Design and Implementation of Ultra-Low Latency Optical Label Switching for Packet-Switched WDM Networks

B. Meagher, G. K. Chang, G. Ellinas, Y. M. Lin, W. Xin, T. F. Chen, X. Yang, A. Chowdhury, J. Young, S. J. Yoo, C. Lee, M. Z. Iqbal, T. Robe, H. Dai, Y. J. Chen, and W. I. Way

***Abstract*—An ultra-low latency, high throughput internet protocol (IP) over wavelength division multiplexing (WDM) packet switching technology for next-generation internet (NGI) applications has been designed and demonstrated. This method overcomes limitations of conventional optical packet switching, which require buffering of packets and synchronization of bits, and optical burst switching methods that require estimation of delays at each node and for each path. An optical label switching technique was developed to realize flexible bandwidth-on-demand packet transport on a reconfigurable WDM network. The aim was to design a network with simplified protocol stacks, scalability, and data transparency. This network will enable the NGI users to send their data applications at gigabit/second access speed with low and predictable latency (<1 μsec per switch node), with a system capacity of beyond multi-Tb/s. Packet forwarding utilizes WDM optical headers that are carried in-band on the same wavelength and modulated out-of-band in the frequency domain.**

***Index Terms*—Optical communication, optical components, optical data processing, optical delay lines, optical fiber switches, optical label switching, packet switching, wavelength division multiplexing (WDM).**

I. INTRODUCTION

THE evolution of optical networks has progressed from single wavelength point-to-point data links, to multiple wavelength fixed data links, and on to fully reconfigurable multiple wavelength systems. Commercially available network elements (NEs) such as wavelength add–drop multiplexers, optical crossconnects, and wavelength amplifiers can be used to build flexible network architectures that can carry terabits of information and handle the interconnection of internet protocol (IP) routers, asynchronous transfer mode (ATM) switches, and synchronous optical network (SONET) NEs. But due to the relatively slow response time of these devices, on the order of hundreds of milliseconds to tens of seconds, these networks are limited to fast circuit switching applications, at best. In order to take full advantage of the possibilities afforded by wavelength division multiplexed data links, new kinds of network elements with vastly improved performance are required.

This report details a network implementation that uses a combination of techniques that have resulted in the ability to switch individual packets through an optical network element without converting the packet from optical to electrical format, and without the need to signal the network element in advance of the impending arrival of the packet. An optical label switching (OLS) testbed has been constructed to realize data switching and packet forwarding with ultra-low latency, simplified protocol stacks, and data transparency. Data applications generated by computer hosts have been transported for the first time over a packet-switched optical network based on information coded in an optical header modulated in the subcarrier domain. This paper summarizes the design and implementation of an ultra-low latency, high throughput packet application over an optical label-switched testbed, utilizing innovative hardware and software solutions. This network will enable its users to send data applications at gigabit per second access speed with low and predictable latency.

The approach described in this paper is based on a scheme employing a subcarrier-multiplexed (SCM) optical header (carried in-band on the same wavelength) attached to each optical packet, with an out-of-band modulation (thus making efficient use of the bandwidth) for packet forwarding. This method overcomes limitations of conventional optical packet switching, which requires buffering of packets, synchronization of bits, and estimation of delays in each node and each path. In the OLS testbed, packets are forwarded in the switch nodes based on the label contained in the SCM optical header without converting and reading the baseband data payload in the electrical domain. In this way the packet forwarding is independent of the payload data format, bit rate, and protocols. Other network features can also be incorporated in the optical header in addition to what has been included in the traditional IP header. Contention resolution techniques in such a network are currently being addressed utilizing deflection routing and wavelength interchange schemes, in addition to a preemption scheme based on packet priority. The header can be swappable

Manuscript received July 21, 2000; revised September 29, 2000. This work was supported in part by the DARPA funding agreement under the NGI initiative F30602-98-C-0216 and also in partby the Defense Advanced Research Projects Agency of the United States.

B. Meagher, G. K. Chang, T. F. Chen, A. Chowdhury, J. Young, M. Z. Iqbal, and T. Robe are with Telcordia Technologies, Red Bank, NJ 07701 USA (e-mail: brian@research.telcordia.com).

G. Ellinas is with the Tellium Inc., Oceanport, NJ 07757 USA.

Y. M. Lin and W. I. Way are with the Department of Communications Engineering, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C.

W. Xin is with Oplink Communications, San Jose, CA 95134 USA.

X. Yang, C. Lee, and Y. J. Chen are with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Catonsville, MD 21228 USA.

H. Dai is with Sorrento Networks, San Diego, CA, 92121 USA.

S. J. Yoo is with the Department of Electrical Engineering, University of California, Davis, CA 95616 USA.

Publisher Item Identifier S 0733-8724(00)10994-6.

at each node and therefore scalable for a large growing network similar to the MPLS approach developed by IETF [1], [2]. New ultra-fast, optical single sideband (OSSB) subcarrier header erasing and replacement techniques are being investigated and demonstrated for swapping the optical labels [3], and an implementation of this feature is discussed in this paper.

Previously, a number of other techniques have been presented that transmit the header ahead of the data payload [4], [5]. These techniques however, require a global knowledge of the network since they have to account for propagation delays at each fiber segment and at each network element in order to calculate the time interval required between the header and the payload. These techniques may also require a different wavelength for the optical header and a different wavelength for the data payload. Using the optical label switching method, the header and the payload packet travel together and there is no need to calculate network propagation delays at any node, thus achieving low-latency and simplified network control.

The remainder of this paper is organized as follows: Section II provides an overview of the system operation; Section III provides details regarding the subcarrier header transmitter and receiver designs; Section IV describes the switch node architecture; Section V describes the switch node controller and the interface between a Network Control and Management (NC&M) system and the network node controller. Section VI outlines the testbed demonstration performed, and discusses network features and considerations for OLS networks. Section VII describes a novel optically erasable label swapping technique, and Section VIII offers concluding remarks.

II. OVERVIEW OF SYSTEM OPERATION

The optical label switching system uses SCM to combine a data payload with a header containing routing information, as shown schematically in Fig. 1. The combined signal presents an analog representation of a data packet to the network that can be routed by switch nodes. The switch nodes inspect the header and set the state of an optical switch just prior to the arrival of the packet. Since the header and packet are coincident, this feat is accomplished by placing an optical splitter at the input to the switch node and removing approximately 10% of the power for header analysis. The other 90% of the power is launched into a short spool of fiber that delays the packet from reaching the switch until the header has been analyzed and the switch is set. The packet may then exit the delay line and travel through the switch and on to the next fiber transmission link.

