of an individual bit, \( u = 1 \). Otherwise, \( u \) can represent the number of bits in a TDM timeslot or an optical packet.
- **Maximum number of such units to be held by a delay-line**, \( k \).

The factors \( b, u \) and \( k \) are determined by the specification of the subsystem. Hence:

\[
\frac{s}{u} = b \quad \text{and} \quad \frac{d}{b} = ku
\]

\( s \) and \( d \) are both constrained by the physical characteristics of the technology, and neither must exceed its limit, for the subsystem is to be realisable. The above formula provides the maximum length of a delay-line in seconds. When measured in bits, the maximum length of a delay-line is \( ku \).

Other factors, such as guard bands, would be considered in reality, but they do not alter the fundamental conclusions. There is an electronic controller which determines the state that each optical switch takes on during each bit period, block period, or packet period.

**TDM switching architectures**

In time-division multiplexing, information is transmitted as “frames”, each composed of a number of equally-sized “timeslots”. The position within the frame of a block of information (i.e. its timeslot number) determines the channel to which it belongs, and hence its destination. The ladder networks shown above can accept one or two such information streams, and rearrange all these incoming streams (each defined by its physical input connection, and its timeslot number) so that they appear on the output connections, on any timeslot defined by the electronic controller. This is a fundamental and crucial function in telecommunications switching.

![Figure 2: An optical TDM switch with 2 inputs, 2 outputs, and switching 4 timeslots per frame. The delays are in units of timeslots.](image)

In this section, we examine the very simple switching subsystem shown in Figure 2 – a TDM switch with 2 inputs, 2 outputs and switching 4 timeslots per frame [11]. This is one of the simplest such switching subsystems that is possible, and we are not claiming that it would necessarily find any real communications application. We study its implementation, with respect to the five parameters discussed above.

Here, optical filters are used as fixed delays but clearly only a total delay of 10 bits is possible with current technology, since this is a reasonable figure for the maximum number of such devices that can be cascaded. We also evaluate the implications of a
total delay of 2500 bits (50 m at 10 Gb/s), in order to see what is possible with the three-dimensional waveguide structures referred to above.

From Figure 2, one would expect that the maximum total delay possible was $8u$, since a total of 8 bits delay exists in the architecture, but in fact the maximum is $7u$, due to the way that the control algorithm is implemented. (This conclusion can be arrived through consideration of the space-time transformation of Figure 2, but this piece of detail is not really relevant to this discussion.) Since we require $7u \leq 10$ due to physical constraints of a 10-bit delay, $u = 1$ is the only possible value, implying bit multiplexing. This means that the $2 \times 2$ switches must change state at the bit rate – probably not desirable or even practical. On the other hand, if 2500-bit delays were possible, then $7u \leq 2500$, i.e. each block of information may contain over 350 bits. This is a useful amount of information for each block, and in fact, if $u$ were made smaller, even larger TDM switching structures could be made, handling a larger number of timeslots per frame.

Hence, even although a particularly simple (perhaps even trivial) architecture was chosen, it is close to the limit of what is possible with today’s resonant optical filter technology. However, the three-dimensional waveguide technology offers the opportunity of constructing useful and practical integrated optical switching systems.

Figure 3: A modified version of the parallel optical TSI architecture.

The so-called “parallel architecture” for timeslot interchange has also been widely studied [12]. This can be adapted to use both fixed and variable delays, as shown in Figure 3. $n$ is the number of timeslots per TDM frame. The demultiplexer consists of optical switches, so that the contents of timeslot 1 are on the top output, timeslots 2 on the next output, and so on. Since the switches in the demultiplexer are implemented a fixed and repetitious pattern it may be easier to make the switches operate at the bit rate than with the ladder structure described above, where the sequence of switch settings is irregular, and liable to change.

After the first stage of delay lines (fixed), the pulses forming the TDM stream are aligned in time, so that they coincide at the same time in each frame. The delays in the second stage (variable) are adjusted electronically to schedule these pulses into the correct timeslots for onward transmission.

If the stage of fixed delay is implemented via some other technology, such as silica delay-lines, then, assuming current technology, the maximum number of cascaded optical filter delays is defined by $n - 1 \leq 10$ hence $n$ can be as large as 11. The total number of resonant filters required in this case is $n(n - 1)$, so if $n = 11$, then 110 resonant filter structures must be fabricated for each TSI.

