Buffering Strategies for Optical Packet Switches

David K. Hunter, Meow C. Chia, Mohamed H. M. Nizam, Ivan Andonovic
Electrical & Electronic Engineering Department, University of Strathclyde, 204 George Street, Glasgow G1 1XW, UK
Phone: +44 141 548 2527 Fax: +44 141 553 1955
d.hunter@eee.strath.ac.uk

Bocang Qiu, A. Catrina Bryce, Yahong Qian, J. Stewart Aitchison, John H. Marsh
Department of Electronics & Electrical Engineering, University of Glasgow, Glasgow G12 8LT, UK

Ken M. Guild, Mike J. O'Mahony
Department of Electronic Systems Engineering, University of Essex, Colchester CO4 3SQ, UK

Abstract: Recently, many optical packet switches have been proposed, to overcome to potential problems of future large electronic switch cores. The need for buffering arises due to the unscheduled nature of packet arrivals at the switch inputs, and several strategies have evolved to implement this buffering. Examples of these are discussed in this paper, including the use of wavelength to assist in contention resolution; where contending packets can usually be transmitted on different wavelengths. Also, the implementation of multi-stage buffers is discussed, using differing technologies, both with and without deflection routing.
OCIS codes: (060.2330) Fiber optics communications; (060.4250) Networks

Introduction

Optical packet switching has been researched for a number of years, in anticipation of solving the EMI and pinout problems that are likely to be encountered in future large electronic packet switch cores. In order to overcome these problems, optical packet switches have an entirely optical data path, although they are electronically controlled. The electronic controller operates at a rate equal to the number of fixed-length packets per second on each link, rather than the optical line bitrate, facilitating its implementation.

A key issue is the manner in which packet contention for each output is resolved [1], since packets arrive at different switch inputs asynchronously. To surmount the problem of contention, several buffering strategies have emerged for fixed-length packets:

- *Deflection routing* – here there are no buffers, and packets are deflected to the wrong output in the case of contention [2]. They are left to reach the destination node by an alternative route.
- *Use wavelength to reduce the amount of buffering* [3,4] – this takes place in the Wavelength Switched Packet Network. Statistical multiplexing between wavelengths reduces the buffering requirement.
- *Use small packet switching modules with a moderate amount of buffering and deflection routing* [5,6], as in Photonic Integrated Gigabit Switches.
- *Imitate electronics by implementing deep buffers*, as in the Switch with Large Optical Buffers [7].

This paper will consider the latter three strategies in turn, discussing an approach to implementing each, and describing work that is taking place their implementation.

A major constraint when implementing these switches is the lack of optical random access memory (RAM). Electronic buffers are often implemented with RAM, and since this technology is not available in optics, it is necessary to implement buffering using delay lines.

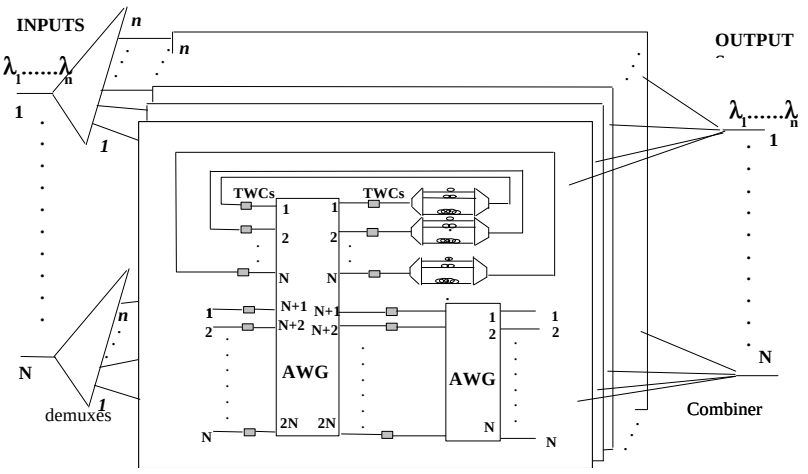

Figure 1: The WASPNET switch. Each plane handles inputs from a different wavelength, although each plane may output packets on more than one wavelength.

