The AMOEBA Switch: An Optoelectronic Switch for Multiprocessor Networking Using Dense-WDM

Ashok V. Krishnamoorthy, *Member, IEEE,* Joseph E. Ford, *Member, IEEE,* Fouad E. Kiamilev, *Member, IEEE,* Richard G. Rozier, *Member, IEEE,* S. Hunsche, Keith W. Goossen, *Member, IEEE,* B. Tseng, James A. Walker, J. E. Cunningham, W. Y. Jan, and Martin C. Nuss, *Member, IEEE*

***Abstract*— We present a single-chip asynchronous multiprocessor optoelectronic bit-sliced arrayed (AMOEBA) crossbar switch. The AMOEBA switch addresses the challenge to produce a large-scale, nonblocking packet switch through dense integration of photonic devices directly onto silicon VLSI circuits. Optoelectronic-VLSI technology is used to integrate the switch fabric, routing controller, packet buffers, line interface circuits, and optoelectronic conversion devices on a single chip. We show how free-space optical interconnects and wavelength-and-space–division-multiplexed networking on single-mode fibers can provide switched interconnection between multiple nodes in a distributed computing environment. An optomechanical transceiver package accomplishes the free-space-to-fiber interfacing. We report the implementation and testing of the key components of a 16-channel AMOEBA prototype switch with a potential capacity of 12.8 Gb/s (or 800 Mb/s/channel), and capable of switching 16 million packets per second.**

***Index Terms*—CMOS integrated circuits, fiber networks, flip-chip devices, integrated optoelectronics, modulators, optical arrays, optical interconnections, quantum-well devices, switching, wavelength-division multiplexing.**

I. Introduction

IT IS BECOMING evident that networking multiple general-purpose processors or workstations is an efficient and cost-effective path to high-performance computing. In such distributed computing environments, the interconnection network is typically the performance bottleneck, especially as electrical interconnect technologies struggle to keep pace with increasing demand for more connectivity and rapidly diminishing processor clock-cycle times. Historically, the network transmission bandwidth (in Megabits per second) has kept pace with the processor computing bandwidth (in millions of instructions per second), as shown in Fig. 1. However, as the performance of both the processors and the network interface cards scale up, the switching/routing functions required of the network will become the limiting bottleneck. The demands made on such a multiprocessor interconnection network will include: 1) high throughput communication; 2) low contention; and 3) simple communication protocols. Additional features of low latency data transport and packet switched operation (potentially with variable-length packets) may be desirable. Markets for such networks at the campus-wide level (i.e., local area networks) as well as within an equipment room (system-area networks) are now emerging.

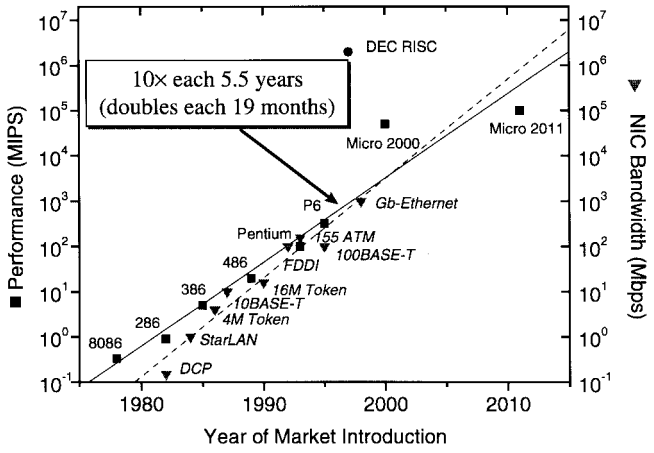


Fig. 1. Evolution of workstation computing performance, measured approximately in millions of instructions-per-second (MIPS), versus network bandwidth, measured in megabits-per-second of the network interface card (NIC).

A number of photonic technologies are candidates for multiprocessor switching and routing networks. Performance, scalability, and cost will be critical. We are investigating interconnection networks based on fiber and surface-normal optical interconnects for data-transport and silicon VLSI for switching. We employ an optoelectronic-VLSI (OE-VLSI) technology capable of supporting several thousand optical input/output (I/O) devices with each device capable of data transmission in excess of 1 Gb/s when driven with standard CMOS circuits. The specific OE-VLSI technology we discuss is based on the hybrid flip-chip bonding of GaAs–AlGaAs multiple-quantum-well (MQW) diodes onto CMOS chips followed by the removal of the GaAs substrate [1]. This technology permits the design of CMOS circuits with submicrometer feature-size and full circuit and layout optimization typical of complex VLSI chips [2], [3].

In this paper, we present the application of this technology to an asynchronous multiprocessor optoelectronic bit-sliced array (AMOEBA) switched network. The AMOEBA switch is intended to function as a high-performance central switch on a data network serving multiple processors or computers in tightly coupled computing environments (single frame), loosely coupled environments ($\sim$100 m), or campus-wide networks ($\geq$1 km, limited by fiber attenuation at the operating

Manuscript received November 16, 1998; revised April 29, 1999.

A. V. Krishnamoorthy, J. E. Ford, R. G. Rozier, S. Hunsche, K. W. Goossen, B. Tseng, J. A. Walker, J. E. Cunningham, W. Y. Jan, and M. C. Nuss are with Bell Laboratories, Lucent Technologies, Holmdel, NJ 07733 USA.

F. E. Kiamilev is with the University of North Carolina, Charlotte, NC 28223 USA.

Publisher Item Identifier S 1077-260X(99)06053-0.

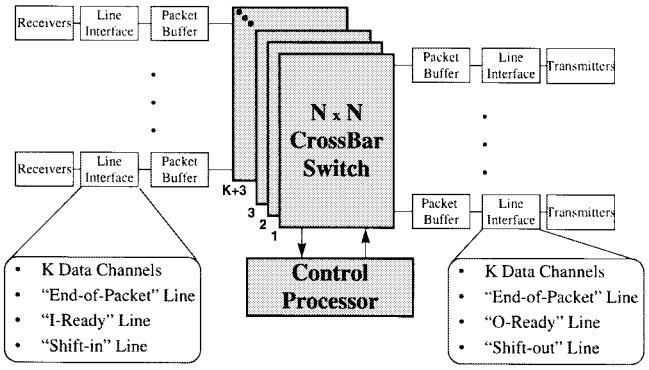


Fig. 2. Architecture of the N-channel AMOEBA switch. Multiple crossbars (one-per-bit), line interfaces, FIFO-buffers, receivers, transmitters, and controller are incorporated into a single optoelectronic chip.

wavelength of 850 nm). The heart of the AMOEBA switch architecture is a scalable optoelectronic VLSI switch chip that integrates the crossbar switch fabric, routing controller, packet buffers, line interface circuits, and optoelectronic conversion devices in onechip.

Section II describes the system architecture. The design and implementation of the AMOEBA switch is presented in Section III. Section IV discusses the design of a wavelength-and-space division multiplexed fiber/free-space optical network for transporting the data. The implementation of a dense-WDM (wavelength division multiplexing) fiber transceiver package is presented in Section V. Testing results for dense-WDM transmission are presented in Section VI. A summary is included in Section VII.

II. System Architecture

The purpose of the AMOEBA chip is to provide byte-wide, switched optical interconnection at the processor clock rate (50 MHz–1 GHz). This central switch uses multiple on-chip NxN crossbars to connect N independent (asynchronous) input channels with K synchronous data bit-streams per input channel to N output channels with a corresponding K data bit-streams per output channel (Fig. 2). Each bit-stream is encoded as a separate optical I/O line to the VLSI chip. An end-of-packet line, and two additional synchronization lines (for clocking) and are included for each channel. End-to-end flow control is performed independently for each channel to avoid packet loss. We anticipate future networks comprising between 16-to-64 processors with 16-to-64-bit words operating at a line rate of 1 Gb/s (1 GByte/s per channel); this would imply a switch with up to 4-Tb/s throughput. For the current AMOEBA chip, crossbars were chosen as the interconnection unit because of their fully nonblocking connection properties and their suitability to a CMOS VLSI implementation. For applications where $N \gg 64$, it may become necessary to replace the crossbars with switch-fabrics that have reduced complexity and nonzero contention, e.g., multistage interconnection networks (MIN's); such networks have been widely studied in the literature. For instance, networks that provide a continuous tradeoff between a MIN and a crossbar, trading internal contention for fanout (or switch complexity) have been investigated elsewhere [4]. Other centralized switching architectures that minimize the probability of packet loss via output-port buffering [5], [6] or multiple parallel buses [7], and distributed network architectures such as wormhole-routing networks [8], have also been considered for implementation with OE-VLSI technologies.

The AMOEBA chip can be used in applications where a large amount of information (potentially originating up to 100 m or more away) would have to be switched and routed to the different processors/servers on the network. Such a system is one where the interconnect has a large aspect-ratio (length of interconnect versus density of connectorization) which puts undue stress on electrical packaging alternatives, and is representative of a class of systems where optical interconnects can be used to advantage [9].