The header recovery and analysis are performed by first sending the 10% sample of the optical signal into an analog optical to electrical converter. The electrical signal is then highpass filtered to remove the packet and leave only the header on the microwave carrier frequency. This signal is then transformed using a homodyne detection technique that results in the recovered header signal. The recovered header signal is launched into a commercial Fiber Channel deserializer and the resulting 10 B data "bytes" are injected into a header logic circuit. This circuit performs a fast lookup in a table that provides the appropriate optical switch setting for each packet address. The header logic circuit also processes priority information carried in the header, so that a high priority packet may interrupt a lower priority packet that is already using the desired output port. Each packet header also contains a duration field that lets the header logic circuit know how long the packet will need the switch. The header logic circuit then outputs a signal to a fast optical switch driver circuit that sets and holds the optical switch for the requested duration.

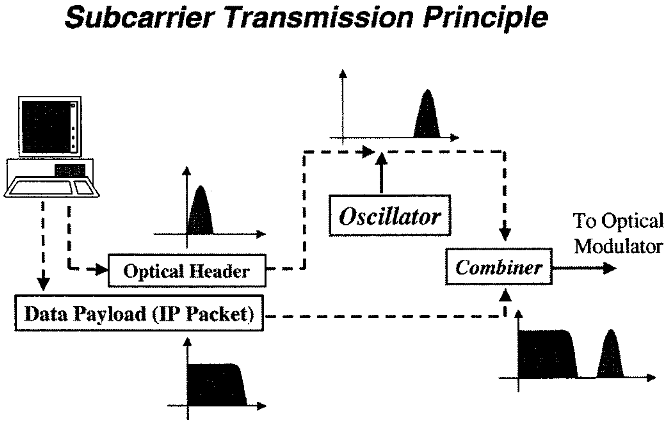


Fig. 1. Subcarrier transmission system.

The optical delay line must have a sufficient length to delay the packet long enough to accommodate the processing time in the header logic circuit plus the latency of the optical switch driver circuit and the switch itself. It is important that all delays be deterministic and bounded, so that the fiber delay line length may be calculated. To achieve this goal, the entire header logic circuit is implemented in one field programmable gate array (FPGA) logic chip that can accomplish its task in 25 ns. The most recent design operates a 2×2 lithium niobate switch in 80 ns, yielding a total switch node latency of 105 ns.

III. SUBCARRIER HEADER TRANSMITTER AND RECEIVER DESIGNS

The OLS system combines a data payload with a header using subcarrier multiplexing as described above. The combined signals are amplified before modulating the optical transmitter. The system uses amplitude shift keying (ASK) for encoding the microwave subcarrier used for optical header transmission. ASK allows for the use of an incoherent receiver, resulting in fast subcarrier detection. This is a better approach than the time consuming coherent detection process needed when either phase shift keying or frequency shift keying are used. Each of these techniques requires locking the local oscillator to the incoming signal.

An optical receiver at the switching node detects the optical header. The electrical output of the optical receiver is low-pass amplified and the baseband payload is filtered out by a high-pass filter. The subcarrier is split into two paths, with one path having a delay line. The two subcarrier signals are mixed and the intermediate frequency port signal is filtered out by a low-pass filter. The output from the low-pass filter is the recovered optical header. Successful operation has been achieved with zero error transmission for more than 10 days with a $2^{23} - 1$ pseudorandom bit stream at a data rate of 250 Mb/s.

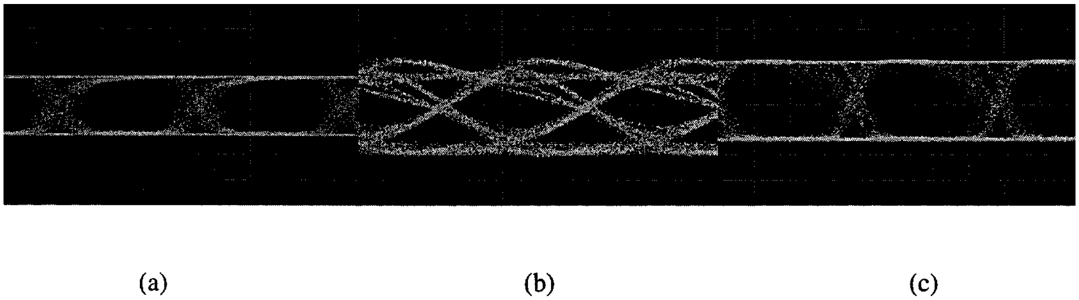


Fig. 2. Eye diagrams for received (a) data payload, (b) the recovered subcarrier header, and (c) the recovered subcarrier header after the digital amplifier.

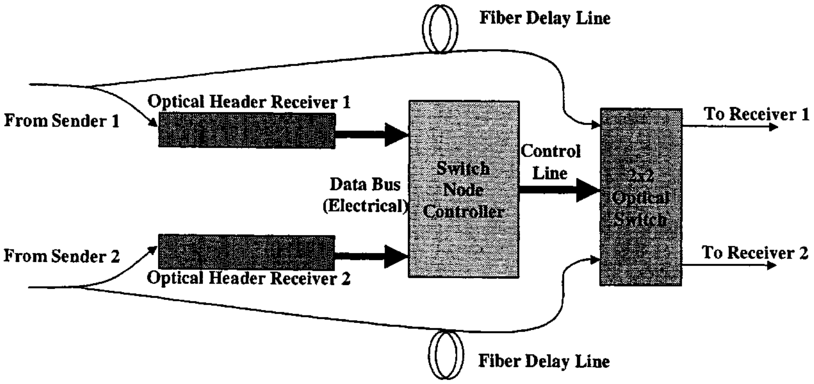


Fig. 3. Switch node architecture.

Fig. 2 shows the recovered eye diagrams for Fig.2(a) the data payload seen by the data link receiver, Fig.2(b) the signaling subcarrier header before a regenerating digital amplifier, and Fig.2(c) the signaling subcarrier header after 2R regeneration by the digital amplifier. The received eye is very wide open, guaranteeing error free transmission.

The subcarrier multiplexing technique becomes even more attractive for future networks, since new microwave components with excellent performance continue to be introduced commercially. The high demand for wireless technology has made these components inexpensive and ubiquitous, and all of the electronic components in the transmitters and subcarrier demodulators in this demonstration are available at low cost in microwave chip form.

