Demonstration of Asynchronous Operation of a Multiwavelength Optical Packet-Switched Fabric

Caroline P. Lai, *Student Member, IEEE*, Assaf Shacham, *Member, IEEE*, and Keren Bergman, *Fellow, IEEE*

***Abstract*—Asynchronous wavelength-striped message routing is experimentally demonstrated for a 4×4 optical packet-switched network test-bed. Asynchronous transmission provides increased interconnection network flexibility for future high-performance computing systems. Multiwavelength optical packets with 6×10-Gb/s payloads are shown correctly routed asynchronously through the network. Error-free operation with bit-error rates less than 10^{-12} is confirmed, with an average induced power penalty of 0.5 dB for the six payload wavelengths.**

***Index Terms*—Optical communication, packet switching, photonic switching systems.**

I. Introduction

As current high-performance computing systems (HPCS) scale, a key challenge lies in designing the interconnect between the multiprocessors and memory nodes. Current interconnection networks are being increasingly constrained by the limited bandwidth, latency, and power performance of electronics. Optical interconnection networks (OINs) have been suggested as a promising solution to these performance bottlenecks [1], [2]. OINs constitute a scalable approach for the construction of a low-latency communications interconnection infrastructure for high-bandwidth multiwavelength optical messages [3]. Deployed OINs in future HPCS may be required to operate asynchronously without the need for a synchronization stage or timeslot alignment module, allowing for more flexible scheduling [4].

Asynchronous operation provides several advantages as compared to synchronous transmission, with no required temporal alignment between optical messages entering the network from different ports. This architectural approach yields two key benefits: first, since the HPCS terminals may be located physically distant in a large room, synchronizing all the transmission ports with a common clock may be costly. Asynchronous operation of the interconnect alleviates the need for fine temporal alignment, calibration, and synchronization. Second, in asynchronous mode operation, variable-length packets are supported, providing greater flexibility and enabling the exchange of small optical packets. This is acutely important for HPCS applications since many messages are control packets with little data. In synchronous networks, the exchange of these messages using full timeslots can lead to underutilization [5]. Thus, asynchronous transmission can lead to more flexible operation.

The OIN architecture here is a multistage optical packet-switched network design comprised of photonic switching nodes. It provides a high-bandwidth, low-latency interconnect for a scalable number of ports by creating transparent lightpaths among terminals [3]. Contentions in the network are resolved by fast message dropping within the packet duration. A novel low-latency message acknowledgement (*ack*) scheme minimizes the latency for message retransmission.

The architecture was originally developed as a synchronous slotted network, transmitting fixed-duration optical messages. The architecture can be straightforwardly adapted to provide asynchronous operation by incorporating simple, minimal modifications to the nodes' electronic circuitry. The network yields simple asynchronous packet switching without requiring control plane signaling, optical buffering, or wavelength conversion. No additional hardware or components are needed to implement the asynchronous transmission as compared to the synchronous case. Unlike designs that require a centralized arbiter to manage routing, the network's unique distributed control nature is leveraged. The issue of signaling between the network nodes and a centrally controlled arbiter is alleviated.

In this letter, we present the first experimental demonstration and performance evaluation of asynchronous transmission of arbitrary-length, wavelength-striped optical messages across a 4×4 optical network test-bed. This allows for the asynchronous routing of multiwavelength optical packets through an optical network without additional hardware. Wavelength-striped optical packets with 6×10-Gb/s payloads are correctly routed through the test-bed, and error-free transmission with bit-error rates (BERs) less than 10^{-12} is confirmed for all payload wavelengths [6]. Sensitivity curves show an average induced power penalty of 0.5 dB for the six payload wavelengths.

