Electrical Ingress Buffering and Traffic Aggregation for Optical Packet Switching and Their Effect on TCP-Level Performance in Optical Mesh Networks

Shun Yao, Fei Xue, Biswanath Mukherjee, and S. J. Ben Yoo, University of California, Davis

Sudhir Dixit, Nokia Research Center

ABSTRACT

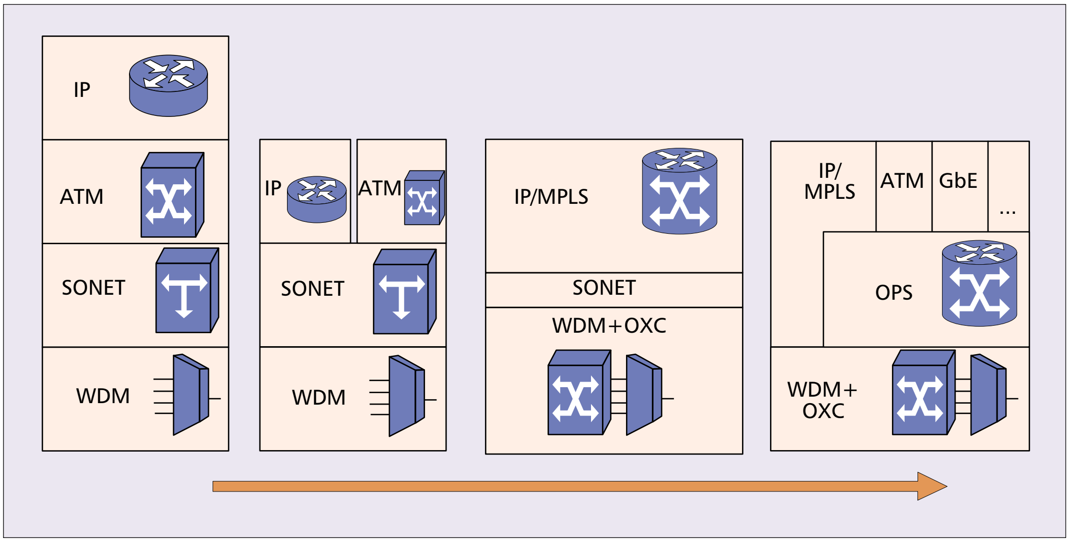
The wide deployment of wavelength-division multiplexing technology and new transmission techniques have resulted in significant increases in the transmission capacity in optical fibers, both in the number of wavelengths and the bandwidth of each wavelength channel. Meanwhile, the fast growth of the Internet demands more data switching capacity in the network in order to deliver high bandwidth to end users. Although the capacity of electronic routers has been increasing consistently in the past, optical switching appears to be a more cost-effective way to switch individual wavelengths. As the bit rate per wavelength channel continues to grow, optical subwavelength switching emerges as a new paradigm capable of dynamically delivering the vast bandwidth WDM offers. This article discusses one of such techniques, namely optical packet switching, and its performance perceived by end users in optical mesh networks. Specifically, our investigation reveals the benefit of using electrical ingress buffering and traffic aggregation to reduce packet-loss rate of optical packet-switched networks. Through simulation experiments, we present an evaluation of the network's TCP-level performance based on the proposed architecture.

INTRODUCTION

Over the past decade, wavelength-division multiplexing (WDM) technology has brought fundamental changes to network design [1, 2]. WDM has evolved from a point-to-point transmission technology to a separate networking layer with its own switching entities and control plane. Wavelength-routed networks, in which lightpaths are set up on specific wavelengths, have been the focus of extensive studies [3, 4]. Within a short period of a few years, these networks have evolved from textbook subjects to real-life products. In order to facilitate easy and flexible access to the high bandwidth such networks offer, both industry and academia are constantly in search of ways to automate and expedite wavelength and bandwidth provisioning in the optical layer. These ongoing efforts indicate the inevitable trends that lead to more intelligent and switch-capable optical networks. Migration of certain switching functionality from electronics to optics will reduce the amount of optical-electrical-optical (OEO) conversions, which become more expensive as the channel bit rate continues to increase.

Until recently, Internet Protocol (IP) routers were interconnected by virtual circuits provided by asynchronous transfer mode (ATM) cell switches. The ATM network is built with links provided by synchronous optical networks (SONET) in the form of time-division multiplexed (TDM) circuits. As the networks evolved, ATM was gradually replaced partially by IP and partially by SONET, while IP routers became capable of directly interfacing with SONET equipment. The multiprotocol label switching (MPLS) protocols further enrich the functionalities of IP. The increase in an IP router's port data rate and aggregate capacity enables the backbone high-performance IP routers to directly use wavelengths as links between routers. However, a wavelength-routed network can only provide a whole wavelength as the smallest