A key feature of the AMOEBA architecture is that each of the crossbars can be identically configured, under the assumption that individual bit-streams of an input channel are destined to the same output channel. Hence, a single controller designed for an $N \times N$ switch can be used to control the multichannel bit-sliced crossbar switch that services $N \cdot K$ bit-streams. Since it can require up to $O(N^2)$ time to determine and configure the switch settings for an $N \times N$ crossbar, the use of K identically configured crossbars, means that an on-chip controller can service a throughput of $K \times N$ bits-per-second in $O(N^2)$ time. The chip can then be made to scale its data I/O bandwidth in proportion to its control computation bandwidth by appropriately scaling the size of the word (K) with the number of I/O channels (N). This is an example of a balanced "firehose" architecture as described in [10]. The controller determines the path for each input channel and configures all K crossbars to switch the bit-stream corresponding to that channel identically. Because of its reduced complexity, this controller can be incorporated onto the same chip. In addition, packet buffering (first-in, first-out buffers) and flow-control is incorporated into each channel. Thus, the data-buffers, switch-controller, and the switch can all reside on a single optoelectronic-VLSI chip with attached optical fiber to deliver the input and output bit-streams to the detectors and modulator, respectively. The amount of packet buffering per channel is a function of the bit-rate, the controller speed and the distance between the switch and the remote trnasceivers. For the purposes of this paper, the on-chip memory has been assumed to be very small, but may ultimately dominate the chip area.

The optical network is shown in Fig. 3 with each data bit-stream depicted as carried on a separate optical fiber. Each input node Pi has a K-bit-parallel transceiver with a linear array of K transmitters and K receivers (drawn as physically separate arrays) connected to linear arrays of optical fiber. All fibers are routed to two-dimensional 2-D arrays of $N \times K$ receivers and $N \times K$ transmitters at the central switch. The optical network implementation we discuss in Section IV differs from this schematic in that it uses wavelength-division multiplexing (WDM) to carry word-parallel data through single fibers.

III. The Amoeba Switch

The AMOEBA switch provides an array of parallel optical interconnects to a switch chip and eliminates the need for

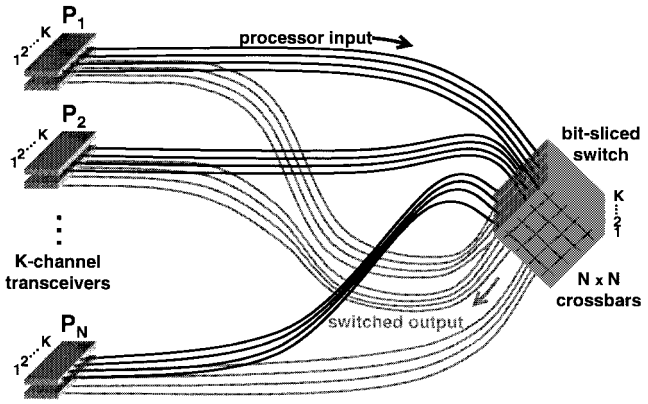


Fig. 3. Architecture of a tightly coupled WDM/SDM multiprocessor network based on the AMOEBA switch.

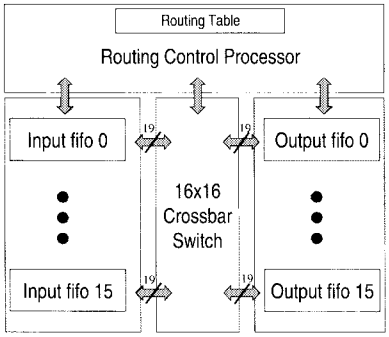


Fig. 4. Block diagram of the AMOEBA switch.

separate chips for physical-layer functions (e.g., encoding, clock-recovery, etc.), optoelectronic transceiver conversions, and packet buffering. This improves switch throughput and reduces package size and power dissipation compared to a fully electronic switch. The switch incorporates several physical and higher level features that directly reduce the latency of the link (compared to other switching technologies). For instance, by removing the need for serializer/deserializer circuits (such as those used in commercial optical links) the latency of an individual link is reduced. Packets do not have to be fully received and buffered before being routed to the output port. Packets of arbitrary length can be routed by the switch, which prevents the need to break a long packet into smaller packets and reassemble them at the other end. The routing protocol also provides the ability to do asynchronous teardown (without assistance from the controller), so that latency associated with link tear-down is minimized.

A. Switch Architecture

The AMOEBA switch implements an N-channel self-routing, packet-switched, asynchronous crossbar. Each channel can transmit a single word in parallel. Each word contains K data bits, one bit to signal end-of-packet, and two bits to synchronize data-flow through the switch and halt traffic when there is output port contention or congestion on either the receiving end or transmitting end of the switch. The addition of the special end-of-packet bit enables the AMOEBA switch to process an arbitrary number of words per packet. This bit is enabled in the last word (Byte) of the packet and it is used by the switch as a "cue" to tear down a connection.

The architecture of the AMOEBA chip was inspired by the IEEE 1355 (ISO/IEC 14 575) interconnect standard. The IEEE 1355 standard was developed to achieve efficient construction of fast, low-cost and low-latency interconnect for high performance systems. One of the distinguishing features of this standard is a simple link protocol that enables a complete line interface to be implemented with very few logic gates. Moreover, the link protocol uses wormhole routing to achieve low latency and to limit memory requirements at the line interface. Unlike typical network protocols, the IEEE P1355 can be used to build a single-chip switch that integrates the routing controller, lines interfaces and the crossbar (or multistage) switching network. A number of electronic implementations of IEEE P1355 switch chips have been demonstrated to date, such as the STC104, a 32-port routing switch from SGS-Thomson. Although these electronic implementations successfully implement a single-chip switch design, they are constrained to use serial connections to carry the data off-chip (due to the limited number of pins available with conventional electrical packaging). In using serial links, these electronic switch chips also incur the additional latency and power consumption penalty associated with multiplexing and demultiplexing the packet data since on-chip processing is typically done on parallel data streams. The AMOEBA architecture overcomes the bottlenecks presented by serial line interfaces. The inputs to the AMOEBA switch come in K-bit parallel format, are processed in that parallel format, and are output in parallel. To simplify system design and testing, we have chosen to implement a simplified version of the IEEE 1355 link protocol. Although this allows us to demonstrate the feasibility of our approach, it also limits the line speed to approximately 50 MHz (50 MHz $\cdot$ 16 bits $=$ 800 Mb/s/channel) in 0.8-μm CMOS technology (a simple firts-in/first-out (FIFO)-based circuit was used). The use of an IEEE 1355 compliant link protocol would allow a substantially higher line rate (250 MHz to 1 GHz).

B. Switch Prototype

We have built two prototypes of the AMOEBA switch in 0.8-μm CMOS. The first generation prototype chip consisted of 16-channels with 8 bits-per-channel and is discussed in [11]. Here, we will discuss the second generation AMOEBA switch with $N = K = 16$.

The overall schematic for the AMOEBA asynchronous, self-routing packet switch chip is shown in Fig. 4. It consists of the following components: 16 input FIFO's, 16 output FIFO's, 17 data crossbars, 2 control crossbars, routing table, and the routing controller. Each FIFO has 19 inputs and outputs: 16 I/O's for data signals; one I/O for the end-of-packet signal; a shift-in (SIN) control input and a shift-out (SOUT) control input for data transfer; an IREADY status indicator output and an OREADY status indicator output to establish handshaking. The addition of these two "handshaking" bits permits the AMOEBA to simultaneously and asynchronously transport packets with different data rates through the switch. Each packet is effectively synchronized on its own and independently of the other packets being carried by the switch. A detailed description of the design and testing of a FIFO

circuit with optical inputs and outputs is contained in [12]. The input FIFO's allow for buffering of data as packets enter the switch and, likewise, the output FIFO's allow for buffering data as packets leave the switch. In order to transmit the 16 data-streams per channel, as well as the end-of-packet line and control/handshaking signals between the input and output FIFO's in parallel, the AMOEBA switch contains nineteen (19) 16×16 crossbars. Each 16×16 crossbar consists of a tree of five 4-input multiplexors replicated for each output port, for a total of 80 four-input muxes per crossbar. During the layout of the crossbars, care was taken to minimize the timing skew between the different bits that are transmitted in parallel by the crossbars.

The routing table consists of two register files that store the current connection configuration for the AMOEBA switch. The routing controller is responsible for setting up and tearing down the switch connections. It takes the routing controller three clock cycles to establish each new connection. However; connection tear down occurs in parallel (e.g., any subset of the connections can be simultaneously terminated within a single clock cycle). These features allow self-routing packet based switching with variable length packets. Operating with 50-MHz clock, the routing controller can process up to 16 million connections per second. A more detailed account of the chip operation is presented in the Appendix.