Wavelength Switched Packet Network (WASPNET)

In the EPSRC-funded WASPNET project, WDM is used to transport packets between nodes, and also to facilitate contention resolution; each packet is dynamically converted to a suitable wavelength in each link. This decision is influenced by the availability of free timeslots on each wavelength when forwarding to the next node. Suppose there are 16 wavelengths within a particular fiber link carrying a number of optical paths. Every packet directed along the fiber can be transmitted on any of the 16 wavelengths; contention between any two packets being transmitted simultaneously is resolved in the first instance by transmitting them each on a different wavelength. If this fails due to all wavelengths being occupied, buffering is employed to delay some of the packets until a free timeslot is available. Employing wavelength to resolve contention in this way reduces the required buffer capacity, for example at a packet loss rate of 10^{-5} and a load of 0.9, this scheme offers a buffer size reduction of over three times compared to other methods. Analytical studies indicate that a packet loss as low as 10^{-12} may be obtained at a load of 0.7 per wavelength and a buffer depth of only 9 per wavelength [4].

A switch architecture has been designed to implement this contention resolution strategy (Figure 1), with a novel aspect being the use throughout of feedback fiber delay lines with arrayed waveguide gratings – AWGs. Feedback delay lines allow the implementation of multiple packet priorities, while wavelength routers exhibit low loss and low crosstalk, offering superior systems performance. The architecture can be said to implement a shared buffer, since each delay line is not associated with a particular output, but implements buffering for all outputs. Further details of the architecture may be found in Reference 4.

A photonic testbed is currently being constructed that allows the WASPNET architecture to be evaluated in terms of control complexity and optical performance, whilst minimizing the necessary hardware required for the routing and buffering functionality (Figure 2). The arrangement emulates a 4×4 switching node by providing a fully functional optical packet path and three sub-equipped electrical paths. The optical packet generator produces a pre-determined traffic distribution such that the effects of uniform and bursty traffic can be physically evaluated on the optical path. At present, a packet comprises a 2.5Gbit/s payload and an eight-bit 155Mbit/s header.

The remaining three inputs are emulated by three electrical header generators which similarly produce a sequence of headers into the control and arbitration module thus allowing various traffic scenarios to be simulated. In this way, various header implementations and packet formats can be investigated by changing only a single optical packet generator and header decoder module. By varying the input traffic distributions, the effect on the monitored optical packet path and the effectiveness of the control and arbitration process can be assessed.

To investigate the effect of cascading multiple routing nodes, the switching architecture is inserted into a re-circulating loop. The principle of operation is similar to a circuit-switched loop in that a fixed length of data is passed through a fixed chain of amplifiers a number of times to simulate a