IV. NETWORK SWITCH NODE DESIGN

The network design is based on the concept of network *edge nodes* and *switch nodes*. The edge nodes have transmitters and receivers as described in the previous section. The switch nodes function as packet switches in the network. Edge nodes are the sources and sinks of data payloads, and switch nodes are the optical label switching points within the network. Optical packets are sent from the edge nodes to a switch node, which then routes the message, using optical labels, to the corresponding outbound port of the switch node. A switch node may also include an optical label swapping function that can replace an existing label with a new one. Optical label swapping is not currently incorporated into the existing demonstration testbed but it is one of the main goals of this research program. A number of experiments have demonstrated the feasibility of a novel approach that erases and low-pass writes headers without disturbing the baseband signal. A detailed description of this work is presented in Section VII.

The system design is intended to separate the contents of an address header in the optical packet from the data conveyed in a message. At the switch node (shown in Fig. 3), most of the optical signal stays in the optical delay line (fiber) while a small portion of the signal is tapped in order to extract the routing information from the subcarrier multiplexed optical header. Wavelength demultiplexers may be used to separate the optical packets on different wavelengths so that packets can be processed on a per wavelength basis in the switch node.

A tap on the input fiber to the switch node sends 90% of the received power to an optical fiber delay line. The data packet waits in the delay line while the switch node controller sets the optical switch to the correct state. At the same time, the remaining 10% of the optical power is sent to a subcarrier receiver as described in Section II. The subcarrier optical header receiver recovers the optical header and sends it in an electrical format to the switch node controller. The switch node controller extracts the address and priority information from incoming messages, correlates it with the information entering the switch node from other switch input fibers, and calculates and implements the appropriate switch connections. Successfully routed packets may travel through several switch nodes before eventually reaching the intended edge node. The packet is received in the edge node by an interface card where the baseband signal is extracted and presented electrically in the same format in which it was originally sent.

V. OPTICAL-LABEL SWITCH NODE CONTROLLER

At the heart of each switch node is an optical label switch node controller, shown in Fig. 4. The switch node control

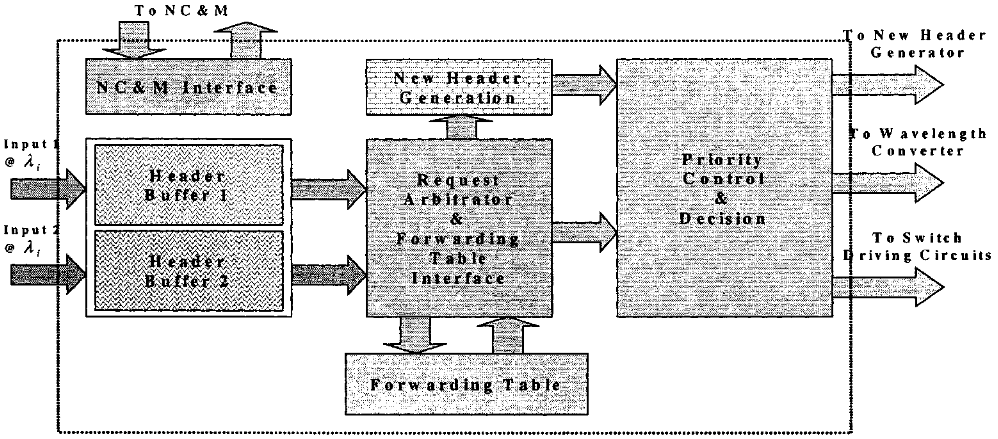


Fig. 4. Optical switch controller.

functions are implemented in hardware in a large field programmable gate array (FPGA) logic chip. The major functions of a switch node controller can be broken into NC&M functions and packet forwarding functions[1].

The NC&M functions include forwarding table update and alarm and status reporting, which occur on a relatively slow time scale, on the order of milliseconds. The packet forwarding functions are extremely time sensitive and must occur in a bounded or deterministic amount of time, on the order of nanoseconds. Packet forwarding functions begin to operate as soon as a new incoming packet header is detected. All forwarding functions must be completed before the packet exits the optical delay line discussed previously. The delay line length required in a switch node can then be calculated once the maximum packet forwarding function time is known.

An NC&M interface is also provided for communication between a switch node controller and a network management system (NMS), as shown in Fig. 4. The NC&M interface provides a path for forwarding table update and maintenance, and also provides a path for reporting traffic statistical information. An NMS sends control messages to the switch node controller using label switch control protocol (LSCP) and messages are carried by a TCP/IP socket connection. One of the primary functions of the NMS system is to update the switch node translation table. This table provides a mapping between optical header labels and switch output ports, and is used in conjunction with data from arbitration logic to route packets through the switch node. Other functions of the NC&M interface include forwarding alarm messages and reporting port usage statistics.

The rest of this section discusses packet forwarding functions.

After the SCM receiver demodulates an optical header, the serial bit stream is passed to a data deserializer. The deserializer recovers the bytes of the OLS header and places them in a header buffer. Each wavelength of each input fiber has a corresponding header buffer for incoming headers. Packet processing is initiated upon reception of a start-of-header byte into a header buffer. The last byte of the header contains an error code byte that is computed by the sender. The integrity of the header may be tested at the switch node by calculating this code and comparing it with the last byte of the received header. Any discrepancy indicates a transmission error.

A header buffer stores a complete header, which includes a source and destination label, priority and authorization bits, and packet duration. In order to obtain the proper switch configuration and new label, the header buffer generates a read request to the forwarding table based on the label and priority bits. The authorization bits can allow for wavelength interchange, alternate path routing, and optical header swapping. A packet duration field provides information that determines the length of time, including interpacket guard band time, that the switch will be occupied. The request arbitrator provides for the orderly access to the forwarding table according to incoming label and priority information. The forwarding table interface generates address and control signals used to access the forwarding table. The table can be implemented using a dual-port SRAM allowing one port for packet processing functions and one port for NC&M functions. If a swappable header scheme is used in the switch node, a new label is found in the forwarding table and buffered in the controller, waiting to be synchronized with the data payload.

The priority control and decision section then decides whether the packet should be forwarded out of a switch node output port, wavelength converted, or dropped, according to the forwarding table content and current output port status. After the decision is made, the controller will produce the appropriate control signals to set the optical switch and release the packet.

VI. PROOF-OF-CONCEPT DEMONSTRATION

A. Network Testbed Architecture

A testbed has been constructed and is fully functional in operation. It provides the first demonstration of computer-to-computer communication using an optical label switching mechanism. A photograph of the testbed is shown in Fig. 5. The testbed and supporting technologies described in this paper were demonstrated in the exhibition area of the Optical Fiber Communications conference held in Baltimore, Maryland on March 5–10, 2000.