Fig. 5(a) presents the microphotograph of the AMOEBA OE-VLSI chip. The 16-channel prototype AMOEBA switch chip integrates 200 000 MOS transistors and 1216 optical devices on a 24-mm^2 die fabricated in 0.8-μm CMOS. It implements a nonblocking, self-routing packet switch with up to 12.5-Gb/s aggregate throughput (or 800 Mb/s/channel) and is capable of switching up to 16 million packets per second. The routing controller on the AMOEBA switch prototype operates at 50 MHz. The line interface for each channel constitutes only 2K transistors, which includes and on-chip buffering of 16 Bytes per channel. The chip layout was a combination of custom design (for the receivers, transmitters, and connections to the crossbar) and automated layout (for the controller and routing of control signals to the crossbar). Most of the random logic (on the right hand side of the chip) and the wiring above and below the chip was for the controller and the distribution of control signals. This limited the clock speed of the controller to approximately 50 MHz. We expect that the area of the chip could be reduced and the speed substantially improved by a full custom design of the chip.

Fig. 5(b) shows all 1024 of the receiver diodes, in a 64×16 arrangement (corresponding to half the array) under forward bias, showing that all are operational. Device pitch is 35 μm in the vertical direction and 70 μm in the horizontal direction (140 m between two consecutive receiver columns in the figure). Similar forward-bias tests verified that all the diodes corresponding to modulators were likewise functional. Each MQW diode was approximately 20 μm $\times$ 50 μm [1].

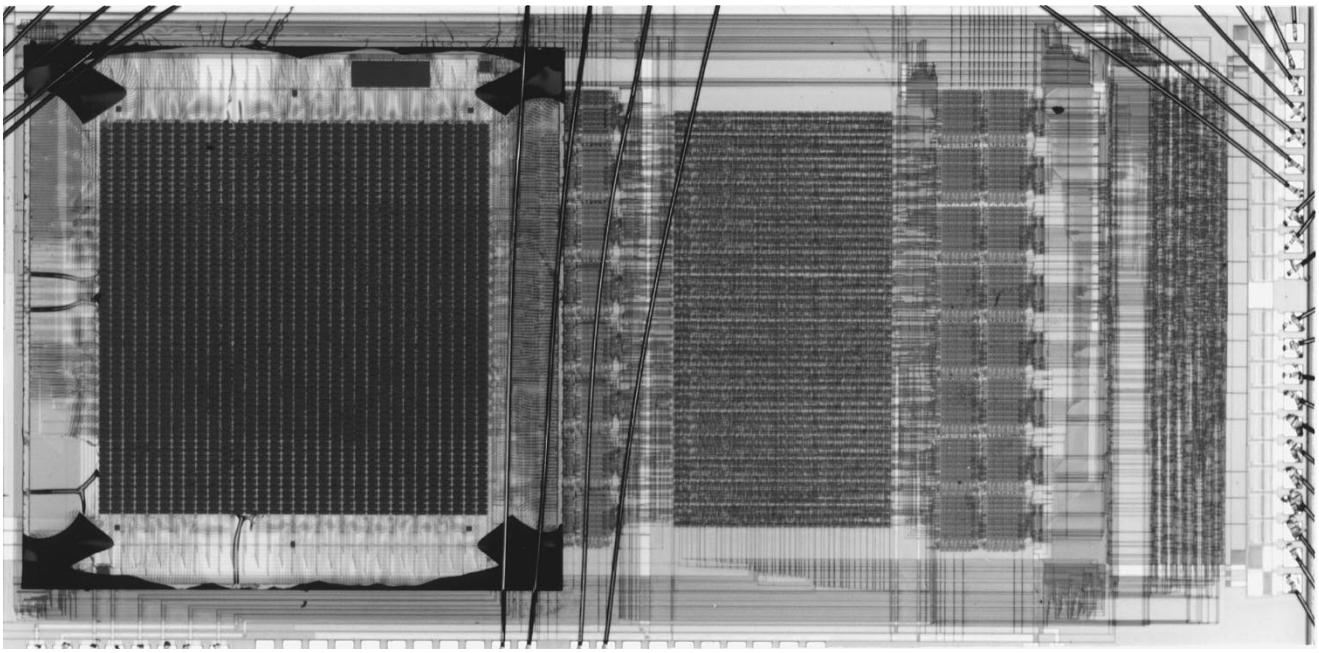
Each p-i(MQW)-n diode required two coplanar contacts for connection to the CMOS chip. The AMOEBA chip used multiple pairs of bias-voltage lines (one pair for every eight diodes) to provide varying bias across the modulator diodes in a single column, corresponding to the different wavelengths (bits) of each channel (word). As described in Section VI, this allowed the contrast of the MQW modulators to be improved for a wide range of wavelengths. Differential signaling was employed in order to increase tolerance to power variations and contrast-ratio variations across the array. A single pair of high-voltage lines ($\pm$5 to 8 V) was used to reverse-bias the detector diodes used by the first generation AMOEBA interconnect. Although only 1216 devices were needed for this switch, photonic chips were produced with all 64×32 devices functional, to demonstrate extensibility of the AMOEBA chip architecture to a larger number of channels and larger word sizes. This chip incorporated three-stage two-beam transimpedance receivers and transmitters with a maximum bandwidth of 1 Gb/s in 0.8-μm CMOS [13].

Initial electrical and optical testing have shown that data can be switched through the chip at 50 Mb/s-per-optical channel. At this bit-rate, the switch was able to provide error-free communication when tested overnight. Fig. 6(a) shows electrical output data from the first bit-stream of output channel 1 that originated from the first bit-stream of input channel 2. Fig. 6(b) shows the corresponding optical output data at 70 Mb/s flowing through the AMOEBA switch prototype. The chip was limited to this data rate due to the FIFO synchronization circuits used for the system. This speed can be increased by reducing feature size and by paying careful consideration to minimizing interconnect lengths. In contrast, previous switch designs have largely ignored the issue of synchronization, leaving it to be performed outside the switch, by the sending/receiving nodes. The data rate in a system that does not provide synchronization is only limited by the optical transceiver circuits, which have been demonstrated to over 1.5 Gb/s (individual) and over 800 Mb/s in system experiments [5].

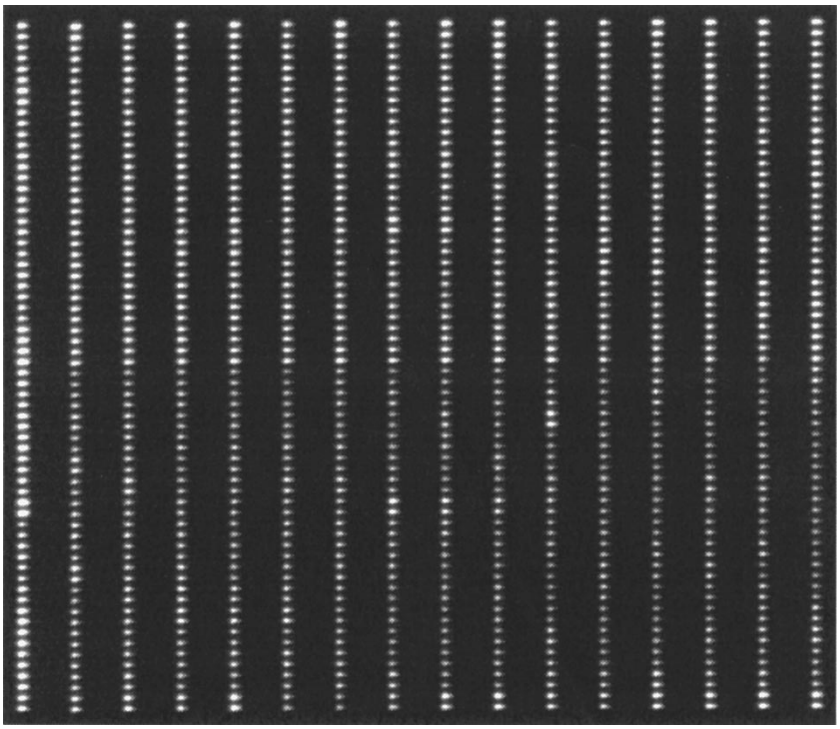
IV. Design of the Fiber/Free-Space Network for the Amoeba Switch

The optical network shown in Fig. 3 can be directly implemented using linear and 2-D arrays of fiber. We have a choice between multimode fiber (MMF) and single-mode fiber (SMF). SMF offers an effectively unlimited modulation bandwidth, but requires submicrometer lateral alignment tolerances. MMF relaxes the lateral alignment by about a factor of 10, but modal dispersion then limits the product of data rate times the transmission distance. The ultimate goal of this network is to provide parallel gigahertz-clock rate connections spanning a several kilometer wide campus. The bandwidth-distance product available from standard graded-index MMF's with overfilled-launch of 850-nm light is roughly 160 MHz·km to 300 MHz·km for 62.5- and 50-μm-core diameters, respectively. MMF would therefore restrict the maximum transmission distance for a 1-Gb/s data rate to less than 500 m, or the maximum data rate for a 5-km-diameter campus to some 200 Mb/s. SMF may, therefore, become necessary.