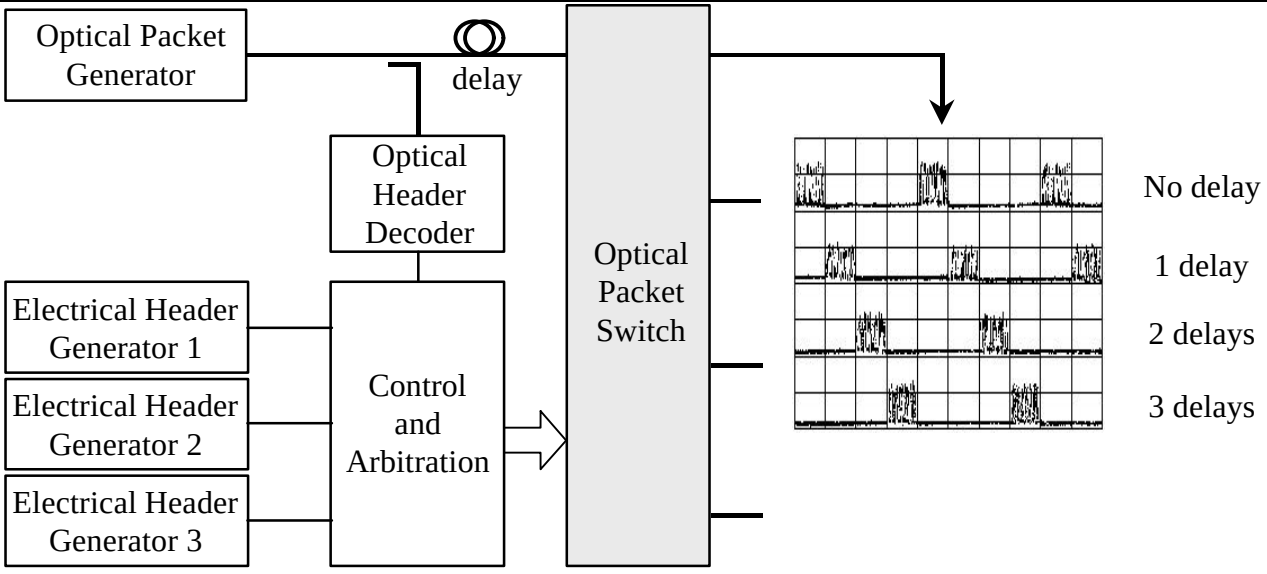

Figure 2: WASPNET testbed showing the output of the node when optical input packets undergo up to three packet delays

longer system. The packet-switched version allows the operator to pre-determine the path of an individual packet through the switching matrix of each node and monitor it at various points throughout the network. The configuration is shown in Figure 3.

The operator is able to specify how many nodes are to be traversed and the respective switch paths to be taken at each circulation. A particular packet is selected for monitoring (identified by a specific header address) and this is transmitted together with arbitrarily addressed dummy packets to fill the loop; once the loop is filled, the optical packet generator ceases transmission. At the node, the header of each incoming packet is decoded and removed by the header detection and removal module, and the packets are routed through the photonic switching matrix according to the routing table within the control electronics. The individual output of the switching matrix, which would normally be separate output ports in a real node, are combined and coupled to the header insertion module. A new header is inserted according to the routing table and the packet exits the node. The packet stream is fed back into the loop and also to the monitoring instruments. By appropriate gating of the instruments, the eye diagram, error performance or optical spectrum can be obtained at the output of any specified node in the chain. The process is repeated a number of times and averaged over the duration of the experiment to achieve realistic results.

The operator is able to specify how many nodes are to be traversed and the respective switch paths to be taken at each circulation. A particular packet is selected for monitoring (identified by a specific address) and this is transmitted together with arbitrarily addressed dummy packets to fill the loop; once the loop is filled, the optical packet generator ceases transmission. At the node, the header of each incoming packet is decoded and removed by the header detection and removal module, and the packets are routed through the photonic switching matrix according to the routing table within the control electronics. The individual output of the switching matrix, which would normally be separate output ports in a real node, are combined and coupled to the header insertion module. A new header is inserted according to the routing table and the packet leaves the node. The packet stream is fed back into the loop and also to the monitoring instruments. By appropriate gating of the instruments, the eye diagram, error performance or optical spectrum can be obtained at the output of any specified node in