---

[1] The switch node controller has been designed with the goal of supporting label swapping features, even though these features are not present in the current version of the switch node. As system development progresses, these features will be integrated in the switch node hardware.

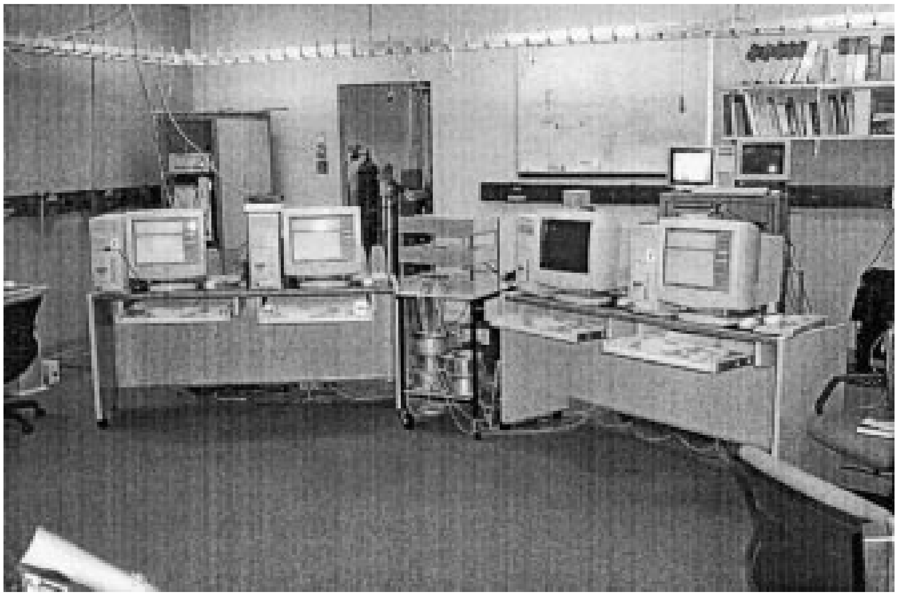


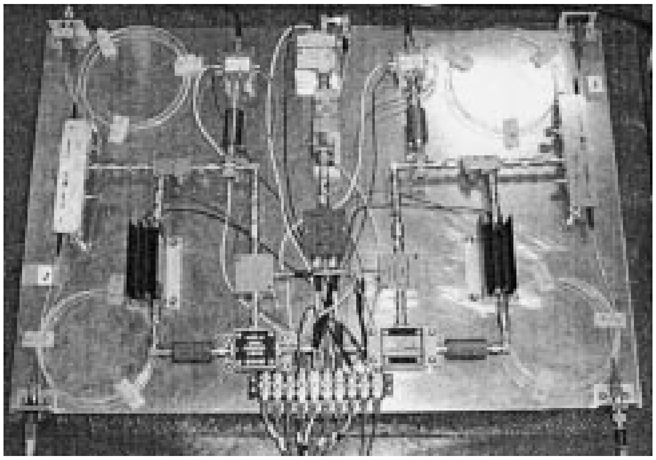
Fig. 5. Optical label switching laboratory.

The elemental building blocks of the network testbed are host/network interface cards (NICs), transmitters, switch nodes, and receivers. The testbed has four Windows NT workstations equipped with fiber channel-based NICs. These cards generate the data payload and optical header that are transmitted over the network. The header information is subcarrier multiplexed with the payload and then used to modulate a continuous wave laser. At the switching nodes, the optical header must be extracted from the signal by a subcarrier receiver. Both of these circuits are shown in Fig. 6.

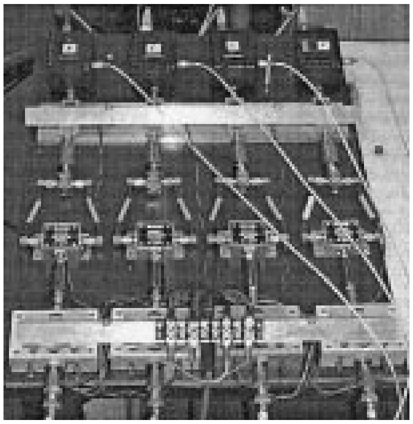
The network hardware is configured as shown below (Fig. 7). Workstations 1 and 2 can send and receive data from Workstations 3 and 4. The switch node logic modules use information contained in the packet header to route the packet and resolve contention between packets that compete for access to the same output port.

Data packets are produced by applications running on the host computers (Windows NT workstations) and are passed down through software drivers to the network interface card. The application also supplies a packet destination address to the NIC. The NIC then provides the packet and header to the transmitter. The NIC card design utilizes commercially available transceiver chips operating at 250 Mb/s. It supplies the packet and header to the transmitter as an 8 B/10 B encoded digital electrical signal.

The transmission scheme uses the subcarrier multiplexing technique described in Section 3 to place the packet and header on the optical fiber simultaneously. The header is mixed with a local oscillator operating at 3 GHz. The resulting signal is mixed with the packet and then injected into the modulation port of a lithium niobate modulator. A continuous wave laser is used as input to this modulator, and the output of the modulator is launched into the transmission fiber. The end result is an analog optical signal that contains the original packet data in the baseband signal, and an easily separated header signal traveling along with the packet, but shifted up in frequency by 3 GHz. This signal may pass through optical amplifiers and switch nodes on its way to the final destination. The amplifiers can compensate for span losses without disturbing the relationship between the header and packet.



(a)



(b)

Fig. 6. (a) Subcarrier Transmitter. (b) Optical header receiver.

Eventually the packet and header reach a receiver at the final destination. An inexpensive receiver is well suited to the task

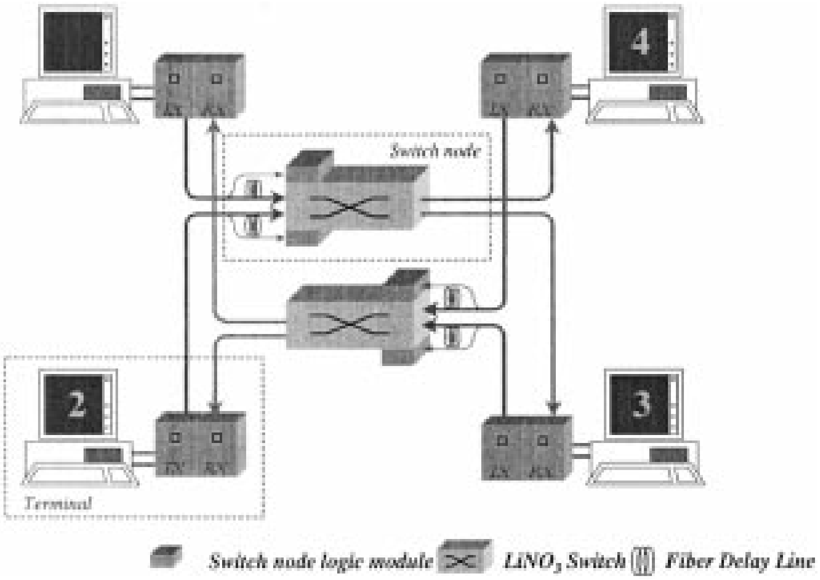


Fig. 7. Schematic layout of demonstration network.

of recovering the packet from the baseband signal. Since such a receiver does not have enough bandwidth to accept the header, it saves the cost of a low-pass filter that would be required if a higher bandwidth receiver was used. The packet, now transformed back into its electrical form, is presented to the host computer NIC card, which has a commercial Fiber Channel receiver chip. The process is complete when the NIC card software driver presents the packet to an application program running on the machine.