0163-6804/02/$17.00 © 2002 IEEE

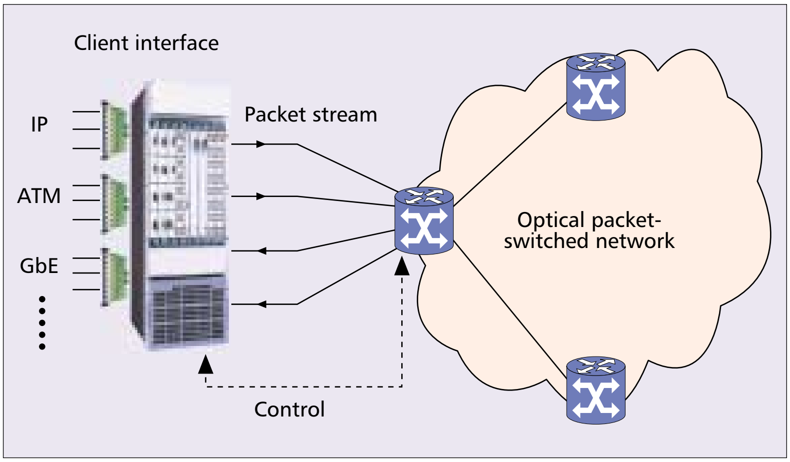


■ **Figure 1.** *The evolution of the protocol stack in telecommunications networks.*

bandwidth unit. A subwavelength circuit (which is more suitable for nonbackbone routers) can only be obtained through traffic aggregation/deaggregation equipment, which is electrical and often expensive at high bit rates. Although SONET performs traffic aggregation well, its TDM-based synchronous aggregation scheme might not be the most cost-effective solution in the new network environment. Hence, various optical subwavelength switching paradigms have been proposed to facilitate easy access to the vast bandwidth WDM offers [5–8]. Among all the optical subwavelength switching techniques, optical packet switching (OPS) [9] is a strong candidate as the building block for the next-generation Internet. It has a fine switching granularity (at the packet level), and provides seamless integration between the WDM layer and IP layer (Fig. 1).

One of the objectives in designing an OPS network is low packet loss rate (PLR, sometimes also denoted as probability of packet loss, PPL). Packet loss is caused by packets dropped in contentions, when there are two or more packets contending for the same output fiber on the same wavelength, at the same time. In electrical packet networks, contention is usually resolved with the store-and-forward technique, which requires the packets that lose a contention to be stored in a memory bank, and sent out at a later time when the desired output port clears. In an OPS network, due to the lack of optical memory, contentions have to be resolved with different approaches. Past work in the field of OPS offers a number of studies on various node architectures and contention resolution algorithms [10]. Most of these studies focus on the optical domain architecture of a single OPS node or a network interconnected with such nodes.

An OPS network, by its name, should perform packet switching while the payload data stays in the optical domain. Nevertheless, it needs to interface with other types of networks to provide end-to-end connectivity (Fig. 2). These client networks are often electrical. Therefore, there needs to be an electrical-to-optical interface at the edge of the OPS network. Such an interface can provide inexpensive electrical buffering. This work proposes to exploit the availability of electrical buffers at the network ingress to resolve contention and lower the network cost.



■ **Figure 2.** *The client interface of OPS networks.*

A well-designed node architecture and contention resolution algorithm can provide low PLR in an OPS network. However, a low PLR is not the only metric for measuring the OPS network's performance. A network's true performance should be measured from the end user's point of view, together with the consideration of network load. In the case of an OPS network, since most of today's data traffic consists of IP-based Transmission Control Protocol (TCP) traffic, TCP performance appears to be a realistic metric. This investigation presents a TCP performance study of OPS networks with electrical ingress buffering and traffic aggregation (defined in the next section). This article is organized as follows. We first describe an architecture that utilizes electrical ingress buffering and traffic aggregation. Then an investigation is presented on TCP performance with the proposed architecture. Finally, we conclude the article.

Both wavelength conversion and optical buffering require extra hardware and control software. Deflection routing can be implemented with extra control software only.