II. Architecture Overview

The OIN design is comprised of 2×2 nonblocking photonic switching nodes arranged as an optical multistage Banyan network [3], which can map a large number of ports using only $\log_2 N$ stages of $N/2$ nodes. Wideband semiconductor optical amplifier (SOA) gates are used as switching elements, offering data-rate and packet-format transparency. The optical packets have lengths in the tens of nanoseconds, spanning across several meters or more. Since these messages may be longer than the switching elements, no storage or buffering capability is available within the nodes. Considering the application's relatively short reach, dispersion is not a significant factor, so multiple wavelengths are used via wavelength-division

Manuscript received February 28, 2010; revised April 30, 2010; accepted May 24, 2010. Date of publication June 10, 2010; date of current version July 23, 2010. This work was supported in part by the NSF Future Internet Design Program under Grant CNS-837995.

C. P. Lai and K. Bergman are with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: caroline@ee.columbia.edu).

A. Shacham is with Qualcomm Inc., New York, NY 10027 USA.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LPT.2010.2052097

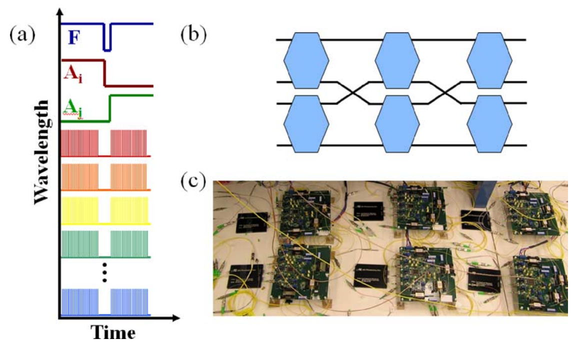

Fig. 1. (a) Wavelength-striped optical packet format. (b) Network block diagram. (c) Photograph of implemented network test-bed.

multiplexing (WDM) to provide very high-transmission bandwidths within the message payload. The architecture supports a wavelength-striped optical packet structure [Fig. 1(a)] in which control information such as frame and address bits are encoded on dedicated wavelengths (a single bit per wavelength), while the payload is segmented and modulated at a high data rate (here, 10 Gb/s per wavelength) on the rest of the available wavelength band [3]. The nodes decode control information immediately upon the reception of the packets' leading edges using fixed wavelength filters and low-speed 155-Mb/s p-i-n photodetectors.

The switching gates in the node are four SOAs, organized in a gate-matrix structure. The framing and address signals are recovered from the incoming optical packet, processed by high-speed electronic circuitry, a complex programmable logic device (CPLD), to gate the appropriate SOA gates. The optical messages are then routed to their desired destination (or dropped upon contention).

Successfully routed messages create transparent lightpaths that extend across the network. When the messages' leading edges reach the output (while the messages are still being transmitted), optical *ack* pulses are sent in the reverse direction, leveraging the bidirectionality of the switching nodes. Thus, an acknowledgement protocol is established, notifying the sources that their messages have been successfully received. No *ack* pulse is received by sources of dropped messages, which can then retransmit at a later time. Due to the simple instantaneous signaling of this protocol, the *ack* pulses are received at the sources with minimal delay, yielding a low latency penalty associated with retransmission. Asynchronous operation allows the input ports to inject varying-length packets in an unslotted manner. At a switching node, priority is given to the first packet to be routed through the node. The node is set appropriately for the duration of the prioritized packet transmission, and other contending messages that wish to propagate through the node are dropped.

III. Experimental Demonstration

A. Experimental Setup

The 4×4 three-stage optical network test-bed is implemented with commercially-available individually-packaged components: SOAs, passive optical components, and high-speed digital electronics. This test-bed realization has

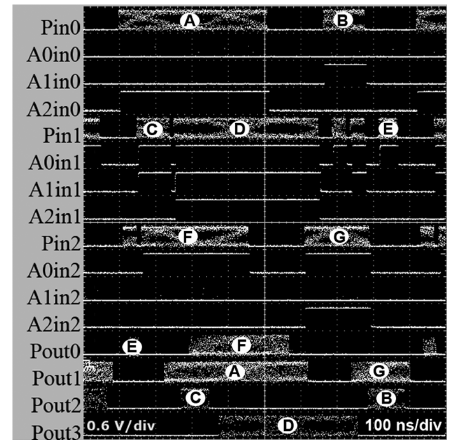

Fig. 2. Experimental optical waveforms of asynchronous traffic.