B. Network Experiments

A terminal communications program was written that allows users to send messages to each other over the optical label switched network. The user types in a message, selects the destination address, sets the priority of the outgoing message (low or high priority), and then hits the "Send" button to initiate transmission.

Packets containing information specified in the user interface program are assembled and an Optical Label Switching header is built. This header contains source and destination address information as well as priority information. The packet contains a message for the remote user application and a checksum that is used to verify the integrity of the transmitted message. The data to be sent (packet and header) are placed in a buffer by the user program. The address of this buffer is then passed to a kernel mode device driver through an application program interface (API) provided by the NIC card vendor. This API contains a function call that passes data out through the NIC card and reports its success back to the application. The user interface program also registers itself with this API to handle the reception of messages from the network. The hardware on the NIC card is configured to transfer incoming messages into a buffer supplied by the user interface program via a high-speed direct memory access (DMA) transfer and then triggers a hardware interrupt which results in a call to the user interface program's receive function. This function is used to place the received data into the Receive Text window of the application program.

Using the network described above various test cases were demonstrated. The functions shown included Normal Send and Receive operation, First-Come First-Served switch operation for packets arriving with equal priority levels, Switch State preemption by High Priority packets, and NC&M system update of switch forwarding tables. These scenarios are described as follows.

1) Normal Send and Receive Operation: An address is assigned to each workstation. We typically assign addresses 1, 2, 3, and 4 as shown in Fig. 7. Messages can be typed in at any workstation. Associated with the message are an Outbound Label and a Priority, assigned via radio buttons from the user interface. A destination workstation is associated with an Outbound Label. The association is created by the Forwarding Table (Table I) supplied by the NC&M system. A power-on default table associates Label 1 to Workstation 4 and Label 2 to Workstation 3, for messages sent from Workstations 1 and 2. The default table associates Label 1 with Workstation 1 and Label 2 with Workstation 2, for messages sent from Workstations 3 and 4. Messages are displayed in the receiving workstation's Receive Text box. For convenience, there is a mechanism for continuously sending test packets with a given Label and Priority. There is also a means for periodically sending packets with a settable (in milliseconds) delay time between packets.

2) First-Come First-Served Switch Operation for Packets Arriving with Equal Priority Levels: In order to demonstrate how the switch resolves contention between packets, a means for guaranteeing the presence of a packet on a switch control port had to be devised. This task is accomplished by sending a packet that the switch node interprets as an infinite length packet. If, for example, Workstation 1 sends an infinite packet to Workstation 3, it will force the switch into the straight through or "BAR" state. A packet from Workstation 2 to Workstation 3 will be unable to affect a change from BAR to CROSS state as long as this

TABLE I
FORWARDING TABLE

Input Port	Input Label	Output Port	NC&M Bits
0	0	0	00
0	1	0	00
0	0	1	01
0	1	1	01
0	0	1	10
0	1	0	10
0	0	0	11
0	1	1	11

second packet has a priority less-than-or-equal to the priority of the packet that arrived first.

3) Switch State Preemption by High Priority Packets: A higher priority packet will always be able to get through to its destination. This can be demonstrated in the network of Fig. 7 by setting up the experiment in the following manner: A low priority message can be sent from Workstation 2 to Workstation 4 as an infinite length packet. If a high priority packet is now sent periodically from Workstation 1 to Workstation 4 (every 500 ms), the switch node controller will hold the switch in the BAR state for the infinite packet, but will switch to CROSS state every half second for the duration of the test packet (approximately 250 μs). A light on the switch controller blinks indicating successful preemption by the high priority packet.

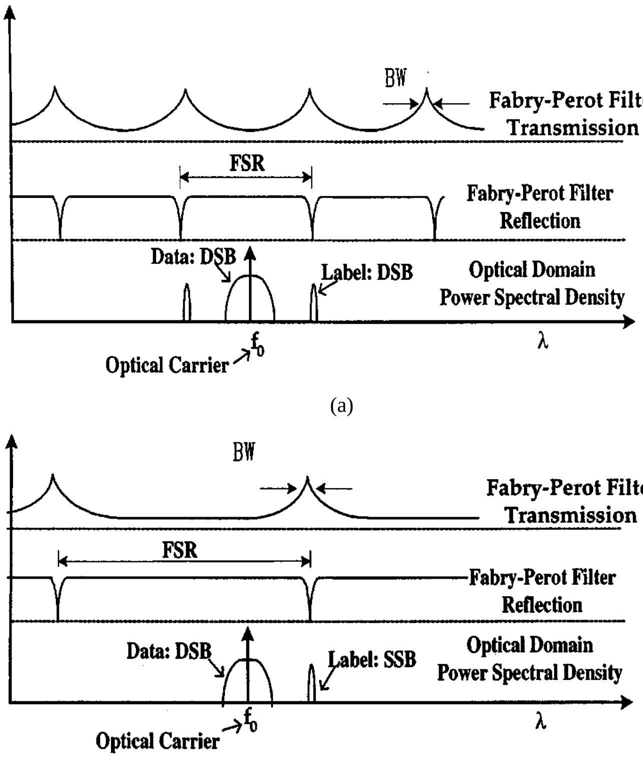
4) NC&M System Update of Switch Forwarding Tables: The NC&M system can send information directly to the switch node controller to change the association between Labels and Outbound Ports (destination Workstations) on a per-fiber basis. For example, an Outbound Label of 1 coming from Workstation 3 can mean the packet will be destined for Workstation 1 or Workstation 2, and an Outbound Label of 2 can mean the packet will be destined for Workstation 1 or Workstation 2. If the Forwarding Table specifies that Label 1 and Label 2 are both destined to Workstation 1, then the Forwarding Table has been set up to make a circuit switch rather than a packet switch. The forwarding table from one of the fibers is shown in Table I.