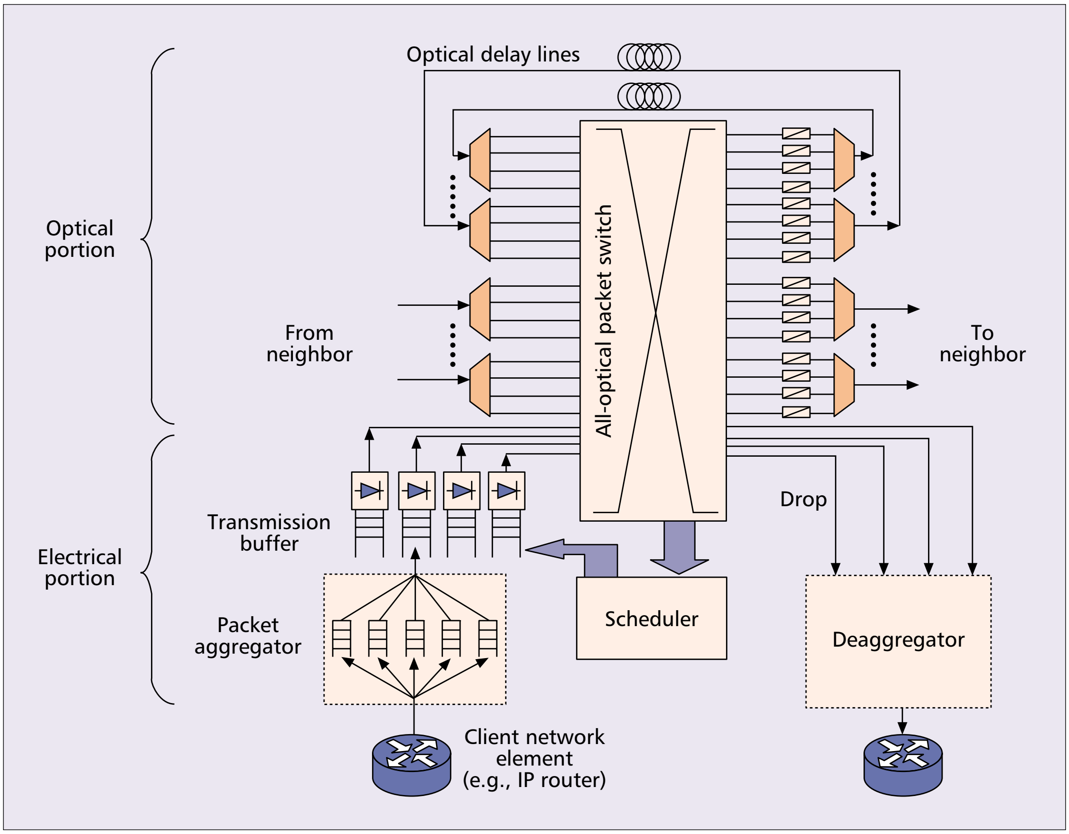


Figure 3. *The architecture of the proposed optical packet switch with an electrical ingress buffer and traffic aggregation.*

NETWORK ARCHITECTURE

In an OPS network, contention occurs at a switching node whenever two or more packets try to leave the switch fabric on the same output fiber, on the same wavelength, at the same time. Since there is no optical equivalent of the electronic random access memory (RAM) technology, the optical packet switches need to adopt approaches other than the store-and-forward technique for contention resolution. Meanwhile, WDM technology offers an additional different dimension, wavelength, for contention resolution. The optical contention resolution mechanisms that can employ three dimensions are outlined below (see the optical portion of the switch architecture shown in Fig. 3).

Wavelength conversion offers effective contention resolution without relying on buffer memory. Wavelength converters can convert wavelengths of packets contending for the same wavelength of the same output port. It is a powerful and the most preferred contention resolution scheme that does not cause extra packet latency, jitter, and packet reordering problems.

Optical delay line (which provides sequential buffering) is a close imitation of the RAM in electrical routers, although it offers a fixed and finite amount of delay. Many previously proposed architectures employ optical delay lines to resolve contentions. Since optical delay lines rely on the propagation delay of the optical signal in silica to buffer the packet in time (i.e., due to their sequential access), they have more limitations than electrical RAM. To implement a large buffer capacity, a switch may need to include a large number of delay lines.

The space deflection approach is a multipath routing technique. Packets that lose the contention are routed to nodes (usually along the second shortest path) other than their preferred next-hop nodes, with the expectation that they will eventually be routed to their destinations. The effectiveness of deflection routing depends heavily on the network topology and offered traffic pattern.

Both wavelength conversion and optical buffering require extra hardware (wavelength converters and pump lasers for wavelength conversion; fibers and additional switch ports for optical buffering) and control software. Deflection routing can be implemented with extra control software only.

When packets arrive from electrical client networks, they need to be converted to optical signals before being sent to the OPS network. This conversion is performed at the client interface of the network (Fig. 2). The optical packet switch performs two types of packet forwarding: the forwarding of transit packets from other optical packet switches, and the forwarding of local packets received from the client interface. A transit packet has to cope with possible contention from the local packets as well as other transit packets. In most proposed architectures, the contention resolution usually requires a large amount of optical resources, such as wavelength converters and delay lines. In the proposed contention resolution scheme, the local packets are first queued in the electrical ingress buffers, which can easily be implemented in the electrical part of the client interface. These packets enter the optical switch only when there is no transit packet occupying the preferred wavelength/output port. Since the switching is still carried out by the optical components and there is no OEO conversion in the network, the use of electrical buffers at ingress does not compromise the all-optical nature of the network.

Figure 3 shows the node architecture that implements electrical ingress buffering and traffic aggregation (as explained below). The packet aggregator assembles client packets into larger entities, referred to as aggregation packets, in a first-in first-out (FIFO) manner. It directly interfaces with the client network elements (typically IP routers), and consists of a number of FIFO subqueues. Each subqueue buffers client packets going to the same destination. The departure of an aggregation packet is triggered by a threshold, measured in either number of packets or number of bits. To avoid excessive queuing delay, the subqueues also adopt a timeout period, after which an aggregation packet departs even if the threshold is not reached. In the case of timeout-triggered departure, the aggregation packet may have a size smaller than the threshold size (either in number of packets or number of bits). This aggregation mechanism can be compared to a bus system (as in public transportation): At any time, there is one bus with one or more empty seats waiting for passengers for each destination. A bus has a maximum capacity for passengers and leaves periodically. If the bus is full before its scheduled departure time, it will leave early and the next empty bus will pull into the station. This aggregator not only preserves the order of client packets, but also shapes the traffic by injecting more evenly sized optical packets at more regular time intervals.