been used to demonstrate essential network concepts including address encoding/decoding, correct routing, and error-free transmission of wavelength-striped optical messages in conjunction with short *ack* pulses [3]. Here, the test-bed is adapted to asynchronously route variable-length messages by incorporating a 2-bit register to encode the node state in each node's electronic routing logic. Under the assumption of synchronous operation, all messages are received simultaneously at the beginning of the timeslot; thus, stateless combinatorial logic is sufficient to process the header and route optical messages. However, when messages arrive at differing times asynchronously, a 2-bit memory is used to denote the nodes' state to give priority to lightpaths that have been set and avoid interference with new messages. Each switching node is set on a per-packet basis according to the headers of the ingressing messages and the state of a given switching node is independent of other nodes.

B. Results

In order to verify the network's capability to route asynchronous traffic with variable packet lengths, a pattern of wavelength-striped packets is injected into the network via three independent input ports. Fig. 2 depicts the optical waveforms corresponding to the input and output signals.

The optical packets (labeled A–G in Fig. 2), with lengths varying between 53.3 and 409.6 ns, are injected from three input ports (*in0*, *in1*, and *in2*). There is no assumed relationship between the individual packets' start and end times. Each packet occupies the full duty length, incorporating a four-wavelength control header and six payload wavelengths. Distributed-feedback lasers (DFBs) are used to generate the incoming messages. All payload wavelengths are simultaneously modulated at 10 Gb/s with a single $LiNbO_3$ modulator with a $2^7 - 1$ pseudorandom binary sequence (PRBS) in a nonreturn-to-zero (NRZ) ON–OFF-keying (OOK) format. The payload wavelength channels range from 1539.6 to 1558.28 nm. On the dedicated control wavelengths, the messages have optically encoded addresses denoting the designated output (*out0*, *out1*, *out2*, and *out3*), modulated at a single bit per wavelength. The optical header has one frame signal (not shown in Fig. 2), a distribution address (*A0*, selecting one of two paths in the three-stage test-bed), and a two-bit routing address (*A1*, *A2*).

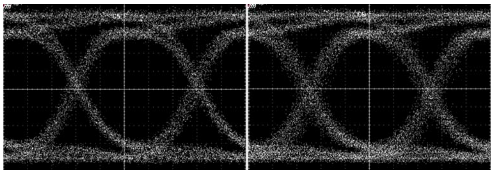

Fig. 3. Eye diagrams at 10 Gb/s of the input (left) and output (right) signals ($\lambda = 1560.2$ nm).

For example, packet C (89.2 ns long) is injected from *in1*, addressed to *out2* ($A1 = 1$, $A2 = 0$), and emerges at *out2*.

Following injection, a packet may be blocked by other packets whose transmission began earlier. The new packet is dropped and its source does not receive an *ack*. In the case of no *ack*, the source recognizes that the packet was blocked and retransmits via a different path. As an example, the first injection attempt from *in2* is blocked; the source retransmits with a different distribution address (*A0* changes to 1 from 0) and another path to the destination is found (packet *F*).

At the output, an optical filter selects one payload wavelength channel which is sent to a direct-coupled 10-Gb/s p-i-n-TIA receiver with limiting amplifier. A bit-error-rate tester (BERT) is then used that is synchronized with the packet gating signals. Error-free routing is verified for all six payload wavelengths of the egressing asynchronously-routed packet, achieving BERs less than 10^{-12}. Fig. 3 shows the 10-Gb/s input and output eye diagrams, with no extinction ratio difference. The power penalty induced by the test-bed is evaluated for all six payload channels. Fig. 4(a) shows a representative set of sensitivity curves for one of the evaluated wavelengths, validating the network's error-free performance. Fig. 4(b) gives the power penalty (at a BER of 10^{-9}) as it relates to wavelength: the average power penalty is 0.5 dB for the three-stage network, ranging from 0.3 to 0.8 dB. The network is not limited by the SOA amplified spontaneous emission (ASE). It is significant to note that the obtained power penalty values are less than 1 dB for all payload channels. The power penalty values verify that this approach to asynchronous routing does not adversely affect the network's operation, as compared to synchronous transmission.