5) Performance: The performance of the OLS system has been characterized. Measurements were made of the time it took for the switch node to perform decision and switching functions. The result was that the switch node could accomplish its task in well under one microsecond. The breakdown is shown as follows:

- 40 ns to acquire destination and priority information from the data stream;
- 25 ns to arbitrate control of the switch and output a new switch state;
- 10 ns for the $LiNbO_3$ switch to operate;
- less than 100 ns for the driver circuit to cause switching in the $LiNbO_3$ switch.

C. Network Features and Considerations

The testbed described here provides the lowest possible latency for the transmission of data. Packets are delayed at switch nodes just long enough to determine the correct outbound port and to set up a connection from the input fiber to that port. If improved technology results in a quicker decision or switching time, then the fiber delay line can be shortened. The delay line can be a different length at every switch node—it need only be appropriate for the type of technology used in each node.



(a)

(b)

Fig. 8. Optical notch filter responses and optical spectra for (a) ODSB and (b) OSSB subcarrier labels.

The network design is all-optical between end stations. When each packet leaves a host computer it sees one long fiber to the destination computer. This "virtual fiber" can be different for every packet that leaves a host.

The steering information in the header is decoupled from the data in the packet. It can be a different format or data rate. Packets do not even have to be digitally encoded. For example, analog video could be sent as the payload (perhaps a better term than packets in this case) as long as the appropriate headers are sent frequently enough to maintain the connection.

Every packet on a given path experiences the same delay in traveling to its destination. Packet sequencing and timing is preserved by the network, which is an important consideration for real time computing applications and for streaming audio and video systems.

In a larger network there may be a choice of paths to a given destination. A switch node could be built that would allow packets to take advantage of this or not, as appropriate for each packet. If the packet is blocked at a given switch node a bit in the header could inform the switch node that it is acceptable to deflect the packet to an alternate path. If the bit is not set, it means it is better to drop the packet than have it arrive late.

The use of optical wavelength translation devices affords many more possibilities for dealing with contention in the switch node. Typical WDM systems use a number of fixed optical wavelengths spaced 50, 100, or 200 GHz apart. With a

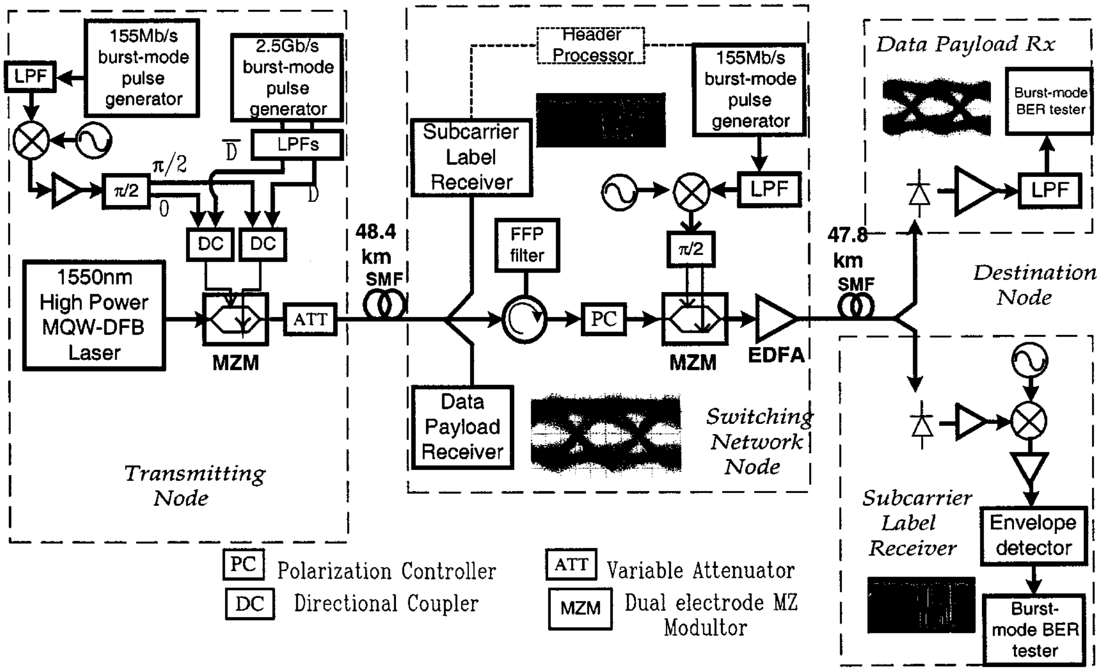


Fig. 9. Experimental setup.

wavelength converter a packet coming in on λ_i could be sent back out on λ_j, for example. The employment of wavelength deflection would work in an analogous manner to path deflection, with the added benefit that the optical packet does not have to perform additional optical hops. Wavelength deflection appears to be very attractive in this regard since it does not significantly affect the transmission distance that the packet travels. If the wavelength translation is sufficiently fast, using solid state conversion techniques for example, the time delay variation of packets taking first choice and alternative paths can be very small.

VII. Label Swapping Using Erasable Optical Single-Sideband Subcarrier Labels

A desirable property for label switching networks is the ability to rewrite all or part of the routing header information at switching points within the network. This technique is used in ATM networks to enable the use of locally significant labels (the VPI/VCI bits) that can be low-pass used with a different meaning in different parts of the network. Internet Protocol [6] uses rewritable hop counters (TTL bits) to help prevent routing loops, and Multi-Protocol Label Switching [7] uses fields in the Label for both of these purposes. Recently, several complicated optical label swapping techniques have been demonstrated [8], [9]. In this paper, we propose a novel technique to simplify and enable numerous repeated optical label swapping operations over a national scale optical network, while at the same time keeping the multi-gigabit/sec data payload intact.

The subcarrier label swapping is accomplished by first erasing the original label in the optical domain, after which the light is low-pass modulated with a new subcarrier label at the same microwave carrier frequency. A notch filter is employed to erase the old subcarrier by using the reflective part of a voltage-tunable fiber Fabry–Perot (FFP) filter, as shown in Fig. 8. However, note that the optical double sidebands (ODSB) of the data payload and subcarrier label are located on both sides of the optical carrier, as illustrated in the lower part of Fig. 8(a). To erase the subcarrier label, both subcarrier sidebands must be notched out. But this requires that the free spectral range (FSR) of the filter be *exactly* equal to the separation of the two subcarrier sidebands. Furthermore, the notch filter must present a sharp and narrow notch so that the data payload is not affected. Consequently, it is difficult to design and manufacture an FFP filter that satisfies both requirements. This problem, can be solved by using an optical single-sideband (OSSB) modulation technique [10] to transport the subcarrier label, so that the resultant optical spectrum contains only one subcarrier label sideband, as shown in Fig. 8(b). Another important advantage of using an OSSB microwave subcarrier label is that, by avoiding the fiber dispersion-induced carrier suppression effect [10], its transmission distance between switching nodes can be more than several hundred kilometers *without* dispersion compensation.

