Slot Timing Considerations in Optical Packet Switching Networks

Benjamin A. Small, *Student Member, IEEE,* and Keren Bergman, *Member, IEEE*

***Abstract*—Slot timing margins and latency characteristics are investigated experimentally in a fully implemented 12-port optical packet switching (OPS) network. It is found that the self-routing packets are shortened by 400 ps for each node hop due to switching transients. The system further demonstrates a 4% range of slot times over which nondestructive packet routing can occur. Relationships for timing margin requirements in self-routing OPS fabrics are developed as well.**

***Index Terms*—Interconnection networks, packet switching, photonic switching systems, timing.**

I. Introduction

In optical packet switching (OPS) networks, lightwave information can traverse an entire system limited only by time-of-flight latencies. Precision in timing within such networks is, therefore, extremely important for maximizing efficiency and hence throughput. Regardless of the architecture, because optical packets cannot be easily slowed down or stopped, OPS networks rely on the lengths of fiberoptic pathways to maintain system synchronization and appropriate packet flow timing [1]–[6].

The successful implementation of a complete 12×12 OPS network based on the data vortex architecture has been recently reported [7]. This complete network is presented here as a case study of the timing flexibility and scalability of photonic switching systems. The data vortex architecture shares the common OPS network requirement of precision in latency for correct timing operation and self-routing. However, this architecture also has an additional timing requirement in order to properly implement its internal deflection scheme [8], [9].

The basic timing structure of the data vortex is similar to other OPS architectures which transmit in a time-slotted synchronous manner and utilize distributed switching nodes for self-routing. This kind of OPS interconnection switching fabric is best suited for implementation as closed, physically proximate systems to ensure that the precise timing relationships can easily be realized, and may not be appropriate for long-haul OPS communications networks.

We study the flexibility of slot timing requirements in the data vortex experimentally and analyze them theoretically, and we discuss the effects on network scalability. Some general scaling relationships for latency-sensitive time-slotted OPS interconnection switching fabrics are also presented.

Manuscript received June 29, 2005; revised July 14, 2005. This work was supported in part by the National Science Foundation under Grant ECS-0322813 and in part by the U.S. Department of Defense under Subcontract B-12-644.

The authors are with the Department of Electrical Engineering at Columbia University, New York, NY 10027 USA (e-mail: bas@ee.columbia.edu; bergman@ee.columbia.edu).

Digital Object Identifier 10.1109/LPT.2005.857595

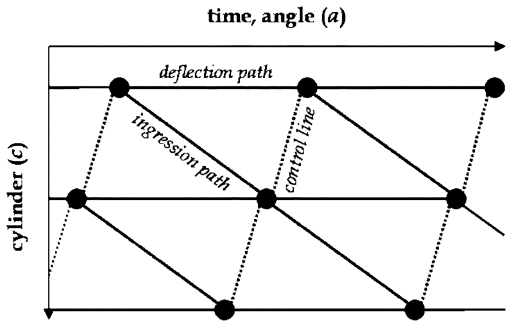

Fig. 1. Graphical representation of the data vortex deflection signal timing requirement, in which packets propagate from node (•) to node on fiber-optic paths (—). Ingression path latencies must be shorter than deflection path latencies in order to allow time for deflection signals to be sent on control lines (- - -).

II. Data Vortex Timing

The data vortex network topology relies heavily on passive unclocked timing constraints between the 2×2 switching nodes for the distributed self-routing of packets [7], [9]. First, packet slot times are preserved by the design of the routing path latencies, as in many OPS systems, since convenient dynamic buffering is not available [1]–[6]. Second, although the individual nodes do not require a clock signal, the electronic deflection signals which are sent between nodes must be timed correctly. For a packet to be deflected properly, the switching node must receive the deflection signal at the appropriate time, so that the routing decision can be executed while the packet is still within that node. This requirement can be illustrated graphically (Fig. 1).

Deflection signals are transmitted from nodes on downstream stages (called *cylinders*) to nodes on the adjacent upstream cylinder in order to prevent packet collisions [8], [9]. The deflection signals are sent only from one node to an adjacent upstream node, resulting in a modular scheme which is very straightforward to implement. The interdependence of the distributed self-routing nodes in the data vortex topology makes timing requirements especially important, although timing in OPS networks in general is very important for system functionality.