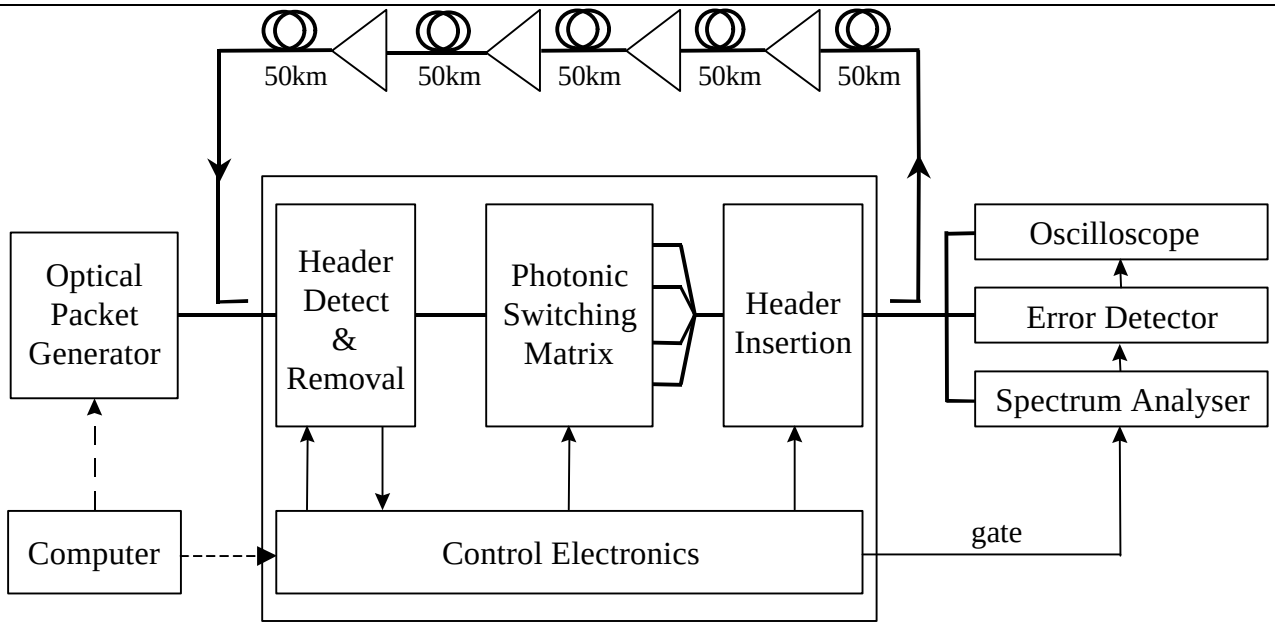

Figure 3: Packet-switched re-circulating loop configuration

the chain. The process is repeated a number of times and averaged over the duration of the experiment to achieve realistic results.

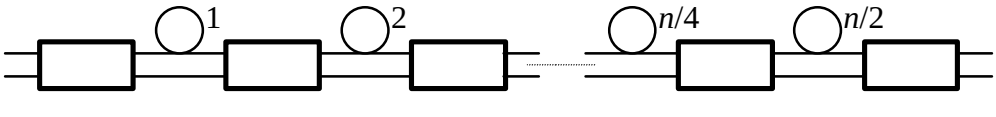

Figure 4: 2x2 switch architecture; the buffer depth rises exponentially with the number of switches.

Photonic Integrated Gigabit Switches (PIGS)

This switch architecture (Figure 4 [6]) implements much the same function as an output-buffered 2×2 optical packet switch, the difference being that the delay may be non-optimal. It is designed to implement a moderate buffer depth, and to operate in conjunction with a limited amount of deflection routing, which occurs if a buffer overflows. The delay-lines between switch devices increase in powers of two, starting at one timeslot, and the size of the shared buffer is equal to the sum of all the delay-line lengths i.e. $2^{k-1} - 1$, where k is the number of switches. One possible application is for add/drop in a ring network. As part of the EPSRC-funded PIGS program, the buffered switch is being fabricated on a silicon substrate with silica delay-lines, which may be up to 4m in length [8].

The 2×2 switch devices have been fabricated from quantum well epitaxial structures based on InGaAs/InGaAsP/InP laser structures and operate in the communications wavelength range around 1.55μm. The switches consist of 2×2 multi-mode interference (MMI) couplers integrated with amplifiers and modulators (Figure 5). The MMI couplers are designed to split a light input to either of its two input ports equally between its two output ports, while the amplifiers compensate for coupling and interface losses, as well as losses in the passive sections of the III-V semiconductor chip. Electroabsorption modulators are used to modulate the signals, since they operate faster than amplifiers.

