# Lookahead Forward Shift Optical Packet Switch

Stephen Suryaputra and Joseph D. Touch
USC/ISI
Marina del Rey, CA, USA
surya@isi.edu and touch@isi.edu

Joseph A. Bannister
Aerospace Corp.
El Segundo, CA, USA
joseph.a.bannister@aero.org

***Abstract*—This paper explains an optical buffer called lookahead forward shift buffer and its application to an optical packet switch. The switch supports variable length packets, and achieves performance similar to electronic virtual output queued switching, using four packets's worth of variable-speed shift-forward delay for a 1 Gbps/port, 32x32 switch.** *(Abstract)*

## I. INTRODUCTION

Lookahead forward shift optical packet switch is a single wavelength, all optical asynchronous packet switch with an electronic controller that buffers/delays all the packets by default, shifting packets forward so they can be lined up contention free to the output ports.

The key problems in all optical switches are constructing buffers and designing scheduling algorithms. There are limited options for optical buffers. By delaying all packets our switch can examine beyond the head of the line. As a result, the switch can decide to take packets from the buffer to line them up contention free. All this is done while the packets are still propagating through the optical medium. Our switch inherently accommodates variable length packets.

## II. PRIOR WORK

All optical packet switches can be differentiated by their contention resolution mechanism, which in turn depends primarily on their buffing.

In *sequential buffering*, packets are buffered once, as with a single FDL or parallel FDLs of different lengths [1], non blocking OTSI [2], feed-forward FDLs [3] used in SDL [4] and SLOB [5] The output time of a packet is committed when it enters the buffer. *In our switch the packet delays can be determined after they are buffered.*

In *recirculating* buffering, packets can be buffered multiple times, as in the SMOP switch [2, 6, 7]. Here the output time of the packets can be augmented when necessary by recirculating the packets back to the buffer. Such switches introduce substantial delay variation and reordering. *Our switch never increases buffering delay. It can even be viewed as decreases the delay.*

In most optical packet switches *packets enter the buffer because there is a contention at that time. Our switch buffers all packets by default, preserving their input line information, and put the packet in contention free order.* Further, in most switches the scheduling decision is made on each packet, either when there is a packet in a slot (synchronous) or when a packet arrives (asynchronous). *Our switch uses batch scheduling, commencing a fixed time after the first packet enters the switch and schedules all buffered packets at that time.*

Another distinction is in how packets are scheduled. In Optical burst switching (OBS) *schedules bursts (a sequence of back to back packets), and we are scheduling individual packet* [8]. OBS also uses a control packet that is sent before the actual burst. Such a control packet is not necessary in our switch because *such a gap is already provided by the lookahead delay*.

## III. LOOKAHEAD FORWARD SHIFT

Our switch has several components, as in Figure 1:

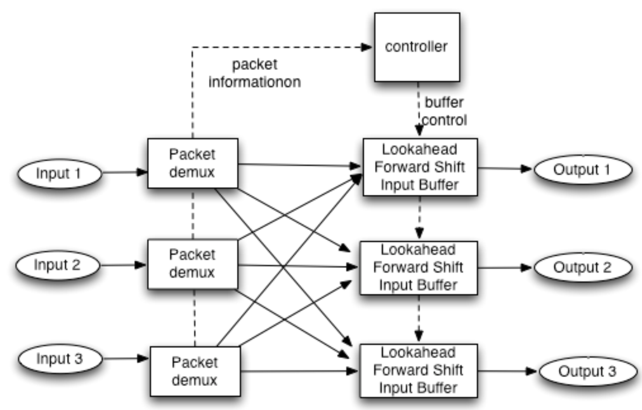


**Figure 1 – 3x3 Lookahead Forward Shift Optical Packet Switch**

- *Packet demux* that routes to the output multiplexers.This component is typical in any optical packet switch.
- *Lookahead forward shift optical buffer;* this is the defining component of the switch.
- *Controller*, that performs contention resolution and shifts selected packets forward to the output port.

The lookahead forward shift buffer is the key to our switch. The buffer is divided into two regions: a look-ahead region and a shifting region (Figure 2). The lookahead region is where the switch controller examines the packets and determines the order. When the first packet arrives at the shifting region the controller sends a signal to the demux to forward shift packets so that they line up properly for the output port. In the next period, the lookahead region becomes the shifting region.

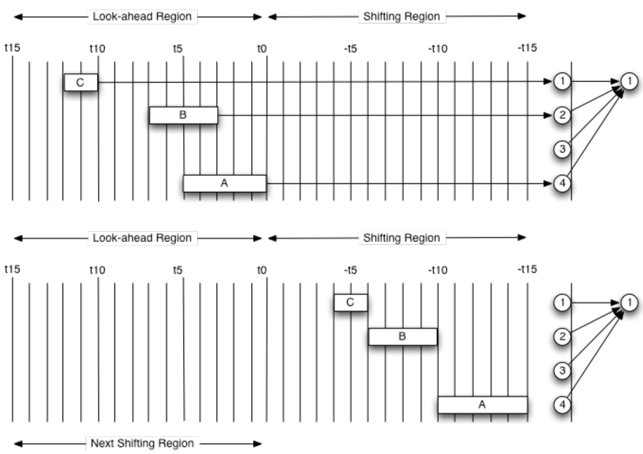


**Figure 2 - Lookahead and shifting regions**

We designed an initial scheduling algorithm known as *earliest fit first* and it achieves FIFO behavior with a simple operation. One controller on each output port keeps information on packets as an ordered set, sorted by arrival time. When the packet with the earliest arrival time hits the shifting region, that packet is shifted forward. The next earliest packet is shifted at the tail of the first packet. The algorithm continues until the region is scheduled. The ordered packets are then sent to the output port. This simple batch algorithm is easy to implement. Other algorithms can also be used.