Once the aggregation packet leaves the subqueue, it enters the ingress transmission buffer (also a FIFO queue), which is designated to a specific local add port. A scheduler constantly monitors the state of the wavelengths at the output ports of the switch. Whenever the FIFO queue is not empty and there is a vacant wavelength on the preferred output port of the aggregation packet at the head of the queue, the scheduler instructs the transmitter to send the packet to the switch fabric. This transmission buffering mechanism ensures that all the wavelength converters and delay lines are used solely for resolving contentions among the transit packets. In the optical portion of the switch, the contention of transit packets is resolved by wavelength conversion, time buffer, and space deflection, as described previously for the all-optical approach.

In order to illustrate the performance improvement of using electrical ingress buffering and traffic aggregation, we present the results from some simulation experiments. Our first set of experiments reveals the advantage of electrical ingress buffering without any traffic aggregation. It is based on a 15-node mesh network (Fig. 4) with different numbers of wavelengths. Each node is capable of performing wavelength conversion, optical buffering, and deflection routing. The number of optical delay lines per node is equal to the nodal degree. The length of the delay line is chosen such that it can buffer the largest packet, which is approximately 1 km at 2.5 Gb/s line speed. The client packets generated in our simulation experiments are IP packets (with the largest packet size of 12,000 bits) with bursty arrivals.

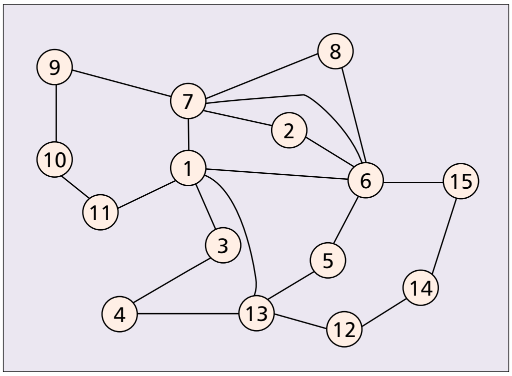
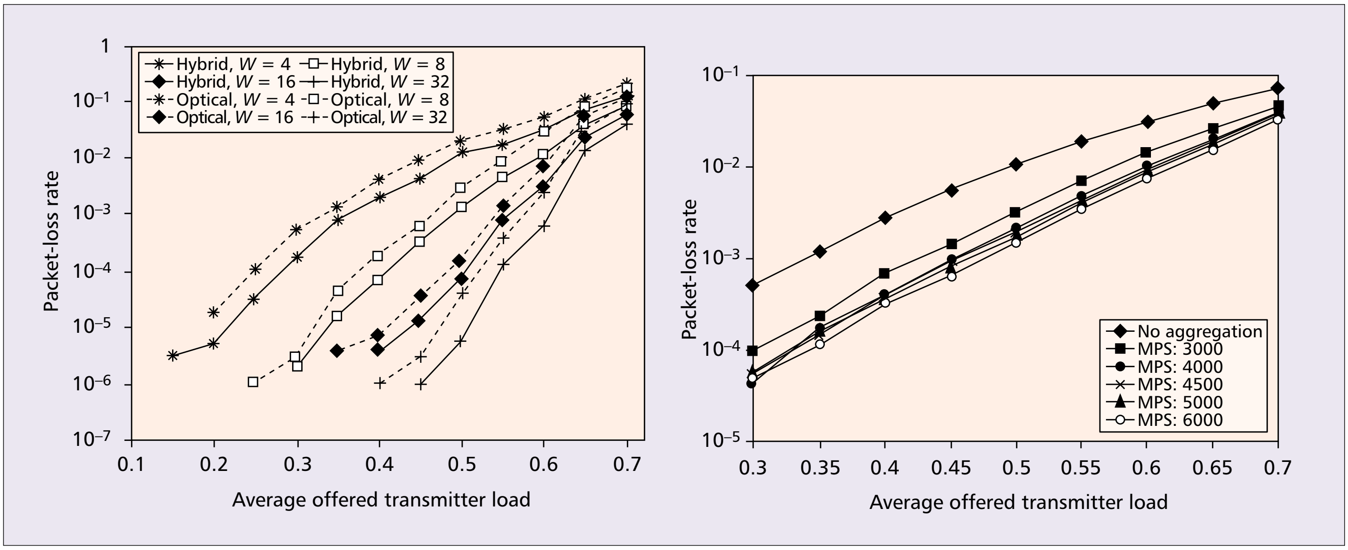