A further advantage of this arrangement is that the amplifiers are run CW, so imposing a constant thermal load on the chip.

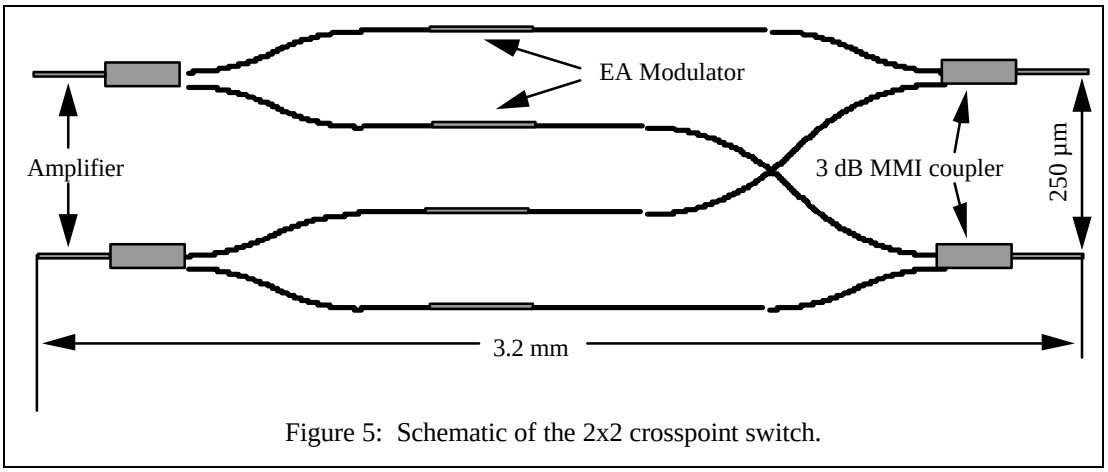

Figure 5: Schematic of the 2x2 crosspoint switch.

The fabrication of the monolithically integrated 2×2 switches requires the realization of three different bandgaps on a single substrate. The optimum bandgap for the modulators is larger than that required for the amplifiers and smaller than that required for the low loss waveguides. In order to engineer these three bandgaps after growth, quantum well intermixing (QWI) has been carried out to tune the bandgap in selected areas of the device chips. Using a sputtered SiO_2 technique (Figure 6), involving the deposition of a thin film (~200nm) of sputtered SiO_2 and subsequent annealing using a rapid thermal annealer (RTA), a reliable means for obtaining post-growth shifts in the band edge of a wide range of III-V material systems can be achieved [9]. During the sputtering stage point defects are generated on the surface of the exposed semiconductor; during a subsequent anneal these point defects diffuse down through the epilayer inducing QWI, the degree of intermixing depending on the initial concentration of point defects. Photoresist protects the surface preventing the formation of point defects, while SiO_2 layers provide partial protection (Figure 6), reducing the number of point defects that are formed during the sputtering process. Samples with sputtered SiO_2 caps start to exhibit bandgap shifting at significantly lower temperatures than those required for thermally induced intermixing of the material. For temperatures >650 C, it is possible to obtain two different bandgap shifts in one annealing stage. For this application, the blue-shift in the modulator section should be ~30nm-50nm compared to the bandgap of as-grown material, while a larger blue-shift is required for the passive waveguide components.