## IV. EVALUATION

We evaluated the throughput performance of a 32×32 switch with FIFO scheduling under quasipoisson and pareto on-off arrivals with uniform output port distribution. We define quasipoisson arrival as a modified poisson arrival process that will not put overlapping packets on the same input port. The shape parameter $\alpha$ for pareto arrivals is set to 1.2 [9] and the mean burst size is 12500 bytes [8]. The lengths of the packets are variable and have bimodal distribution. 80% of the packets are 40 bytes long and the rest are 1500 bytes., based on recent observations [10]. Other performance evaluation is reported in detail elsewhere [11].

Figure 3 shows the result when we set the lookahead region size equal to the shifting region size. We chose 50Kb regions so that uniform traffic can achieve 100% throughput. This size represents approximately 4 packets of delay.

We compare our performance with a VOQ switch with modified Parallel Iterative Matching (PIM) modified for variable length packets [12], chosen due to its reasonably well throughput performance. The results are shown in two groups of lines plotted in Figure 3. The upper group is for quasipoisson arrivals. The lower one is for pareto arrivals. The top graph in each group is the throughput of the PIM switch. Since the packets are *variable length*, the number of bytes from one input port to an output port is not exactly $1/N$. Hence, the shifting region size is not big enough. Therefore, we observed less than 100% throughput in the first group. In the second group, more severe drop is observed on pareto traffic due to the same reason. When we use discrete shifting points, obviously there are more drops. These remain issues that we need to explore. The graph also shows continuous shifting to any bit boundary, and the impact of shifting only to 320 or 640-bit boundaries

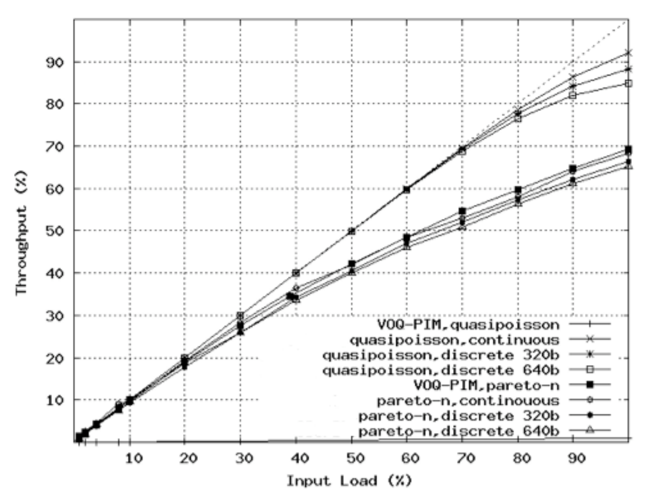


**Figure 3 -The throughput of our 32x32 lookahead switch**

## V. FUTURE WORK

We are currently planning to explore a variety of scheduling algorithms, the impact of traffic on the size of the lookahead and shift-forward regions, and how to handle hot-spot traffic bursts.

## REFERENCES


[1] Haas, Z., *The "Staggering Switch": An Electronically Controlled Optical Packet Switch.* Journal of Lightwave Technology, 1993. **11**(5/6): p. 925-926.

[2] Liew, S.Y., G. Hu, and H.J. Chao, *Scheduling Algorithms for Shared Fiber-Delay-Line Optical Packet Switches - Part I: The Single-Stage Case.* J. Lightwave Technol., 2005. **23**(4): p. 1586.

[3] Gee-Kung, C., et al., *Enabling Technologies for Next-Generation Optical Packet-Switching Networks.* Proceedings of the IEEE, 2006. **94**(5): p. 892-910.

[4] Chlamtac, I. and A. Fumagalli, *An Optical Switch Architecture for Manhattan Networks.* IEEE Journal on Selected Areas in Communications, 1993. **11**(4): p. 550-559.

[5] Hunter, D.K., et al., *SLOB: a switch with large optical buffers for packet switching.* Lightwave Technology, Journal of, 1998. **16**(10): p. 1725-1736.

[6] Karol, M.J. *A Shared-Memory Optical Packet (ATM) Switch.* in *Local and Metropolitan Area Networks, 1993., Proceedings of the 6th IEEE Workshop on*. 1993.

[7] Xiaoliang, W. *Maintaining Packet Order in Reservation-Based Shared-Memory Optical Packet Switch*. 2008.

[8] Qiao, C. and M. Yoo, *Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet.* Journal of High Speed Networks, 1999. **8**(1): p. 69-84.

[9] Crovella, M.E., M.S. Taqqu, and A. Bestavros. *Heavy-tailed Probability Distributions in the World Wide Web*. 1998: Chapman and Hall.

[10] Sinha, R., C. Papadopoulos, and J. Heidemann, *Internet Packet Size Distributions: Some Observations*. 2007, USC/ISI.

[11] Suryaputra, S., J.D. Touch, and J.A. Bannister, *The Tetris Switch*. 2009, USC/ISI-TR-662: Marina del Rey, CA.

[12] Ge, N., M. Hamdi, and K.B. Letaief. *Efficient scheduling of variable-length IP packets on high-speed switches*. in *Global Telecommunications Conference, 1999. GLOBECOM '99*. 1999.