Figure 4. *An example mesh network topology used in the simulation experiments.*

Figure 5a shows the packet loss rates plotted against the average offered transmitter load, that is, the total number of bits offered per unit time divided by the line speed. (Once the network topology is given, the average link load is proportional to the average transmitter load.) For the four-wavelength scenario, the packet loss rate is kept below 0.01 when the offered transmitter load is less than 0.5. For the 32-wavelength case, the transmitter load can be as high as 0.65 while the packet loss rate is kept below 0.01. These results indicate the benefit of wavelength conversion when there are more wavelengths in the network. It is also in line with the fact that most network operators avoid loading their networks more than 50 percent. For comparison, Fig. 5a shows the packet loss rate of the same network with only all-optical contention resolutions, without any electrical ingress buffers. In general, the use of an ingress buffer helps reduce the packet loss by 50 percent from that of the all-optical scheme.

Figure 5b shows the packet loss rates (for original data packets, not the aggregation packets) with traffic aggregation for a six-node network with four wavelengths (Fig. 6). The transmission of aggregation packets is triggered by a maximum payload size (MPS) measured in bytes. The traffic aggregation results in noticeably smaller PLRs than in the baseline network without aggregation, and this improvement becomes larger as the MPS value increases. For example, with a transmitter load of 0.5, the PLR with MPS of 6000 bytes is 0.0015, while the PLR without aggregation is 0.1. Traffic aggregation usually leads to a severalfold benefit of reducing the PLR. When the transmitter load is 0.5, the obtained gains (ratio of PLR of no-aggregation case to aggregation cases) for MPS = 4500, 5000, and 6000 bytes are 5.56, 6.43, and 7.31, respectively. This benefit of aggregation is mainly due to its traffic-shaping property, as explained below. The Internet traffic is known to be bursty under a large range of timescales. With this type of traffic the interval between packet arrivals could be very large or very small. Meanwhile, the sizes of IP packets are awfully irregular. Statistical data [11] shows that almost 75 percent of the IP packets (in terms of number of packets)

If the bus is full before its scheduled departure time, it will leave early and the next empty bus will pull into the station. This aggregator not only preserves the order of client packets, but also shapes the traffic by injecting more even-sized optical packets at more regular time intervals.



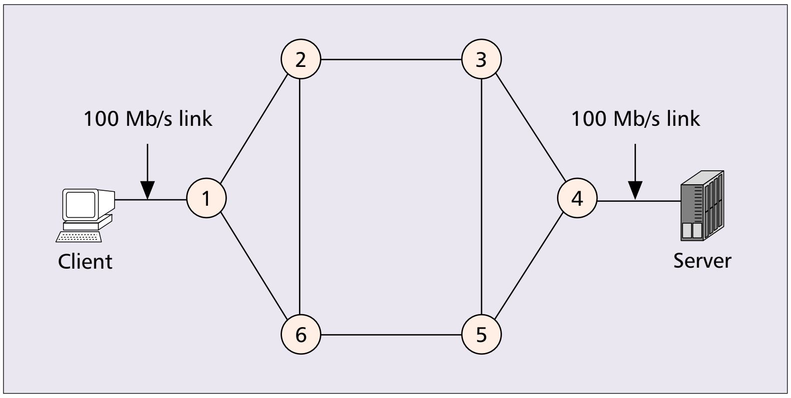
■ **Figure 5.** *Packet loss rate with a) electrical ingress buffering only; b) electrical ingress buffering and traffic aggregation.*

are smaller than 552 bytes. Nearly half of the IP packets are 40–44 bytes in length, due to TCP acknowledgment segments, TCP control segments, and telnet packets carrying single characters. On the other hand, over half of the total traffic (in terms of number of bits) is carried in packets of 1500 bytes (which is the typical maximum size of packets generated by Ethernet-attached hosts) or larger. Both bursty arrivals and irregular packet size distribution result in less efficient contention resolution inside the optical packet switch. The proposed traffic aggregation scheme helps relieve the burden on the optical contention resolution by shaping the traffic to more even-sized packets arriving at more regular intervals. It was observed that a larger aggregation size offers more benefit. For any specified value of MPS, the rate of gain tends to decrease with an increase in load because, at higher load the traffic becomes less bursty, and the effect of traffic shaping from aggregation is less prominent.

TCP PERFORMANCE

With IP being the main traffic in today's data networks, designers of OPS networks need to consider performance beyond the OPS network itself. The IP layer provides a best-effort connectionless packet delivery service. TCP was specifically designed to provide a reliable end-to-end connection over an unreliable internetwork (the IP network). An internetwork differs from a single network because different parts of the internetwork may have different topologies, bandwidths, delays, and other parameters. TCP was designed to dynamically adapt to properties of the internetwork and to be robust in the face of failures. It provides an acknowledgment-based reliable service to most Internet applications. TCP entities reside on end hosts as a part of the operating system. They accept user data streams from local processes, break them up into pieces usually not exceeding 1500 bytes, and throttle packet transmission according to the network bandwidth and round-trip delay. TCP-based traffic accounts for approximately 90 percent of the total Internet traffic [11]. Therefore, we believe it is important to investigate the effect on the TCP performance of a new OPS architecture.