The optical system requirements for this application are to connect a 16×38 array of modulators at the AMOEBA switch chip to 1×38 arrays of detectors at each of 16 transceivers located at distances from a few meters to several kilometers,



(a)

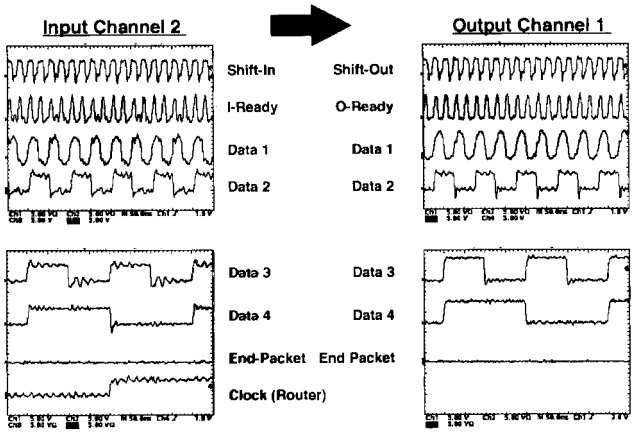


(b)

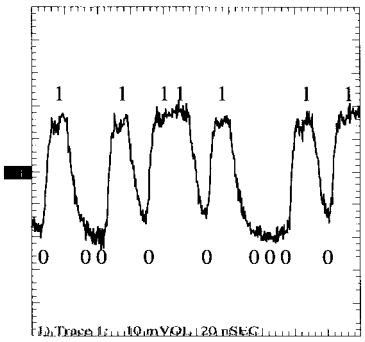
Fig. 5. (a) Microphotograph of the 16-channel AMOEBA switch (~200 K transistors) fabricated in 0.8-μm CMOS with a 64×32 array of MQE diodes flip-chip bonded to it. Chip dimensions are 4.5 mm $\times$ 8 mm. (b) Forward biased array of 1024 receivers arranged in 64×16 array showing that all the receiver diodes were functional. The entire diode array contained another array of 64×16 modulators interleaved with the modulators (not shown).

and similarly connect 1×38 arrays of modulators at the transceivers to a 16×38 array of detectors interleaved with the modulators on the switch chip, all at data rates scaling up to 1 Gb/s. The usable device aperture is 18×18 μm. The pitch between devices is at least 35 μm (constrained by fabrication tolerances) and no more than 125 μm (constrained by the CMOS chip area).

One possible implementation for the free-space-to-fiber interconnect interface is 2-D fiber arrays for the input and output connections [14]. The problem comes from alignment tolerances. To couple 90% of the signal from a perfectly mode-matched source to 850-nm SMF requires a decentration of less than 1.4 μm. This tolerance is achievable with active alignment of a single fiber, and is also possible to obtain using a 1-D array of fibers (assembled using silicon V-grooves for alignment). But there is currently no practical way to manufacture 2-D SMF arrays with single-micrometer positioning accuracy.



(a)



(b)

Fig. 6. (a) Measured electrical data from packets flowing through switch from input channel 2 to output channel 1 demonstrating switch operation. Shown are 4 of the 16 data channels at 50 MHz per bit-stream. Shift-in, shift-out, end-of-packet, and the clock to the switch controller (router). (b) Measured optical data from input channel 16 flowing through switch to output channel 16 on the data1 channel.

The combination of wavelength multiplexing with 1-D fiber arrays shown in Fig. 7, however, can achieve 2-D array interconnection. An input array of fibers is imaged through a planar diffraction grating onto 2-D array of devices, so that each row of devices is associated with a corresponding optical wavelength. In the simple optical system drawn in Fig. 7, each MQW modulator is optically powered from the same optical fiber which will carry the modulated output data signal. The back-reflected data signal is separated from the input fiber using a 2×2 splitter (or, more efficiently, with an optical circulator). In this configuration, the lateral alignment tolerance of the fiber to the modulator is determined not by the fiber mode field diameter, but by the useable aperture of the modulator, as illustrated in Fig. 8.

For the crossbar switch, a 1×32 array of fibers would hold 16 input data fibers interleaved with 16 fibers carrying modulator power and the back-reflected (and modulated) data output signals. The fiber array uses a 70-μm pitch, exactly matching the lateral device spacing, because the optical system produces 1-to-1 imaging for each wavelength. A prototype fiber array was fabricated by drawing a conventional SMF preform to half the normal 125-μm diameter, which reduced the core diameter from 8.2 μm to the 4.1-μm size necessary for single-mode guiding at 850 nm. The fibers were packaged in

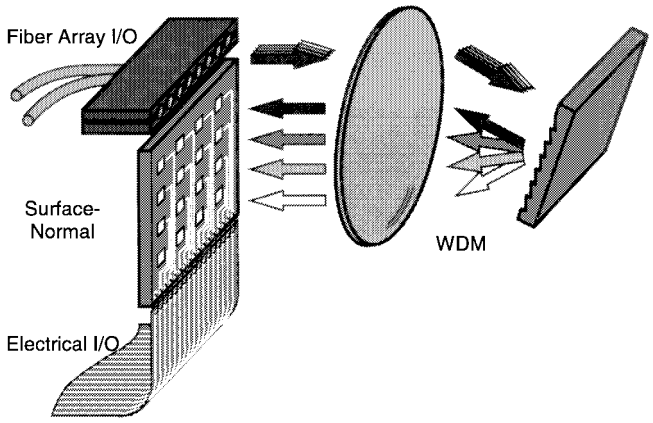
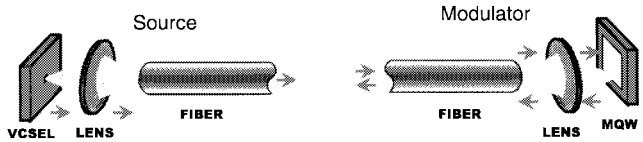


Fig. 7. The combination of space and wavelength division multiplexing enables a 1-D fiber array to access a 2-D array of photonic devices on the optoelectronic switch surface. (Note that the single-fiber transceiver uses a different optical system.)



(a) (b)

Fig. 8. When coupling an optical source to a fiber (a), the lateral alignment tolerance is determined by the propagating fiber mode. However, when light is coupled out of the fiber and reflected from an oversized modulator back into the source fiber, the lateral alignment tolerance is determined by the modulator aperture.

a silicon V-groove array using standard silicon optical bench techniques.

The use of WDM with modulators introduces two immediate requirements: a multiple-wavelength source, and broad-spectrum MQW modulators. One source of broad spectrum light in the 850-nm range is an mode-locked Ti:Sapphire laser, whose 100-fs pulses have some 30-nm spectral width, which, in theory, is sufficient for 100 or more wavelength channels. While GaAs–AlGaAs MQW modulators are normally considered narrow-bandwidth devices, they can maintain moderate contrast operation over the width of the exiton peak. A 16-channel dense WDM data transmitter using a short-pulse laser as a broad spectrum source, free-space optics to distribute the wavelengths, and a surface-normal modulators to apply the data signals was previously demonstrated with FET-SEED modulators and a mode-locked Ti:Sapphire laser [15].

A Ti : sapphire laser with 1-GHz repetition rates, 110-fs pulses, and 450-mW output power has been demonstrated [16]. This type of Argon-pumped laser is expensive and bulky, requiring high drive currents and water cooling. However, compact and inexpensive lasers have recently been demonstrated using diode pumped Cr : LiSAF with saturable Bragg reflector mode locking [17]. Such lasers can already provide over 100 mW of 850-nm output with 100-fs pulses at 300 MHz, and further development of this type of laser should yield a reliable 1-GHz rep-rate source compatible with computer environments. In fact, a diode-pumped Cr : LiSAF contained within a total laser footprint of 28×22 cm producing 133-fs pulses at 450 MHz has recently been reported [18]. A more

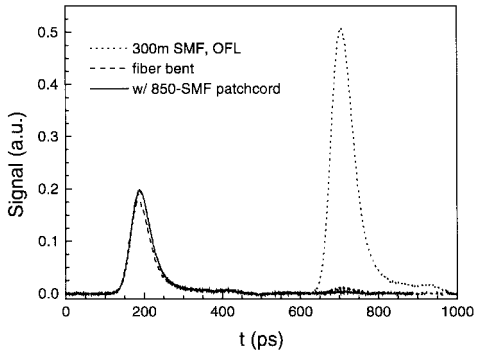


Fig. 9. Temporal response of 300 m of standard 1.3-μm SMF with 40-ps pulse input from an 850-μm laser using a 9-GHz detector. Under over-filled launch (OFL) conditions, the mode delays from the two propagating modes in the fiber, LP01 and LP11, can be seen. This differential mode delay reduces the effective bandwidth of the fiber. When the fiber is looped near the receiver or when an 850-nm SMF patch cord is used after the fiber (before the receiver), then the higher order mode (LP11) is severely attenuated, and the bandwidth is determined by the impulse response of the fundamental mode (LP01).

conventional approach to the laser source would be to use a set of discrete wavelength-stabilized lasers combined using a arrayed waveguide router. There is no fundamental difference between this and the 1.5-μm and multiwavelength lasers (MFL) demonstrated for telecommunications applications [19], although a significant development effort would be necessary to bring 850-nm MFL's to market.