On the other hand, if both stages of delay lines are implemented via resonant filters, we have $2(n - 1) \leq 10$ i.e. $n$ is limited to 6. The total number of resonant filter structures required is now:
\[ \sum_{d=0}^{n-1} d + n(n-1) = \frac{3n(n-1)}{2} \]

So if there are 6 timeslots per frame, a total of 45 resonant filter structures are required. This is entirely feasible with current microstructure technology.

As before, three-dimensional waveguide technology permits much more practical systems to be constructed, due to the increased number of bits that can be stored, with \( u > 1 \) i.e. many bits in each timeslot.

**Optical packet switching**

In packet switching, each packet contains information (in its header) which defines its destination. Hence there is no “frame” structure, and the packet’s destination is not defined by its position within a frame. The ladder network of Figure 1 can be used as a simple packet switch [13]. With four delay-lines, for example, the sequence would be \( d_1 = 1 \) packet, \( d_2 = 2 \) packets, \( d_3 = 4 \) packets, \( d_4 = 8 \) packets. Adding more delay lines provides more packet buffering.

Since each packet would be hundreds or even thousands of bits in length, \( u \) would also be of the same size. The length of delay-line required would be far too long for this type of technology i.e. it is not viable, since the maximum of 10 bits delay rules out switching blocks of packets.

With 2500 bits of delay, useful optical packet switches could be made with small packets (less than 250 bits) and small optical buffer depths. As we shall see, smoothing of the data at the edge of the network will permit small optical buffers to be used in the core network, and this approach has in fact been investigated by several groups, especially SOAPS, in which Essex and Intel are both involved. Various options leading to this conclusion are discussed below.

One approach, which has been considered in the past, involves emulating most of the functionality of an electronic IP router in optics [10]. This requires delay-lines of many millions of bits in length, which is clearly not possible with the technology discussed in this paper, given the limitation of 10 bits’ delay, or, in fact, even 2500 bits’ delay.

Hence, it is generally agreed that trying to build an IP router with entirely optical buffers is not a sensible idea. However, there has recently been some research on sizing router buffers which makes it worthwhile to re-evaluate this conclusion. Conventional wisdom states that the buffer on each router port should be RTT × the data rate. Currently, router linecards can have as much as 10 Gbits of memory.

Appenzeller et al [14] showed that this can be divided by \( \sqrt{n} \), where \( n \) is the number of TCP flows through the port. Furthermore, Enachescu et al [15] showed that if certain minor modifications were made to TCP, the memory requirement for a linecard can be further reduced to \( O(\log W_{\text{max}}) \), where \( W_{\text{max}} \) is the maximum window size. Both these studies assume that all traffic is TCP, although they both mention that a small amount of UDP will not affect their conclusions.

Enachescu’s results suggest that an IP router could be built with a buffer depth of 12. Assume that each datagram is 1500 bytes long i.e. the maximum size allowed by Ethernet. (Even our longer delay of 2500 is not nearly big enough to hold a single packet: \( 1500 \times 8 = 12,000 > 2500 \)). Then the maximum delay requirement will be \( 12 \times 1500 \times 8 = 144,000 \) bits. Optical fibre stores 50 bits per metre at 10 Gb/s, so
2.88 km would be required to store 144,000 bits. Use of multiple DWDM wavelengths would require a shorter length of fibre to store the required bits, but this requires extra complexity.

Besides TCP, the other most important transport layer protocol running over IP is UDP, although there are others. Many different application layer protocols can run over UDP, including RTP, RIP, SNMP, echo, NTP, DNS, BOOTP and daytime. All of these can be regarded as producing Poisson traffic, except for RTP, which carries voice and video traffic. We will see that by imposing certain restriction on the traffic that applications using RTP generate, the original buffer depth of \( O(\log W_{\text{max}}) \) may be retained.

The results derived by Enachescu [15] may be extended to UDP traffic, however the UDP traffic must be paced, that is, the time between consecutive UDP packets from a UDP flow must be no less than \( O(\log W_{\text{max}}) \). \( W_{\text{max}} \) is the TCP maximum window size. This satisfies Enachescu’s theorem 2 [16]. (Theorem 1 is concerned with the dynamics of TCP traffic, which does not apply to UDP.)