Our initial attempt to measure the TCP performance is to carry out a set of simulation experiments based on the OPNET simulation software. Figure 6 shows an example network topology used in this study. Each link is bidirectional, 20 km in length, and carries four wavelengths, each operating at 2.5 Gb/s. The OPS nodes are as described in Fig. 3, with the aggregation threshold C measured in number of packets. A client and a server are connected to the network through 100 Mb/s links. For purposes of illustration, we choose a file transfer protocol (FTP) session to capture the TCP performance. The main performance metric is the transfer time of a large file, assumed to be 1.6 Mbytes long in this example. We also assume that both hosts have Ethernet interfaces; therefore, the maximum transfer unit (MTU) is 1500 bytes. To simulate a realistic network scenario, the network also carries bursty IP traffic in the background on all links. Each node is equipped with four transmitters fed independently by four traffic generators. The intensity of the background traffic is controlled by the average offered transmitter load.



■ **Figure 6.** *The network topology for the TCP experiment.*

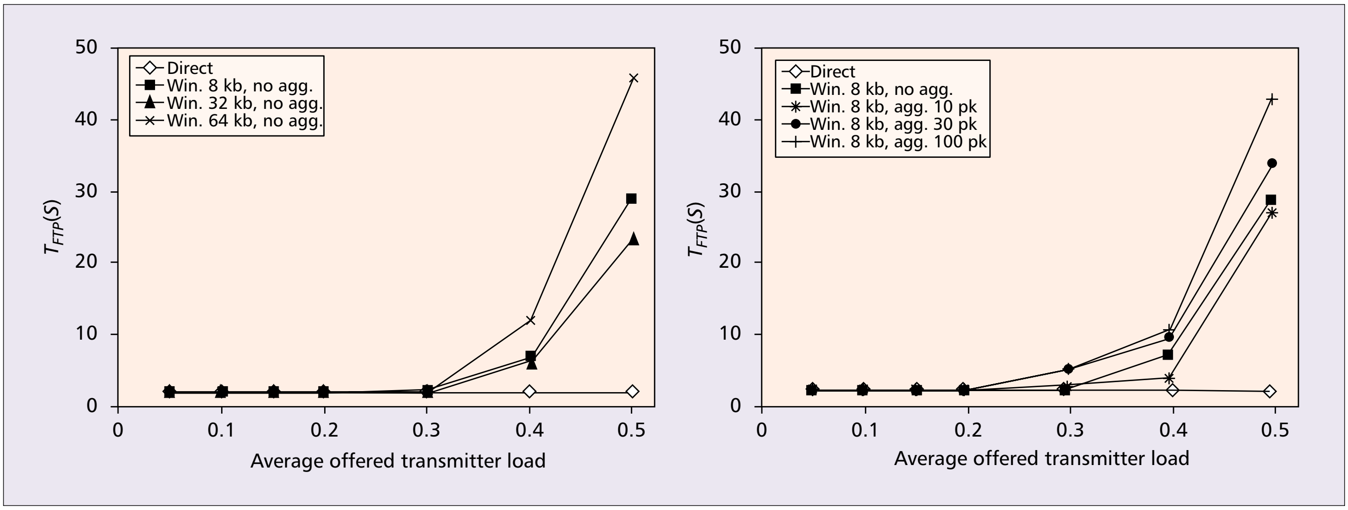


Figure 7. *A comparison of TFTP a) for different TCP window sizes; b) for different aggregation schemes.*

One of the main factors that affect TCP performance is the receiver window size, whose typical values are 8, 32, or 64 kbytes. The aggregation threshold C can also impact the TCP performance. When C is varied, the timeout value and delay line size should be adjusted accordingly. In our experiments, we set both the timeout period and delay line length to be equal to the transmission delay of C packets with maximum length (1500 bytes each). The running time for each data point varied between 4 and 75 hours on a 500-MHz Pentium III machine, depending on transmitter load. The maximum transmitter load was 0.5 because larger values made the simulation time prohibitively long.

Figure 7a compares the file transfer time T_{FTP} for different TCP window sizes with electrical ingress buffering only and no aggregation. For reference, it also shows the T_{FTP} without any background traffic for a client-server pair directly connected through a 100 Mb/s link with the same propagation delay (i.e., a link length of 60 km). Without aggregation, a window size of 32 kbytes provides the best result because the measured TCP round-trip time (RTT) is approximately 3 ms, and the TCP connection's data rate is 100 Mb/s. (Note that the optimal window size should be the product of RTT and data rate.) The figure shows that with average transmitter load exceeding 0.4, T_{FTP} increases considerably faster. This is because, for a given network, TCP performance would deteriorate significantly after the packet loss rate reaches a certain value [12].