When using low-cost SMF, not-optimized for 850-nm transmission, the issues of skew and, more importantly, modal dispersion and chromatic dispersion become relevant, particularly as the data rates and the number of WDM channels are increased. First, it is interesting to look at the requirements on the SMF carrying the 850-nm signal. The core diameter necessary to support only the TEM_{00} mode scales linearly with wavelength (approx. 5 μm for 850 nm). Standard SMF used in campus networks is designed for 1.3-μm wavelength light, becoming "MMF" at 850 nm, with two distinct mode groups: LP_{01} and LP_{11}. Just as in gradient index MMF, the higher order modes have a significantly different effective path, resulting in modal dispersion. However, it has long been known that it is possible to transmit high bandwidth data through 1300-nm SMF with 800-nm light with a short section of 800-nm SMF, or simply by twisting the fiber into a coil, to strip off the higher order modes just before entering the receiver [20]. Provided that alignment tolerances suitable for 850-nm SMF are maintained at the transceivers, this means that standard low-cost telecom SMF (optimized for 1300-nm transmission) could be used in this network.

To determine the bandwidth that such a system could support, we have performed differential-mode-delay (DMD) measurements that involve exciting the fiber with a short pulse and examining the temporal response of the fiber to determine the delay between the different mode-groups. Fig. 9 shows the temporal response of standard 1300-nm SMF (9-μm core) with 40-ps pulse input from an 850-μm Ti : Sapphire laser after propagating a distance of 300 m. For these measurements, a large-area detector and a receiver with a bandwidth of 9 GHz was used. Under over-filled launch (OFL) conditions, the mode delays from the two propagating

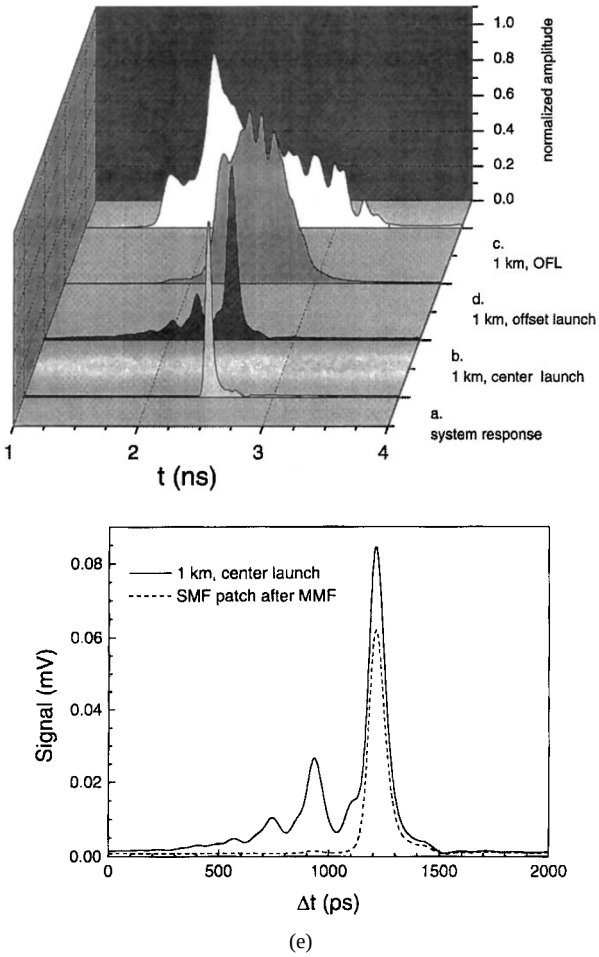


Fig. 10. Normalized temporal response of 1 km of standard 1.3-μm (9/125) SMF using a short stretch of 850 nm (5/125) SMF to launch input. (a) System response (without 130 nm fiber). (b) Response when 850-nm fiber is centered on the 1300 nm fiber. (c) Response under overfilled launch when laser is directly coupled into 1300 nm fiber. (d) Response when 850-nm fiber is offset (misaligned) with respect to the 1300 nm to launch higher order modes. (e) Response under center launch as in (b) with and without another 850-nm SMF patch-cord after the standard fiber. The 850-nm patch-cord acts as a mode-stripper which removes the higher order mode.

modes in the fiber, LP_{01} and LP_{03}, can be seen. This differential mode delay reduces the effective bandwidth of the fiber since the duration of the output pulse is effectively 1.25 ns. When the fiber is looped near the receiver or when a short stretch of 850-nm SMF (with a 5-μm-core diameter) is used just before the receiver, then the higher order mode (LP_{11}) is severely attenuated, and the effective bandwidth is determined by the impulse response of the LP_{01} mode (approximately 200 ps limited by the detector bandwidth).

We also examined the DMD behavior under various launch conditions by using the 850-nm single-mode patch cord to launch into a 1-km stretch of 1300-nm SMF. Fig. 10(a) depicts the system response (without 1300-nm fiber) showing the minimum resolvable mode separation. Fig. 10(b) shows the response when the 850-nm fiber is centered on the 1300-nm fiber. This result suggests that power can be effectively coupled into the fundamental mode and that there is little mode-mixing between the two mode groups, i.e.; light coupled to the fundamental mode tends to stay there despite relatively long paths and multiple (single-mode) connectors. Fig. 10(c)

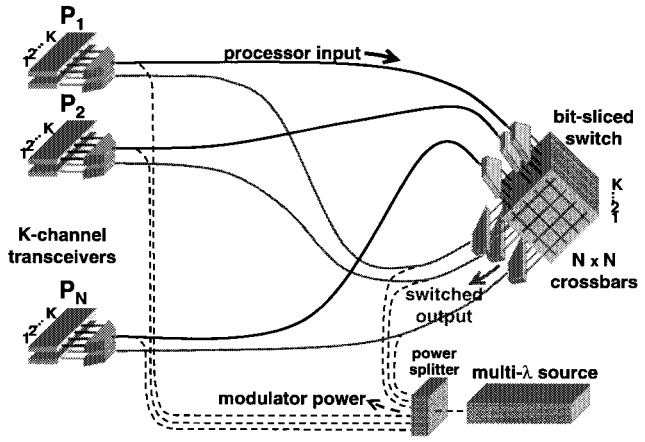


Fig. 11. Fiber-optic WDM/SDM network based on modulators. A short-pulse (100 fs) laser can be used as the broad-band multiwavelength source.

shows the response under overfilled launch (both mode groups excited), when laser is directly coupled into 1300-nm fiber. Fig. 10(d) is the response when the higher order modes are selectively launched by decentering the 850-nm fiber with respect to the 1300 nm (offset launch). Note that the higher order mode group (LP_{11}) experiences more dispersion than the fundamental mode. Fig. 10(e) shows the response of the 1300 nm fiber under center launch as in (b) with and without another 850-nm SMF patch-cord placed before the receiver. The 850-nm patch-cord again removes the higher order mode, confirming that large bandwidths can be achieved even after 1 km of propagation in standard 1300-nm SMF by using center-launch and stripping the LP_{11} mode group before the receiver.

If 850-nm optimized SMF were used in the network, thereby mitigating the modal dispersion issue, the effective bit-rate would be limited by chromatic dispersion. The chromatic dispersion of standard silica 9/125 1300-nm SMF is approximately 100 ps/nm/km at 850 nm. Assuming a source spectrum of 0.4 nm [full-width at half-maximum (FWHM)], this would limit the bit-rate-distance product to approximately 6 Gb·km, corresponding to a minimum bit-period of approximately 170 ps.

A third issue is skew. There are two sources of skew: electrical and optical. Note that the bits of one channel are synchronous (with the clock embedded as one of the transmitted WDM channels) whereas there is no interchannel synchronization requirement. It is worth mentioning that the use of modulators with a pulsed optical laser such as the one described in this paper may help to remove the skew associated with the electrical on-chip signals because the signals are essentially retimed by the sampling optical pulse. In terms of the optical skew, assuming that the clock signal were embedded at the center wavelength of the usable spectrum of the MQW modulators, and that the center-to-edge separation was on the order of 10 nm (corresponding to a total of 40 WDM channels at 0.5-nm spacing), then the skew between the bits would be a negligible fraction (less than 10 ps) of the minimum bit-period for 1-km transmission.