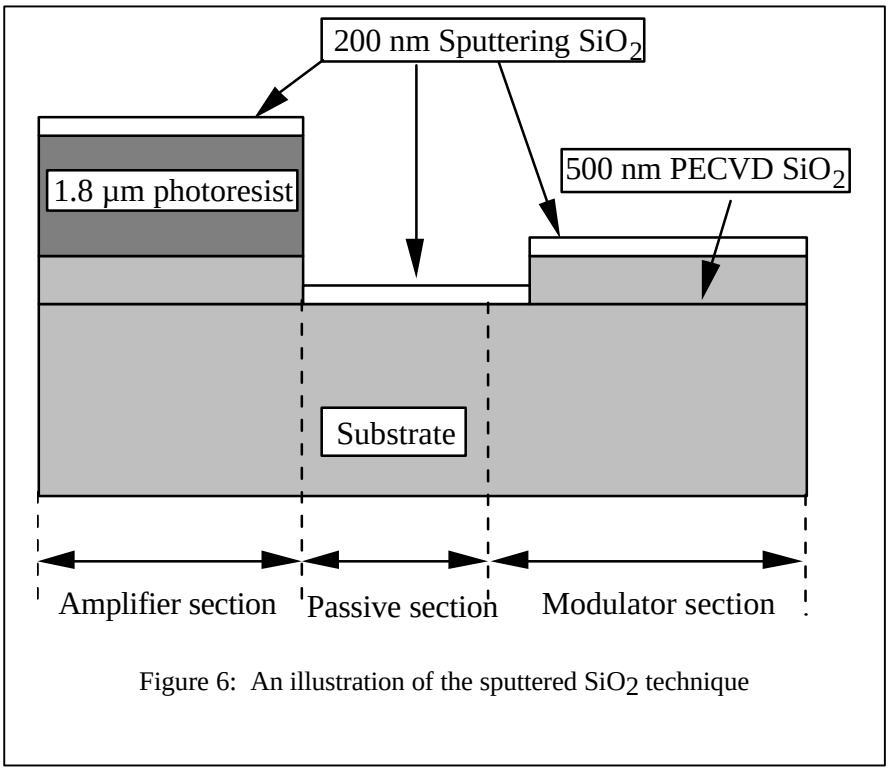

Figure 6: An illustration of the sputtered SiO_2 technique

The processing stages begin by depositing a layer of SiO_2 using plasma enhanced chemical vapour deposition (PECVD). This layer is then patterned, using photolithography and etching of the SiO_2, to expose the regions of the sample where low waveguide losses were required whilst the modulator section of

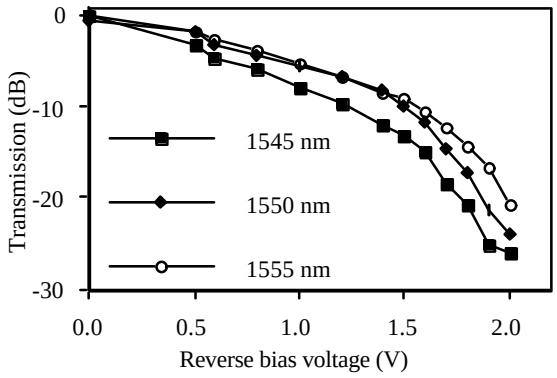

Figure 7: The modulation depth of the modulators in the 2x2 crosspoints.

the sample remained protected with 500nm of PECVD SiO_2. The samples were then annealed at a temperature of 720 C. A blue-shift of 40nm occurs for the modulator region, while the blue-shift in the passive region is around 102nm, which together meet the requirements for the modulator and passive sections. Finally, the MMI couplers and waveguides were defined using a single photolithographic step and a Ti/Si_3N_4 mask. The devices were reactively ion etched in a $CH_4:H_2:O_2$ plasma [10].

Initial performance characterization of the electro-absorption modulators has been performed. The light from a tunable semiconductor laser, emitting wavelengths between 1480nm-1580nm, was endfire coupled into the device. A polarization controller was used to ensure that only the TE mode was excited in the waveguide. The output power of the device was monitored as a reverse DC bias, between 0Volts and 2.5Volts, was applied to a modulator. Figure 7 shows the modulation depth as a function of reverse bias voltage for the modulators within the crosspoint geometry.