In any case, TCP is very much the dominant network traffic type, for example, in February 2000, measurements from a single site showed that TCP represented 18 times more traffic in bytes than UDP [17].

Electronic buffering at the network edge is also a viable way of reducing optical buffering requirements in the network [18]. With this approach, the data is segmented into fixed-length slots, with only 10 slots’ storage per core switch port being necessary. If the correct trade-offs were made between optical device technology, networking, and systems engineering, this would imply a packet size of 250 bits (less in practice due to guard bands), which seems to offer reasonable opportunities for building a practical system.

### Optimising optical storage capacity

The above discussion has assumed that single delay-lines are used, and that 1 bit/s represents exactly 1 Hz of bandwidth. Here, we examine these assumptions, and show how greater capacity than the 10 bits discussed above could be achieved.

#### Spectrally efficient coding techniques

Spectrally efficient coding schemes such as ASK (Amplitude Shift Keying) and PSK (Phase Shift Keying) offer the possibility of increasing the number of bits/s per Hz, and thus increasing the capacity in bits of a delay-line.

A recent tutorial paper by Gnauck et al [19 – Table II] states that, even for 40 Gb/s DPSK (Differential Phase Shift Keying), the highest efficiency that has been achieved experimentally is 0.8 bits/s/Hz. (10 Gb/s experiments generally have lower spectral efficiency). Theoretical studies show that in order to obtain a much higher efficiency of over 2 bits/s/Hz, techniques such as 8-level PSK (phase Shift Keying), 8-level QAM (Quadrature Amplitude Modulation) or 16-level QAM are required [20].

For example, an experiment with 16-level transmission at 10 Gsymbol/s was demonstrated recently by Hitachi [21]. There were 4 levels in the amplitude domain, and 4 in the phase domain. The optical hardware required to implement such coding schemes is relatively complex, and it is not clear whether it would be justified, unless such a spectrally efficient coding scheme were already used for transmission over links between nodes themselves. Furthermore, this would probably only offer a
benefit for structures based on resonant filters, rather than optical delay-lines, which have a much wider bandwidth.

Parallel delay-lines

Multiple structures (such as that of Figure 2) may be placed in parallel, forming several parallel “planes”, with a corresponding reduction in the bitrate carried and capacity of each delay-line. For example, if four switching systems were put in parallel, the bitrate in each plane would be a quarter of the original, and each delay-line would hold a quarter as many bits as the original. While this would increase the amount of storage that is possible while still adhering to our constraint of 10 bits’ delay in series, there would the added cost and complexity of the additional optical switching devices required to implement multiple planes. Furthermore, multiplexing and demultiplexing hardware would be required at the edge of the switch fabric to convert the incoming bit-streams to the required parallel format, and vice-versa.

Conclusions

In this report, we have considered some of the more interesting and challenging applications of optical memory. Several other excellent applications have not been considered because they are already well under investigation elsewhere:

- Fine synchronisation in optical TDM and optical packet switching.
- Optical CDMA [22]
- Multiplexing and demultiplexing of optical TDM
- Dispersion compensation
- Variable optical delay-lines for radar and phased-array antennae

It is clear, especially given the discussion above, that photonic memory technology has severe limitations, and that single-bit delays are not the best fundamental unit for constructing memory systems, due to the cascadability limitations we described.

If three-dimensional waveguide structures can indeed realise delays of 2500 bits, then it is possible that by means of careful architectural design, useful optical switching subsystems could be developed. For example, if 10 packets’ delay were required [18], this would imply that each packet could be no more than 250 bits (approximately 15 bytes) long, with implications for switching speed and network protocol design. If such an OPS network were designed, it would be based upon the following principles:

1. Use of delay-line architectures, perhaps in a recirculating architecture, to facilitate regeneration.
2. Reducing the buffer depth required for TCP, or, more likely, using electronic edge smoothing with slotted packets in the core, as in the SOAPS project.
3. Perhaps using multi-level coding, as it may have the potential the double to capacity in bits of optical memory.
References