Next, in Fig. 7b, we study the effect of the aggregation threshold by varying C to 10, 30, and 100 packets, with window size equal to 8 kbytes. For transmitter load less than 0.2, the aggregation threshold does not seem to have much effect on system performance. As the transmitter load increases, the 10-packet aggregation scheme has the lowest T_{FTP}, followed by the 30-packet and 100-packet schemes. The 10-packet scheme also performs better than the one without aggregation, indicating that aggregation improves TCP performance. However, with more packets aggregated, the performance deteriorates because more queuing delay is introduced in the packet aggregator. Intuitively, one would think a good aggregation scheme should collect all the TCP segments sent within one window size and send them out in one aggregation packet. If this was the case, the aggregator would have to hold the first segment for at least the whole transmission delay of all the segments in that window. Such a scheme would defeat the purpose of pipelining in the TCP sliding window mechanism, because the total transmission delay for the whole TCP window is determined by the slowest link (in this example the 100 Mb/s link) regardless of how fast the rest of the network is. Therefore, the benefit of traffic aggregation on TCP performance is not directly caused by aggregating TCP segments within one window, but rather by its traffic shaping effect (and the consequent reduction of packet loss in the network).

CONCLUSION

This article presents a novel optical-electrical hybrid contention resolution scheme for OPS networks. This scheme exploits the inexpensive electrical buffer available at the ingress of the network to buffer packets before they enter the optical domain. Without introducing any noticeable extra latency under normal load, electrical ingress buffering can significantly improve the efficiency of optical contention resolution resources by allowing them to be used solely for resolving contentions among transit packets. In the example network, the networkwide packet loss rate is reduced by approximately 50 percent with the hybrid approach. The article also investigated the benefits of performing traffic aggregation at the ingress. An aggregation scheme was shown to lower PLR severalfold. Such a scheme can smoothe optical packet size and reduce the burstiness of Internet traffic. Based on the proposed node architecture, the TCP performance was evaluated for an OPS network. It was observed that under normal network load, an OPS network can provide good-quality transport service to the TCP applications. Moreover, traffic aggregation can significantly improve the TCP performance because of the low PLR.

It was observed that under normal network load, an OPS network can provide good-quality transport service to the TCP applications. Moreover, traffic aggregation can significantly improve the TCP performance because of the low PLR.

REFERENCES

[1] B. Mukherjee, *Optical Communication Networks*, McGraw-Hill, 1997.
[2] R. Rawaswami and K. Sivarajan, *Optical Networks: A Practical Perspective*, Morgan Kaufmann, 1998.
[3] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath Communications: An Approach to High Bandwidth Optical WAN's," *IEEE Trans. Commun.*, vol. 40, no. 7, Jul. 1992, pp. 1171–82.
[4] B. Mukherjee *et al.*, "Some Principles for Designing a Wide-area WDM Optical Network," *IEEE/ACM Trans. Net.*, vol. 4, no. 5, Oct. 1996, pp. 124–29.
[5] B. Meagher *et al.*, "Design and Implementation of Ultra-low Latency Optical Label Switching for Packet-switched WDM Networks," *IEEE/OSA J. Lightwave Tech.*, vol. 18, no. 12, Dec. 2000, pp. 1978–87.
[6] C. Qiao and M. Yoo, "Optical Burst Switching (OBS) — A New Paradigm for an Optical Internet," *J. High Speed Net.*, vol. 8, no. 1, 1999, pp. 69–84.
[7] I. Chlamtac *et al.*, "Scalable WDM Access Network Architecture based on Photonic Slot Routing," *IEEE/ACM Trans. Net.*, vol. 7, no. 1, Feb. 1999, pp. 1–9.
[8] A. Detti, V. Eramo, and M. Listani, "Performance Evaluation of a New Technique for IP Support in a WDM Optical Network: Optical Composite Burst Switching (OCBS)," *IEEE/OSA J. Lightwave Tech.*, vol. 20, no. 2, Feb. 2002, pp. 154–65.
[9] S. Yao, B. Mukherjee, and S. Dixit, "Advances in Photonic Packet Switching: An Overview," *IEEE Commun. Mag.*, vol. 38, no. 2, Feb. 2000, pp. 84–94.
[10] *IEEE/OSA J. Lightwave Tech.*, vol. 16, Dec. 1998.
[11] http://www.caida.org/outreach/papers/Inet98/
[12] T. V. Lakshman and U. Madhow, "The Performance of TCP/IP for Networks with High Bandwidth-Delay Products and Random Loss," *IEEE/ACM Trans. Net.*, vol. 5, no. 3, June 1997, pp. 336–50.

BIOGRAPHIES

SHUN YAO (shyao@ece.ucdavis.edu) received his B.S. degree in electronic engineering from Ts-inghua University, China, in 1997 and his Ph.D. degree in electrical engineering from the University of California, Davis in 2001. From 1998 to 2000 he worked part-time at Nokia Research Centers in Helsinki. Finland. and Boston, Massachusetts. Currently he is a post-doctoral researcher at the University of California, Davis. His main research interests include optical switch architectures, network planning, traffic engineering, and fault management in WDM networks.