Although no direct measurement of the losses in the fabricated switches has been performed as yet, an estimate of the loss in the switch is possible, by comparing the losses in a straight waveguide with those of the switches. The straight waveguides, which had the same structure as that of the switches, fabricated using an identical process, were measured to have losses of 11±3 $dB.cm^{-1}$. From this result, it can be deduced that the insertion loss of the whole switch is around 9.6 dB, which includes the 3dB losses due to the MMI couplers. The fabrication process can easily allow the inclusion of amplifiers in the switch; the optical gain of the amplifiers can compensate for the measured insertion

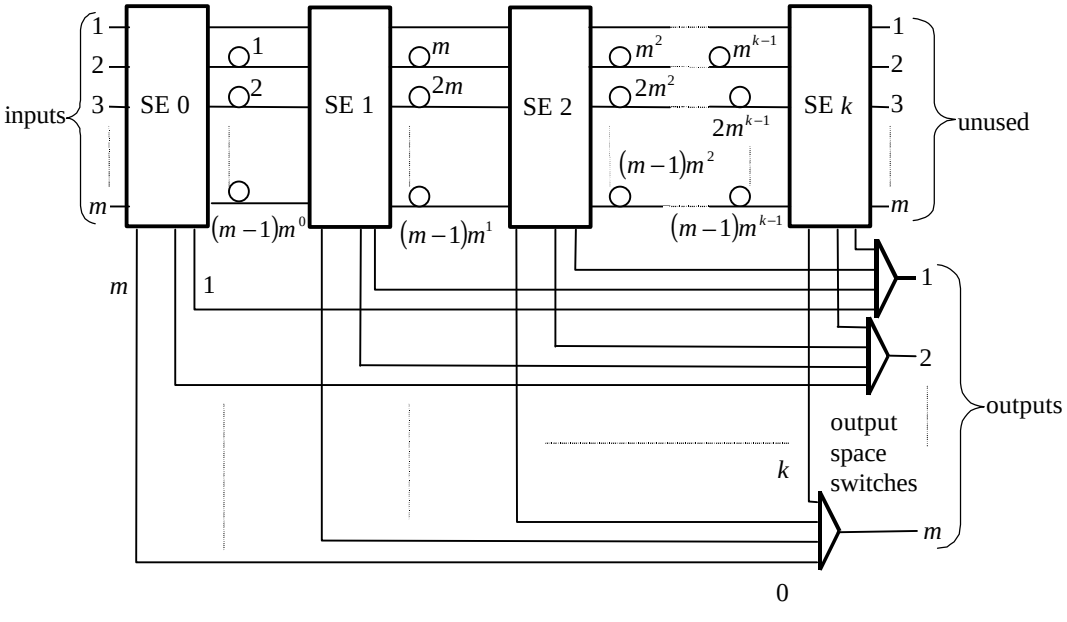

Figure 8: SLOB architecture; each stage is a modified OASIS switch (SE or switch element).

losses.

Switch with Large Optical Buffers (SLOB)

A final approach involves implementing deep buffers, in order to attain the buffer depth required for bursty traffic. This is implemented by the Switch with Large Optical Buffers (SLOB [7] – Figure 8) which cascades many small switch elements, forming a larger switch with a greater buffer depth. A modified Alcatel OASIS switch [11] is chosen as the basic switch element, although other switches could be used. SLOB is electronically controlled but has an optical packet data path, with *m* inputs and outputs. The delay line lengths increase exponentially from left to right along the structure, and each buffer has a depth of $m^k - 1$ packets, where *k* is the number of stages in the architecture, hence the number of stages increases with the logarithm of buffer depth. When a packet has experienced the correct delay, it is sent to the appropriate output space switch via a lower output of a stage. SLOB emulates an output-buffered switch, thus exhibiting optimal delay/throughput performance. Experiments have shown that 40 similar stages may be cascaded with the aid of optical regeneration [12], further justifying this concept.

Conclusions

The three major buffering strategies for optical packet switching have been identified, and examples of work in each have been discussed. These were based on using statistical multiplexing among many wavelengths to relax buffering requirements, using deflection routing in conjunction with some buffering, and implementing deep buffers. It is certainly too early to say which technique will predominate, however it is likely that WDM will be used in conjunction with optical packet switching in order to enhance the link capacity.