IV. Conclusion

Using minimal electronic logic circuitry modifications, the asynchronous mode operation of an optical interconnection network test-bed is successfully demonstrated with variable-length wavelength-striped optical packets. Correctly routed packets deliver six wavelength channels of 10-Gb/s payloads error-free, with BERs less than 10^{-12}. The achieved power penalties for asynchronous transmission are similar to the original slotted operation. This experimental demonstration paves the way for the enhanced, scalable performance of future HPCS interconnection network infrastructures.

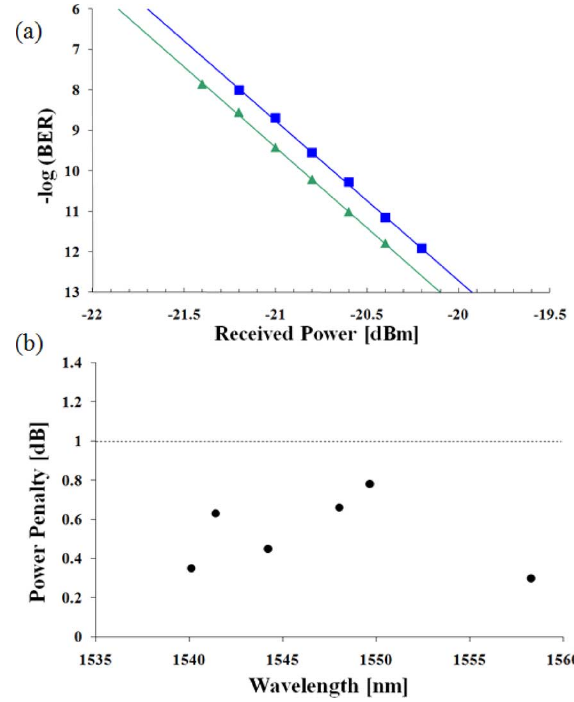

Fig. 4. (a) 10-Gb/s sensitivity curves for one representative wavelength ($\lambda = 1558.28$ nm) [▲ back-to-back; ■ through]. (b) Power penalty with respect to all six payload wavelength channels.

References

[1] A. F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta, and M. B. Ritter, "Exploitation of optical interconnects in future server architectures," *IBM J. Res. Dev.*, vol. 49, no. 4/5, pp. 755–775, Jul. 2005.

[2] R. Luijten, W. E. Denzel, R. R. Grzybowski, and R. Hemenway, "Optical interconnection networks: The OSMOSIS project [invited]," in *Proc. 17th Annu. IEEE Lasers & Electro-Optics Soc.(LEOS 2004)*, Puerto Rico, Nov. 2004, Paper WM1.

[3] A. Shacham and K. Bergman, "An experimental validation of a wavelength-striped, packet switched, optical interconnection network," *J. Lightw. Technol.*, vol. 27, no. 7, pp. 841–850, Apr. 1, 2009.

[4] L. Tancevski, S. Yegnanarayanan, G. Castanon, L. Tamil, F. Masetti, and T. McDermott, "Optical routing of asynchronous, variable length packets," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, pp. 2084–2093, Oct. 2000.

[5] D. Dai and D. K. Panda, "How can we design better networks for DSM systems?," in *Lecture Notes in Computer Science*, Jun. 1997, vol. 1417, pp. 171–184.

[6] A. Shacham, C. P. Lai, and K. Bergman, "Experimental demonstration of an optical interconnection network with asynchronous transmission," in *Eur. Conf. Optical Commun. (ECOC 2007)*, Berlin, Germany, Sep. 2007, Paper 6.5.6.