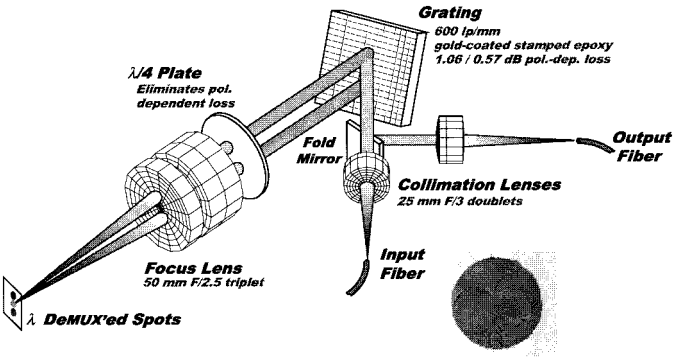
Fig. 11 shows an overview of the resulting WDM data network (although not the actual physical layout of the optical components). $2N$ SMF's connected to a central switch break out into N duplex fiber pairs routed to each node. At each fiber terminal, free-space optics distribute individual wavelength signals onto MQW modulators and detectors to carry word-parallel data. Light from a single short pulse (or multifrequency) laser is passively split and distributed to power the entire network's modulators, with one fiber to each of the N transceivers and N channels to the switch. This source would provide automatic clock and wavelength synchronization, and could potentially be housed in a single box with AMOEBA switch.

V. Dense-WDM Fiber-to-Free-Space Transceiver Package

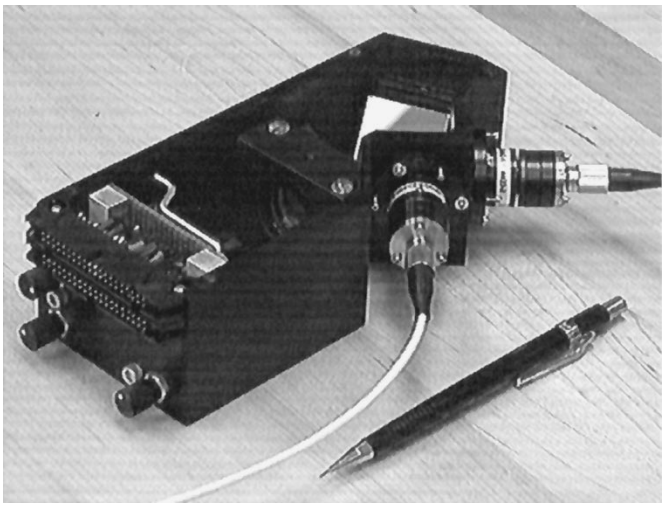
The optical system shown in Fig. 7 is an effective way to provide interconnection to large 2-D arrays of devices. In particular, the relatively loose lateral and axial alignment tolerances arising from the design's optical symmetry make it practical to interconnect large device arrays with 850-nm SMF. But this free-space optical design does not separate the backreflected output data from the modulator power. Instead, it requires external fiber-optic 2×2 splitters (which are lossy) or optical circulators (which are expensive). For the transceiver nodes, however, it is straightforward to modify the optics to direct the backreflected data signal into a distinct output fiber. In particular, the modulator illumination can be angled so that the backreflected (and modulated) signal does not overlap the input illumination, and can be redirected to a separate output [21].

The resulting optical system, used for the AMOEBA transceivers, is shown in Fig. 12(a). Light from an input fiber is collimated by a $f = 25$ mm lens and illuminates a 1200 lp/mm blazed grating in the Fourier plane. The diffracted light is focused by a $f = 50$ mm triplet lens onto the device plane so that a 5-nm spectral band creates a 300-μm stripe on the device. Light reflected from the device is recollimated by the focus lens, and recombined by a second pass through the grating. The axis of the focus lens is slightly lower than the collimator, so that the reflected signal is displaced from the input. This allows a fold mirror placed just below the collimated input to pick off the reflected output and direct it into a separate fiber collimator. From this point on, the output signal does not retrace the input light path, so the fold mirror and output coupler must held in precise alignment. This optical system has no intrinsic loss. In practice, surface reflections, grating and optical aberrations, and misalignment are all possible sources of loss.

The optical system diagram shows a single input source. However, a second fiber placed adjacent to the first will also be dispersed by wavelength and imaged onto the device plane. For the AMOEBA transceiver, two input fibers are needed. The first fiber carries data to a column of detectors. The second fiber carries optical power to an adjacent column of modulators, and is aligned so that the reflected light will be coupled into a single output fiber. The required 1×2 fiber "array" is easily handled using a standard optical fiber connector body modified by doubling the diameter of the ferrule hole so that two fibers can be inserted and polished.



(a)



(b)

Fig. 12. (a) The optical system layout for the dense-WDM transceiver consists of a 4-f imaging system with the grating placed in the Fourier plane and the OE-VLSI chip placed in the image plane. Pupil-division multiplexing together with a fold mirror is used to separate the modulated wavelength bit-streams into the output fiber. (b) Packaged WDM transceiver.

The photograph inset into the drawing shows a pair of SMF's packaged side-by-side in a fiber-optic connector with an oversize (250 μm) diameter hole. The only additional degree of freedom required is to be able to rotate the ferrule to that the two columns of spectrally dispersed light are separated by the right spacing at the device plane.

Fig. 12(b) shows the transceiver's optomechanical package, including electrical connections for device control and two FC-connectorized optical fiber ports. One port accepts a dual-fiber connector (one fiber for input data, the other for modulator power) while the other port accepts the output data fiber. There are also numerous tip/tilt/translation controls for optical alignment. This implementation used off-the-shelf lenses intended for on-axis, small field-of-view operation. The total fiber-to-fiber insertion loss for this package, measured using a gold mirror at the device plane, is 10.3 dB, with 0.7-dB polarization dependent loss. The chip is held in a 84-pin ceramic package in a zero insertion force (ZIF) socket. This package is intended for laboratory environments. However, once the mechanics are aligned and locked down, the package can be handled and fibers can be attached and removed with moderate ($\sim$1 dB) change in insertion loss. With minor modifications, such a package might be engineering to mount directly onto a circuit board inside the workstation. Optical losses due to the diffraction grating are difficult to reduce, but it is reasonable to expect that the optical losses could be cut to under 4 dB using custom lenses. Fig. 13 shows the resolution of the dense-WDM transceiver package measured with an optical

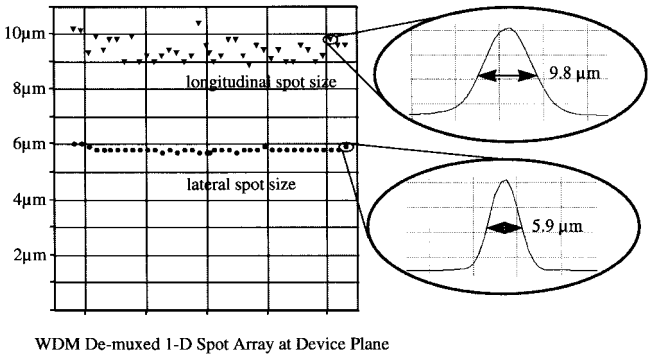
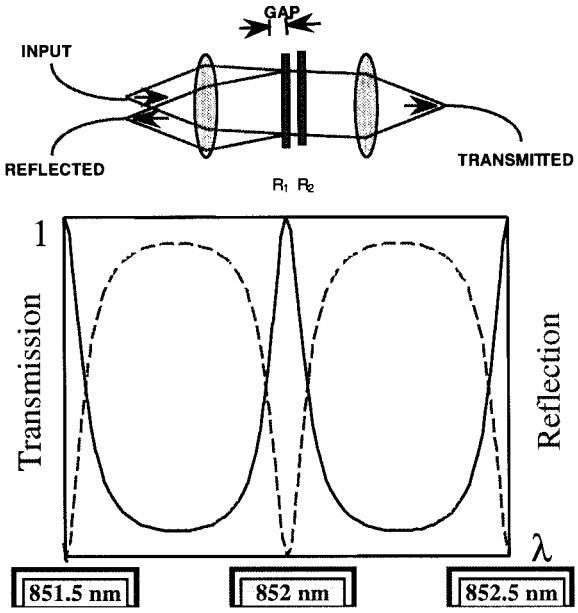
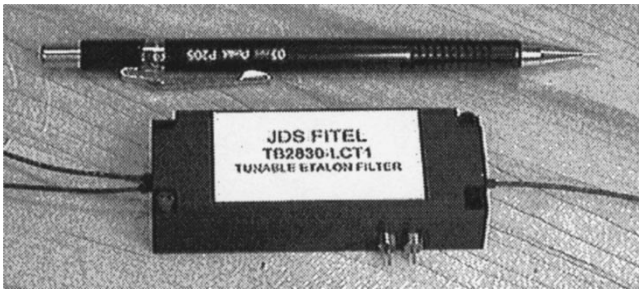


Fig. 13. Performance of the dense-WDM transceiver package. The spot size is relatively independent of wavelength over 38 channels spaced at 0.485 nm (35 μm at the chip image plane). The spot position, in addition, is distortion free over a 20 nm range between 843 and 862 nm. Total insertion loss was approximately 10 dB with SMF, and approximately 8 dB with MMF.