References

1. D. K. Hunter, M. C. Chia, I. Andonovic: "Buffering in Optical Packet Switches", *IEEE/OSA Journal of Lightwave Technology*, vol. 16, no. 12, December 1998, pp 2081-2094
2. D. J. Blumenthal, R. J. Feuerstein, J. R. Sauer: "First Demonstration of Multihop All-Optical Packet Switching", *IEEE Photonics Technology Letters*, vol. 6, Mar. 1993, pp. 457-460
3. S. L. Danielsen, C. Joergensen, B. Mikkelsen, K. E. Stubkjaer: "Analysis of a WDM packet Switch with Improved Performance Under Bursty Traffic Conditions Due to Tuneable Wavelength Converters", *IEEE/OSA Journal of Lightwave Technology*, vol. 16, no. 5, May 1998, pp729-735
4. D. K. Hunter, M. H. M. Nizam, M. C. Chia, I. Andonovic, K. M. Guild, A. Tzanakaki, M. J. O'Mahony, J. D. Bainbridge, M. F. C. Stephens, R. V. Penty, I. H. White: "WASPNET: A Wavelength Switched Packet Network", *IEEE Communications Magazine*, vol. 37, no. 3, March 1999, pp120-129
5. I. Chlamtac, A. Fumagalli, L. G. Kazovsky, P. Melman, W. H. Nelson, P. Poggiolini, M. Cerisola, A. N. M. M. Choudhury, T. K. Fong, R. T. Hofmeister, C.-L. Lu, A. Mekkittikul, D. J. MN. Sabido IX, C.-J. Suh., E. W. M. Wong: "CORD: Contention Resolution by Delay Lines", *IEEE Journal on Selected Areas in Communications*, vol. 14, June 1996, pp1014-1029
6. D. K. Hunter, D. Cotter, R. B. Ahmad, W. D. Cornwell, T. H. Gilfedder, P. J. Legg, I. Andonovic: "Buffered Switch Fabrics for Traffic Routing, Merging and Shaping in Photonic Cell Networks", *IEEE/OSA Journal of Lightwave Technology*, vol. 15, no. 1, Jan. 1997, pp86-101
7. D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, I. Andonovic: "SLOB: A Switch with Large Optical Buffers for Packet Switching", *IEEE/OSA Journal of Lightwave Technology*, vol. 16, no. 10, October 1998, pp1725-1736
8. Y. Hida *et al*, *Integrated Photonics Research*, 1995, Dana Point, USA, paper IThC6
9. J. H. Marsh, O. P. Kowalski, S. D. McDougall, B. C. Qiu, A. McKee, C. J. Hamilton, R. M. De La Rue, A. C. Bryce: "Quantum Well Intermixing in Material Systems for 1.5μm", *J. Vacuum Science and Tech. (A)*, vol. 16, no. 2, March-April 1998, pp. 810-816

10. Y. H. Qian, A. C. Bryce, J. H. Marsh, C. D. W. Wilkinson, M. Owen, R. V. Penty, I. H. White, M. J. Robertson: "Fabrication of Test Structures for 4×4 InP/InGaAsP Integrated Optical Matrix", *CLEO Europe 1998*, 14-17 Sept 1998, pp52-
11. J. M. Gabriagues, J. B. Jacob: "OASIS: A High-Speed Photonic ATM Switch – Results and Perspectives", *15th International Switching Symposium*, Berlin, Apr 1995, vol. 2, paper C8.4, pp457-461
12. D. Chiaroni, B. Levigne, L. Hamon, A. Jourdan, F. Dorgeuille, C. Janz, E. Grard, M. Renaud, R. Bauknecht, C. Graf, H. P. Schneibel, H. Melchior: "Experimental Validation of an All-Optical Network Based on 160Gbits/s Throughput Packet Switching Nodes", *24th European Conference on Optical Communications (ECOC) '98*, September 20-24 1998, Madrid, Spain