III. Experimental Setup

In order to generate packets of the correct format for the implemented 12×12 data vortex architecture, five routing header wavelengths are required. The routing header wavelengths are modulated with the header information, which is constant throughout the duration of the packet [7]–[9], and coupled to the 16-wavelength packet payload. The whole packet stream is divided in two by a 50 : 50 coupler, injecting packets into

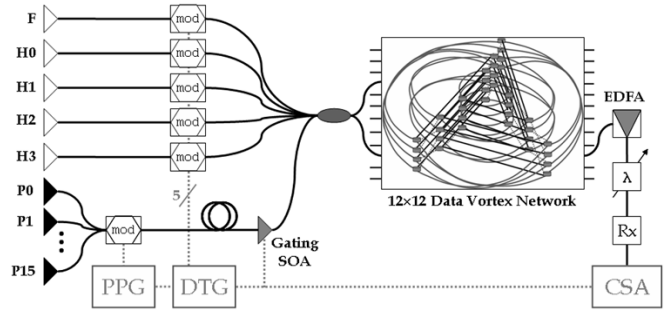

Fig. 2. Schematic of experimental setup with the appropriate packet generation and detection subsystems containing five routing header wavelengths (*F* and *H0* through *H3*) and 16 payload wavelengths (*P0* through *P15*), with $LiNbO_3$ modulators (mod), pulse pattern generator (PPG), data timing generator (DTG), communications signal analyzer (CSA), erbium-doped fiber amplifier (EDFA), tunable filter (λ), and high-speed dc-coupled postamplified receiver (Rx).

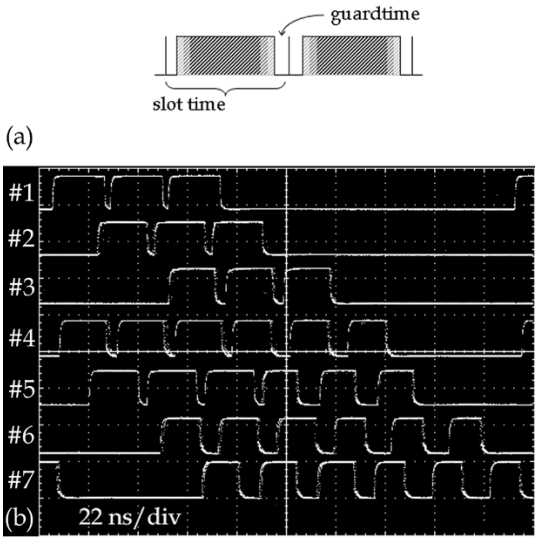

Fig. 3. (a) Diagram of the packet structure. (b) Electrical waveforms generated by the seven nodes as packets are incident on them. The first three packets enter the system at node #1 and are deflected at node #2; the result is seen at node #3. At node #4, the packets which caused the deflections (at node #2) enter the stream and propagate through the remaining nodes immediately following the first three packets. The packets are seen to move forward in time as they propagate through the switching nodes.

two different input ports of the 12×12 data vortex network (Fig. 2). The packet streams are logically identical, so they are accordingly routed to the same network destinations. But because they are injected into different input ports, their routing paths differ. Therefore, deflection signals are necessary to prevent the packet streams from colliding. This deflection test case is a typical example of the deflection routing situations which are common in the data vortex architecture.

The two packet streams contain three packets each. In the absence of the other packet stream, both would require four node hops to reach the addressed output port. However, one stream is made to deflect the other, resulting in a longer, seven-hop path through the network (three-hop deflection penalty). This example illustrates a common situation arising from the aforementioned data vortex deflection structure. The electronic waveforms generated by each packet incident on particular switching nodes allows for the observation of the progression of these packet streams (Fig. 3).

The duration of the packets and their slot times are also varied. The slot time is defined as the duration between the beginning of one packet and the beginning of the next, including guardtime for spacing [Fig. 3(a)]. For the data vortex architecture, this slot time matches the latency between adjacent switching nodes, so that packets subtend exactly one node. The duration is easily adjusted by tuning the packet generation clock. Thus, the optimal slot time for meeting the data vortex timing constraints, as discussed above, can be determined.