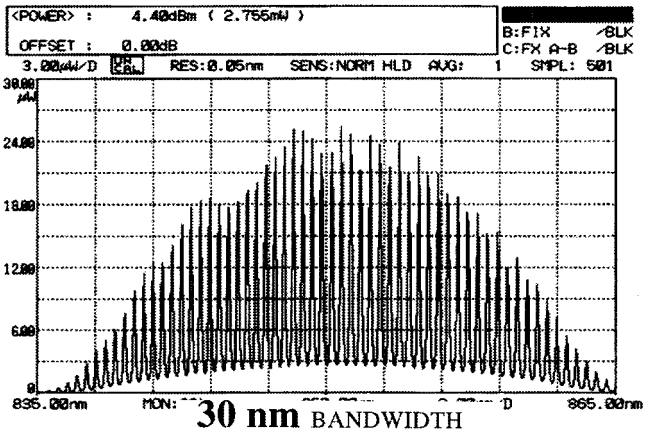


(a)

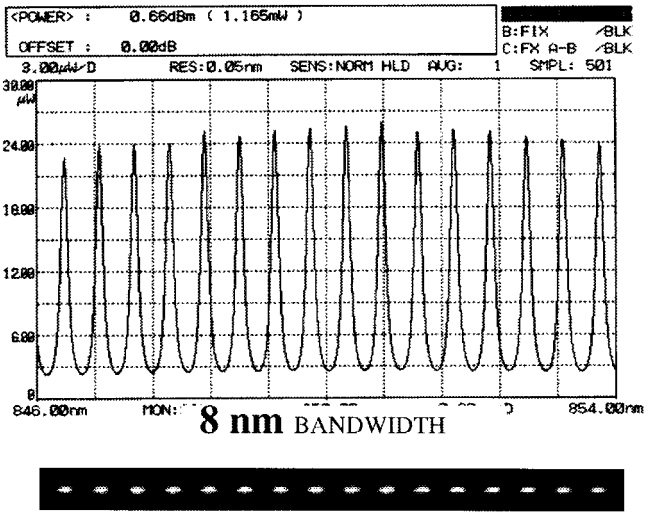


(b)

Fig. 14. (a) A tunable etalon filter can be used in this system to provide a narrow, periodic passband for the desired transmission wavelengths. This prevents unwanted light from being imaged onto the chip, and saves power, because the reflected light can again be reused for another transceiver. (b) Picture of the filter fabricated as a custom product by JDS Fitel.



(a)



(b)

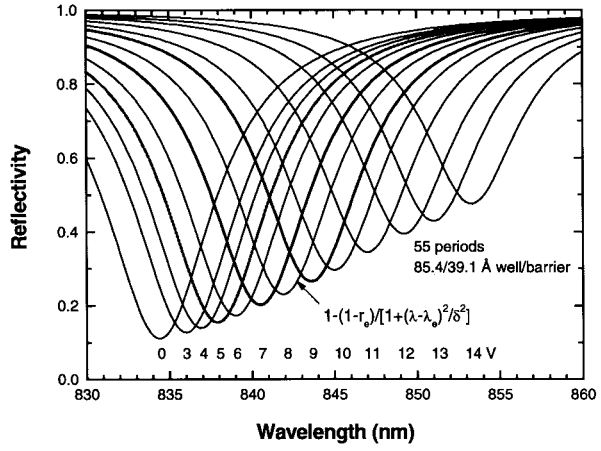
Fig. 15. Transmitted output from the filter: (a) Showing over 30 nm of bandwidth obtained from a 100-fs pulsed laser source. (b) Magnified view of central 16 channels in an 8-nm bandwidth, and the resulting spots dispersed by the grating onto a CCD camera.

spot profilometer at the device plane. Multiple wavelength input produces a distortion-free linear array of uniform 6×10 micron (FWHM) spots.

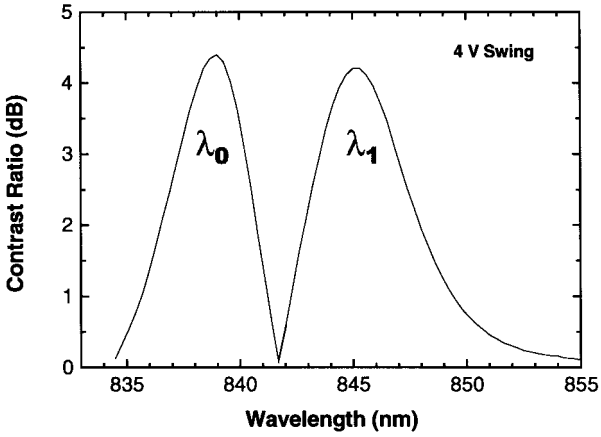
Femtosecond sources produce a continuous spectrum, but the modulators and hence the data channels are discrete. An external filter can remove the unwanted light between data channels. Fig. 14 shows a tunable etalon filter designed to provide a narrow periodic passband at a wavelength spacing of 0.485 nm, corresponding to approximately 35 nm at the device plane. By discretizing the passband, the filter prevents unwanted light from being imaged onto the chip surface between the modulators, reduces chromatic dispersion, and also saves power by allowing the reflected power to be reused for other channels at the same wavelength spacing. The measured output from the filter can be seen in Fig. 15, using a conventional (Coherent Mira-900) mode-locked Ti : sapphire laser source.

VI. Dense-WDM Transmission with MQW OE-VLSI Modulators

Compatibility of the MQW OE-VLSI devices with dense-wavelength division multiplexing is a key concern. Fig. 16



(a)



(b)

Fig. 16. (a) Reflectivity spectra as a function of reverse bias voltage for an MQW modulator with 55 periods and a 85.4/39.1 well-to-barrier ratio. (b) Predicted contrast ratio of a linear array of MQW modulators as a function of wavelength, for a fixed bias and a fixed voltage swing of 4 V. The modulator can be operated either at a exciton peak (λ-0 operation) or at a wavelength beyond the exciton peak (λ-1 operation).

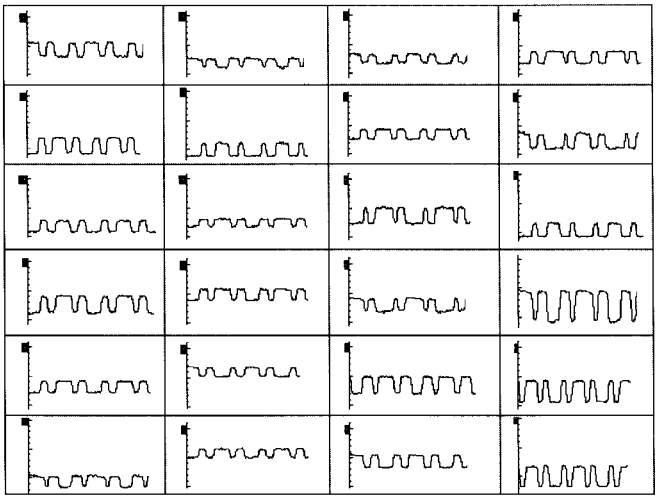


Fig. 17. Optical outputs of a linear column of modulators at 155 Mb/s per bit-stream, configured as a 24-bit dense-WDM transmitter.

shows measurements from fabricated MQW modulators (not bonded to the VLSI chip) which reveal a normal operating wavelength range of approximately 4 nm, at a fixed bias