The experimental setup for this approach is shown in Fig. 9. The transmitting node uses a 50 mW, 1551 nm DFB-MQW laser and a two-electrode LiNbO_3-external Mach–Zehnder modulator (MZM) with a 3-dB bandwidth of 20 GHz and an insertion loss of 4 dB. A bursty 155-Mb/s ASK subcarrier label at 12 GHz is applied to a hybrid coupler whose two outputs have $\pi/2$ phase shift with respect to each other. These two outputs are then combined with a bursty 2.5-Gb/s data and inverted data signal, respectively, through two directional couplers. Note that the label and payload bursts are both randomly generated, with each label and payload burst consisting of 20 bits and 53 bytes, respectively. Two low-pass filters (LPFs) with a 3 dB bandwidth of 2.4 GHz are used to prevent the tails of the 2.5 Gb/s NRZ data from interfering with the ASK subcarrier. The

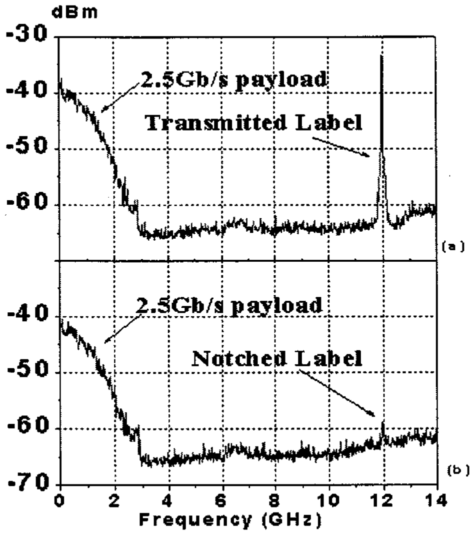
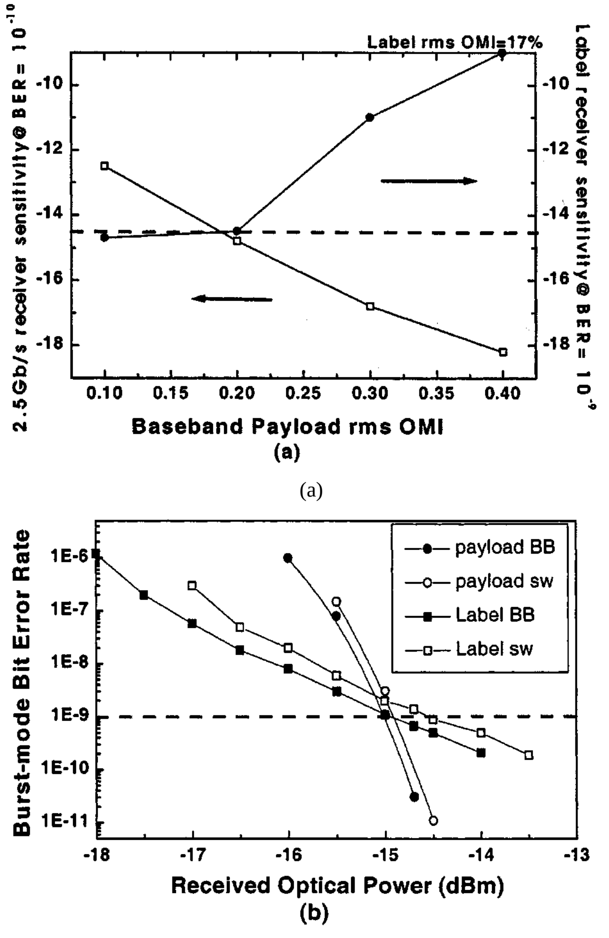


Fig. 10. Spectrum analyzer displays (a) before and (b) after the OSSB subcarrier label is suppressed.



(a)

(b)

Fig. 11. (a) 2.5 Gb/s and subcarrier label receiver sensitivities as functions of the 2.5 Gb/s rms OMI. (b) Burst-mode BER versus receive optical power for 2.5 Gb/s payload (random, bursty 53-byte sequence), and for subcarrier label (random, bursty 20-bit sequence) before and after remodulation.

two combined NRZ data and ASK subcarrier outputs are then used to drive the two electrodes of the MZM, respectively. Fig. 10(a) shows the spectrum of the transmitted baseband 2.5 Gb/s data and the ASK microwave subcarrier label.

In the switching node, the received optical signal is split into three paths. The first path is simply a data payload receiver which consists of a 50-Ω-terminated photodiode, a DC-14 GHz amplifier, and a LPF. The second path is a label receiver which consists of a 50-Ω-terminated photodiode, an 8–12 GHz amplifier, a downconverter with IF frequency at 550 MHz, and an ASK envelope detector. The eye diagram of the received data payload and the bursty 20-bit label are shown in the insets of Fig. 9. The third path is the path where the label swapping takes place. The optical signal is reflected by an FFP filter via an optical circulator, so that the old subcarrier label is notched out. The FFP filter has an FSR of 1500 GHz, a finesse of 100, and a reflection loss of 1.5 dB. Fig.10(b) shows the microwave spectra after notching. The subcarrier label is suppressed by 25 dB at 12 GHz, while the 2.5 Gb/s payload experiences only 2 dB loss. The optical signal which had its subcarrier label suppressed is subsequently passed through a polarization controller and another MZM. At the MZM, the optical signal is remodulated with a new ASK subcarrier label, which has the same carrier frequency and optical modulation index (OMI) as those of the old subcarrier label. At the output of the switching node, an EDFA with an output power of 12 dBm is used. Note that the lengths of the first and second fiber spans in Fig. 9 are 48.4 and 47.8 km, respectively.