IV. Results

A. Packet Truncation

The first observation from this experiment is that the self-routed packets grow slightly shorter as they propagate through the network. This is an expected result of the finite rise and fall times of the semiconductor optical amplifier (SOA) switching elements, which have been measured to be approximately 0.9 ns each. However, at each node hop, the packets are actually truncated by a total of 0.4 ns on average. This figure is not the full 1.8 ns (rise plus fall times) because the transition edges are gradual slopes, and the low-speed detectors at each switching node have very high sensitivity, so that the routing decision is triggered on the slightest rise of incident optical power for an incoming packet; the reverse case also applies to the trailing edge. With faster SOA switching elements, the packet truncation can substantially be reduced further.

The packets are constructed with guardtimes at the leading and trailing edges to accommodate this switching truncation. For this system, these guardtimes are 1.6 ns each (3.2 ns total) [7], [9]. This timing margin is sufficient for eight node hops (3.2 ns divided by 0.4 ns). In this 12×12 data vortex network under moderate load conditions, only 0.3% of the packets require more than eight node hops.

In general, as the data vortex scales to larger numbers of input and output ports $N \times N$, the number of routing hops H required for each packet scales logarithmically [10]

$$H \sim \log_2 N + 1. \tag{1}$$

Therefore, the packet guardtime T_G required for a large system based on the data vortex architecture is on the order of

$$T_G \geq \tau \log_2 N \tag{2}$$

where τ is the average truncation at each switching node (0.4 ns in this case). This result is very encouraging, especially as improved switching element speeds yield direct reductions in τ.

B. Packet Slot Time

Next, the optimal slot time is determined by monitoring the length of the packets which emerge from the implemented network. With the particular packet streams used [Fig. 3(b)], it is easy to ascertain when packets are truncated in error due to timing mismatches. That is, when packets are too long or too short, the aforementioned deflection signal timing requirement breaks down and leads to improperly timed deflection signals. Other timing mismatches also accumulate, resulting in outgoing packets being shortened due to collisions at the packet head and tail. Thus, the optimal slot time can be determined by measuring

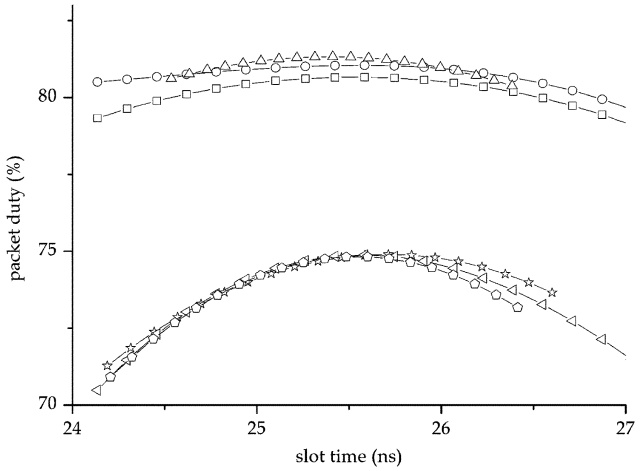

Fig. 4. Plot of the maximum packet duty versus slot time for the given six-packet sequence, with curves showing the manifestation of a generally parabolic relationship near the optimal slot time. The first three packets (#1 □, #2 ◯, #3 △) are longer than the last three (deflected) packets (#4 ◁, #5 ☆, #6 ⬠).

the packet duty, given as the ratio of the packet length to the slot time, which is an indicator of system efficiency.

In order to compare packet lengths consistently, the electronic signals generated by each packet as it enters a switching node are recorded (Fig. 4); these are the same signals that enable the SOA switching elements. The first three packets in the sequence traverse three node hops before reaching the final switching node (Fig. 3), truncated by 1.2 ns. The second three packets traverse an additional three node hops due to deflections, as discussed above, and experience a total truncation of 2.4 ns. This 1.2-ns difference can be seen in Fig. 4 as the 5% difference between the first three packets and the last three packets.

The results of Fig. 4 show that the optimal slot time, based on maximizing the packet duty, is 25.6 ns. Because of the guard-time inserted between packets, collisions are not severe as long as the slot time remains within approximately 4% of the optimal slot time (i.e., 24.6–26.5 ns). Beyond this range, significant packet collision and truncation occurs.