FEI XUE (fxue@ucdavis.edu) received his Ph.D. in computer engineering from Tianjin University, China, in 1998. He is currently a research scientist at the University of California, Davis. From 1998 to 1999 he worked as a research associate at the Chinese University of Hong Kong. From 1999 to 2000 he was a postdoctoral fellow at Simon Fraser University, Canada. His research interests include traffic modeling, performance evaluation of high-speed networks, optical IP, and architecture.

BISWANATH MUKHERJEE (mukerjee@cs.ucdavis.edu) received his B.Tech. (Hons.) degree from the Indian Institute of Technology, Kharagpur, in 1980, and his Ph.D. degree from the University of Washington, Seattle, in June 1987. At Washington he held a GTE Teaching Fellowship and a General Electric Foundation Fellowship. In July 1987 he joined the University of California, Davis, where he has been a professor of computer science since July 1995, and served as chair of computer science from September 1997 to June 2000. He is co-winner of paper awards presented at the 1991 and the 1994 National Computer Security Conferences. He serves or has served on the editorial boards of *IEEE/ACM Transactions on Networking*, *IEEE Network*, *ACM/Baltzer Wireless Information Networks*, *Journal of High-Speed Networks*, *Photonic Network Communications*, and *Optical Network Magazine*. He also served as editor-at-large for optical networking and communications for the IEEE Communications Society. He served as technical program chair of IEEE INFOCOM '96. He is author of the textbook *Optical Communication Networks* (McGraw-Hill, 1997), which received the Association of American Publishers, Inc.'s 1997 Honorable Mention in Computer Science. He is a member of the Board of Directors of IPLocks, Inc., a Silicon Valley startup company. He has consulted for and served on the Technical Advisory Boards of a number of startup companies in optical networking. His research interests include lightwave networks, network security, and wireless networks.

S. J. BEN YOO [SM] (yoo@ece.ucdavis.edu) joined UC Davis as associate professor of electrical and computer engineering in March 1999. He is currently serving as UC Davis branch director of a new intercampus center, the Center for Information Technology Research in the Interest of Society (CITRIS). His current research involves advanced switching techniques and optical communications systems for the next-generation Internet. Prior to joining UC Davis he was a senior scientist at Bellcore leading technical efforts in optical networking research and systems integration. His research activities at Bellcore included optical label switching for the next-generation Internet, power transients in reconfigurable optical networks, wavelength interchanging crossconnects, wavelength converters, vertical cavity lasers, and high-speed modulators. He also participated in ATD/MONET systems integration, the OC-192 SONET Ring studies, and a number of standardization activities. For his work, he received Bellcore CEO Award in 1998 and DARPA award for sustained excellence in 1997. Prior to joining Bellcore in 1991, he conducted research on nonlinear optical processes in quantum wells, four-wave mixing study of relaxation mechanisms in dye molecules, and ultra-fast diffusion driven photodetectors. During this period he also conducted research on lifetime measurements of intersubband transitions and nonlinear optical storage mechanisms at Bell Laboratories and IBM Research Laboratories, respectively. He received a B.S. degree with distinction in electrical engineering, an M.S. degree in electrical engineering, and a Ph.D. degree in electrical engineering with a minor in physics, all from Stanford University. His Ph.D. thesis at Stanford was on linear and nonlinear optical spectroscopy of quantum well intersubband transitions. He is a member of OSA and Tau Beta Pi.

SUDHIR D. DIXIT [SM] (sudhir.dixit@ieee.org) received a B.E. degree from Maulana Azad College of Technology (MACT), Bhopal, India, an M.E. degree from Birla Institute of Technology and Science (BITS), Pilani, India, and a Ph.D. degree from the University of Strathclyde, Glasgow, Scotland, all in electrical engineering. He also received an M.B.A. degree from Florida Institute of Technology, Melbourne. He is currently a senior R&D manager and site manager at Nokia Research Center, Burlington, Massachusetts. His main areas of interest are mobile/wireless Internet, optical networks, and content delivery networks. From 1991 to 1996 he was a broadband network architect at NYNEX Science and Technology (now Verizon Communications). Prior to that he held various engineering and management positions at other major companies such as GTE, Motorola, Wang, Harris, and STL (now Nortel Europe Labs). He has published or presented over 150 papers and has 23 patents either granted or pending. He has edited a book on IP over DWDM published by John Wiley in January 2002, and co-edited a book with Ramjee Prasad on wireless IP published by Artech House in December 2001. He was technical co-chair and general chair of the IEEE International Conference on Computer, Communications, and Networks in 1999 and 2000, respectively, technical co-chair of the SPIE Conference on Terabit Optical Networking in 1999, 2000, and 2001, general chair of the Broadband Networking in the Next Millennium Conference in 2001, and general co-chair of OptiComm 2002. He has also been an ATM Forum Ambassador since 1996. He has served as a guest editor in *IEEE Network*, *IEEE Communications Magazine*, and *Optical Networks Magazine*. Currently, he is lightwave series editor of *IEEE Communications Magazine*, and is on the Editorial Boards of *Wireless Personal Communications Journal* and *International Journal on Wireless and Optical Communications*.