In the destination node, the same data payload and subcarrier label receivers as those in the switching node are used. The bit-error-rate (BER) of the payload and the label are measured by a burst-mode BER tester. With its burst clock and trigger, the BER tester can synchronize the received payload or label pattern to the stored pattern such that the corresponding BER per burst can be obtained. The bursty 20-bit sequence of the received new subcarrier label (after one-time label swapping and 47.8 km of transmission) and the eye diagrams of the baseband payload (after 96.2 km of transmission) are shown in the insets of Fig. 9. Since ~25 dB of the old label power is suppressed by the notch filter, the residual old label has negligible effect on the low-pass modulated new label. However, the carrier-to-noise ratio (CNR) of the new label is degraded by the intermodulation noise caused by the beating between the baseband data and the new subcarrier label. This CNR degradation results in power penalties in receiver sensitivity, and is dependent on the route-mean-square (rms) OMI of the baseband data. Experiments showed that if the data payload receiver and the ASK subcarrier label receiver require the same received optical power, and if the ASK subcarrier label's OMI is 17%, the optimum rms OMI of the baseband data is 20%, as shown in Fig. 11(a). Under this condition, there is less than 0.5 dB power penalty (for BER $< 10^{-9}$) when comparing the burst-mode BER performance of the bursty subcarrier label before and after remodulation (which is due to the intermodulation noise just mentioned), as can be seen in Fig. 11(b). The bursty 2.5 Gb/s payload is essentially unaffected by the switching node, as can be seen from the burst-mode BERs before and after label swapping.

VIII. CONCLUSION

We have implemented, for the first time, a proof-of-concept optical-label switching network testbed demonstration using four computer-based hosts and two optical cross-connects as the switch nodes, in order to perform data switching and packet forwarding. This testbed provided the first ever demonstration of computer-to-computer communication using an optical-label switching mechanism. Data applications generated by the hosts have been transported over a label-switched optical network based on information coded in optical headers modulated at 250 Mb/s and multiplexed at 3 GHz subcarrier frequency. At the switching nodes, the optical headers are stripped and examined, and the optical switches are switched accordingly. The total delay of approximately 105 ns at the switching node is dominated by the speed of the driver circuit for the lithium niobate optical switches. The header content in this proof-of-concept, early demonstration, carries preamble bytes, forwarding labels, packet duration, and priority information. This network will enable the future users to send data at high access speed of >1 Gb/s with low and predictable latency of <1 μsec per node in high throughput, scalable NGI networks.

Significant progress has also been achieved in the area of label swapping. We have successfully demonstrated a novel subcarrier label swapping technique by using OSSB modulation and an optical notch filter. The label swapping can be repeated numerous times over a long distance network without dispersion compensation because the subcarrier label is reset at each switching node and because of the inherent transmission advantage of OSSB modulation. Future work includes the integration of optical label swapping in the switch node hardware and in turn in the testbed demonstration.

The demonstration system is evolving as new technological growth areas are identified. In particular, attention is being paid to putting greater capability and more intelligence in the switch node. Larger routing tables and the ability to gather statistics for network management functions will be required. Multiple wavelengths and higher line speeds are also anticipated. Furthermore, we have been studying and designing network control interfaces, as well as label distribution and switching control protocols between the IP and WDM layers, which will interoperate with the existing Internet protocol.

The final goal of this project is a testbed demonstration that will incorporate optical-label switching with network control interfaces, and feature multiple wavelengths, multiple priorities, contention resolution techniques, optical label swapping and multigigabit packet rates.

ACKNOWLEDGMENT

The authors would like to thank J. Wei, W. Leland, M. Post, S.-Y. Park, K. Liu, J. Pastor, A. Broscius, R. Talpade, S. Khurana, B. Mason, and M. Meagher for their contributions.

REFERENCES

[1] A. Viswanathan, N. Feldman, Z. Wang, and R. Callon, “Evolution of multiprotocol label switching,” *IEEE Commun. Mag.*, vol. 36, no. 5, pp. 165–173, May 1998.

[2] D. J. Blumenthal, A. Carena, L. Rau, V. Curri, and S. Humphries, “WDM optical tag switching with packet-rate wavelength conversion and subcarrier multiplexed addressing,” in *Proc. Tech. Dig., OFC’99*, San Diego, CA, Feb. 1999, ThM1, pp. 162–164.

[3] W. Way, Y. Lin, and G. K. Chang, “A novel optical label swapping technique using erasable optical single-sideband subcarrier label,” in *Proc. Tech. Dig., OFC’00*, Baltimore, MD, Mar. 2000, WD6, pp. 1–4.

[4] J. Wei, “The role of DCN in optical WDM networks,” in *Proc. Tech. Dig., OFC’00*, vol. , Baltimore, MD, Mar. 2000, FI1, pp. 1–3.

[5] J. S. Turner, “WDM burst switching for petabit data networks,” in *Proc. Tech. Dig., OFC’00*, Baltimore, MD, Mar. 2000, WD2, pp. 1–4.

[6] (1981, Sept.) Internet Draft. [Online]. Available: http://www.ietf.org/rfc/rfc0791.txt

[7] (1999, Sept.) Internet Draft. [Online]. Available: http://www.ietf.org/internet-drafts/draft-ietf-mpls-framework-05.txt

[8] D. J. Blumenthal, A. Carena, L. Rau, V. Curri, and S. Humphries, “All-optical label swapping with wavelength conversion for WDM-IP networks with subcarrier multiplexed addressing,” *IEEE Photon. Technol. Lett.*, vol. 11, pp. 1497–1499, Nov. 1999.

[9] X. Jiang, X. P. Chen, and A. E. Willner, “All optical wavelength independent packet header replacement using a long CW region generated directly from the packet flag,” *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1638–1640, Nov. 1998.

[10] G. H. Smith and D. Novak, “Broad-band millimeter-wave (38 GHz) fiber-wireless transmission system using electrical and optical SSB modulation to overcome dispersion effects,” *IEEE Photon. Technol. Lett.*, vol. 10, pp. 141–143, Jan. 1998.

B. Meagher, photograph and biography not available at the time of publication.

G. K. Chang, photograph and biography not available at the time of publication.

G. Ellinas, photograph and biography not available at the time of publication.

Y. M. Lin, photograph and biography not available at the time of publication.

W. Xin, photograph and biography not available at the time of publication.

T. F. Chen, photograph and biography not available at the time of publication.

X. Yang, photograph and biography not available at the time of publication.

A. Chowdhury, photograph and biography not available at the time of publication.

J. Young, photograph and biography not available at the time of publication.

S. J. Yoo, photograph and biography not available at the time of publication.

C. Lee, photograph and biography not available at the time of publication.

M. Z. Iqbal, photograph and biography not available at the time of publication.

T. Robe, photograph and biography not available at the time of publication.

H. Dai, photograph and biography not available at the time of publication.

Y. J. Chen, photograph and biography not available at the time of publication.

W. I. Way, photograph and biography not available at the time of publication.