The curvature of these plots is significant, and illustrates that less variability is permitted for packets with longer paths through the system. For example, consider the range of slot times over which the packet duty falls 0.5% from its peak. For the first three packets, which traverse three node hops, this range is 1.6 ns; for the six-hop packets, the range shrinks to 1.1 ns (Fig. 4). For a 1% decline in packet duty, the ranges are 2.4 and 1.7 ns, respectively. It is evident that this relationship tends to be parabolic, which makes sense if we consider the manner by which variances are summed

$$\sigma = \sqrt{\sum_i \sigma_i^2} \tag{3}$$

for independent Gaussian distributions. The cause of the parabolic relationship for the timing margin is therefore likely attributed to fabrication error and variances in the fiber lengths.

In general, for a latency-sensitive distributed OPS network of fibers with length mismatches based on a uniform Gaussian distribution, the timing range $|\Delta T|$ of a particular path containing H node hops can be as much as

$$|\Delta T| \sim \sigma\sqrt{H} \sim \sigma\sqrt{\log_2 N} \tag{4}$$

for an $N \times N$ network by (1), where σ is the variance of a single node latency in time units. Thus, the margin allowed between packets must grow with network size to accommodate for this statistical mismatch. Or, conversely by (4), for a particular packet timing format which accounts for a fixed $|\Delta T|$, the certainty in node latency must improve with increasing network size.

V. Conclusion

The scalability and robustness of the implemented 12×12 data vortex architecture in terms of timing and latency requirements are analyzed. The system shows reasonable tolerance to packet timing variations, with nondestructive routing results within 4% of the optimal slot time. The induced packet truncation is measured and is shown to scale logarithmically with network size. We also present the scaling of the timing flexibility for a latency-sensitive self-routing distributed OPS switching fabric, correlated to mismatches in fiber lengths or in packet timing. These characterizations of timing requirements are important considerations in the design and implementations of large-scale OPS interconnection networks.

References

[1] R. S. Tucker and W. D. Zhong, "Photonic packet switching: An overview," *IEICE Trans. Electron.*, vol. E82-C, pp. 202–212, Feb. 1999.

[2] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *J. Lightw. Technol.*, vol. 21, no. 3, pp. 672–683, Mar. 2003.

[3] F. Masetti, D. Chiaroni, R. Dragnea, R. Robotham, and D. Zriny, "High-speed high-capacity packet-switching fabric: A key system for required flexibility and capacity," *J. Opt. Networking*, vol. 2, pp. 255–265, Jul. 2003.

[4] R. Rajaduray, S. Ovadia, and D. J. Blumenthal, "Analysis of an edge router for span-constrained optical burst switched (OBS) networks," *J. Lightw. Technol.*, vol. 22, no. 11, pp. 2693–2705, Nov. 2004.

[5] R. Hemenway, R. R. Grzybowski, C. Minkenberg, and R. Luijten, "Optical-packet-switched interconnect for supercomputer applications," *J. Opt. Networking*, vol. 3, pp. 900–913, Dec. 2004.

[6] T. Lin, K. A. Williams, R. V. Penty, I. H. White, M. Glick, and D. McAuley, "Self-configuring intelligent control for short reach 100 Gb/s optical packet routing," in *Proc. Optical Fiber Comm. Conf. (OFC)*, Anaheim, CA, Mar. 2005, Paper OWK5.

[7] B. A. Small, O. Liboiron-Ladouceur, A. Shacham, J. P. Mack, and K. Bergman, "Demonstration of a complete 12-port terabit capacity optical packet switching fabric," in *Proc. Optical Fiber Comm. Conf. (OFC)*, Anaheim, CA, Mar. 2005, Paper OWK1.

[8] Q. Yang and K. Bergman, "Traffic control and WDM routing in the data vortex packet switch," *IEEE Photon. Technol. Lett.*, vol. 14, no. 2, pp. 236–238, Feb. 2002.

[9] B. A. Small, A. Shacham, and K. Bergman, "Ultra-low latency optical packet switching node," *IEEE Photon. Technol. Lett.*, vol. 17, no. 7, pp. 1564–1566, Jul. 2005.

[10] Q. Yang and K. Bergman, "Performances of the data vortex switch architecture under nonuniform and bursty traffic," *J. Lightw. Technol.*, vol. 20, no. 8, pp. 1242–1247, Aug. 2002.