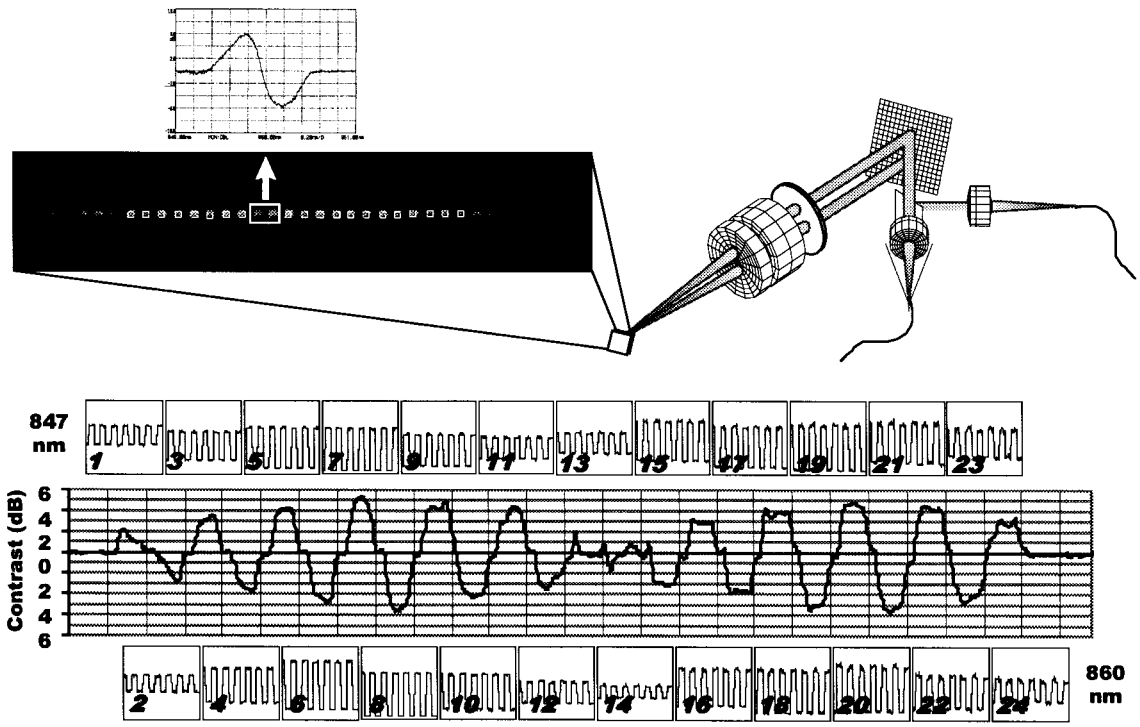


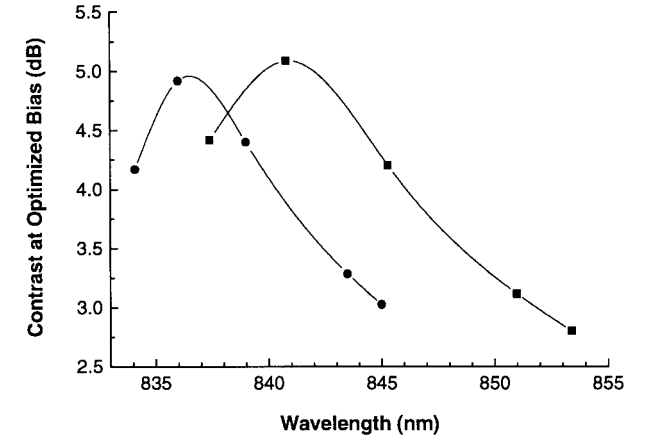
Fig. 18. Experimental results from the 24-channel differential dense-WDM transmitter array when operated in the transmitter package with a fixed bias voltage and a fixed swing of 3 V. Shown are the 24 individual waveforms and the output of an optical spectrum analyzer configured to measure contrast ratio as a function of wavelength.

voltage (based on a minimum contrast ratio of 2:1). Changing the bias voltage causes the exciton to shift and broaden. We have adopted two techniques to increase the available spectral bandwidth of the MQW modulators (typically only 5 nm wide). First, we use both λ-0 (on the exciton absorption peak) and λ-1 (off-peak) operation simultaneously in different parts of the array. Second, we use several adjustable-bias levels along the array to stagger the voltage levels across the modulator array in order to remove the low-contrast valleys that occur between λ-0 and λ-1 peaks. In order to test the feasibility of these techniques, a separate transceiver chip was designed and implemented with exactly the same diode-spacing as the AMOEBA chip for the specific purpose of testing as many WDM channels as possible. This transceiver chip was built for the optomechanical package (to be situated at the processor nodes) and designed to communicate with the AMOEBA switch. Multiple bias-voltage lines (one per eight diodes) were supplied to this chip. Fig. 17 shows the testing of each modulator in a 24-bit transmitter chip at 155 Mb/s per channel, where we used a monochromatic laser source probing the OE-VLSI device prior to insertion in the optical package.

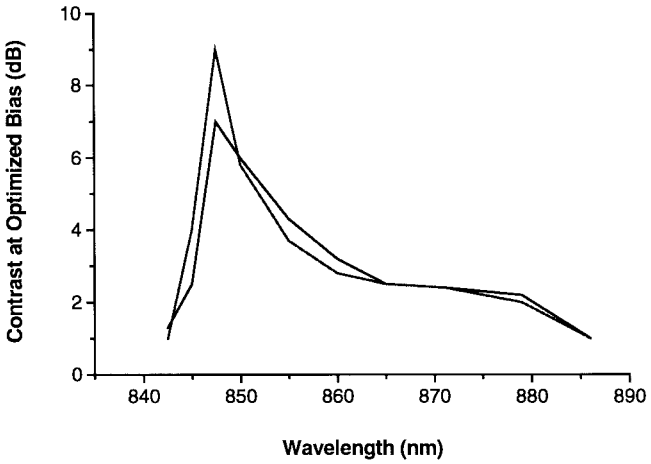
Measurements were then taken with the transmitter chip in the WDM transceiver package to determine the maximum usable wavelength range under different biasing conditions. First, all 24 channels were simultaneously modulated with a fixed voltage swing and a fixed bias voltage. The results are presented in Fig. 18, which shows the 24 simultaneous bit-streams. Also displayed in Fig. 18 is the contrast ratio measured for each modulator. This was accomplished using an optical spectrum analyzer (OSA) with the modulation speed set to the sampling interval of the OSA (approximately 0.5 Hz), and the OSA configured to read the difference of two successive traces. Because the modulators were configured for differential transmission, adjacent modulators have opposite polarity. As is evident from the Fig. 18, we can see two regions of high contrast separated by a region of low contrast, the same behavior predicted in Fig. 16(b). Because the MQW modulators used in the experiment had a larger number of periods than that of Fig. 16, the measured values of contrast ratio are higher and the exciton peak is shifted toward the red.

It is possible to span a much larger operating range using several independent bias voltage lines for a linear array of modulators. The predicted contrast of the modulator as a function of wavelength at the optimum bias is shown in Fig. 19(a), where the two curves correspond to λ-0 and λ-1 operation. The peak contrast would simply be the maximum of the two curves at any given wavelength. This was verified by experiments made with the dense-WDM transmitter package. The results are displayed in Fig. 19(b). By separately tuning the adjustable bias voltage lines across the transmitter array to maximize contrast ratio across the array, a 2:1 contrast ratio was achieved over as much as a 30–35-nm bandwidth. This would potentially allow up to 60 wavelengths to be transmitted in parallel.

This estimate assumes that either the temperature of the chip could be stabilized or that modulators with reduced temperature sensitivity could be fabricated. Both of these options may be feasible. Indeed, for the purposes of this paper, no special attention was given to temperature stabilization for the MQW modulators. In [22], the authors show that a $\pm 30^{\circ}$ temperature variation as well as a ± 1 nm in laser variation can, in principle, be tolerated by using voltage feedback in the



(a)



(b)

Fig. 19. (a) Predicted performance of a linear MQW modulator array as a function of wavelength, when the voltage swing is constant (4 V), but the bias voltage can be varied across the array. (b) Experimental results from the packaged dense WDM transmitter, verifying that an acceptable contrast ratio (e.g., 2 dB) can be obtained over a wide wavelength range, when multiple bias lines are provided to the array.

modulator and by paying special attention to MQW design. This temperature range corresponds to a wavelength range of approx. 17 nm (for λ-1 operation). This theory (as well as the experimental verification presented in our paper) suggests that a wide wavelength operating range can be achieved if one is willing to actively stabilize the chip operating temperature and bias the modulators appropriately.

VII. Summary and Conclusion

We have described the AMOEBA switch: a novel implementation of a digital, asynchronous, optoelectronic bit-sliced, crossbar array switch. It is the first switch that integrates a switching fabric, optical I/O devices, packet buffers, and routing controller on a single silicon CMOS die. This device can function as a high-performance single-chip central switch on a data network serving multiple processors or high-performance workstations. The switch features a routing algorithm designed to route packets of any length with end-to-end flow control. The key enabling technologies for this AMOEBA system are: 1) an optoelectronic device technology that permits the integration of several thousand optical inputs and outputs onto a state-of-the-art submicrometer silicon CMOS integrated circuit; 2) an optical WDM networking technology that uses SMF to transport dense WDM data between nodes and that uses either multifrequency lasers or a short-pulse laser functioning as a broad-spectrum power source; and 3) a 1-D-fiber to 2-D-free-space interconnect/packaging design that can convert between dense WDM data on a SMF and space–division-multiplexed data in free space.

Although the complete system has not yet been assembled, we have demonstrated key components of the system, including a 16-channel, nonblocking switch chip, the packaged transceiver, and the principle of dense-WDM transmission with flip-chip bonded MQW modulators. Even at a modest clock speed of 50 MHz, the controller enables the switch chip to process packets at a peak rate of 16 Mega-connections/s. An array of 64×32 GaAs–AlGaAs MQW devices was bonded to the AMOEBA switch. Testing of the chip showed that electrical and optical data can be sent through the switch at a line-rate of 70 Mb/s. This provides a peak user throughput of approximately 18 Gb/s (at a peak chip optical I/O bandwidth of 36 Gb/s). Based on the demonstrated performance of the hybrid optoelectronic-VLSI technology, we expect that the number of data channels as well as the bandwidth per channel can be scaled up to provide terabit/s throughputs.

We have also designed the fiber network that will allow the AMOEBA chip to be networked with the multiple processing nodes. A combination of wavelength-and-space-division multiplexing with SMF's is used for data transport. The use of 850-SMF or standard 1300-nm SMF's (with mode-stripping) will allow much higher bandwidth-distance products than standard MMF solutions. Each processor encodes its data onto a SMF using a dense-WDM technique; one fiber from each processor is aggregated into a 1-D fiber ribbon. We have implemented a optomechanical transceiver package that accomplishes this SMF-to-free-space WDM/SDM interface.

We have adopted two techniques to increase the available spectral bandwidth of a linear array of modulators. We use both λ-0 (on the exciton peak) and λ-1 (off-peak) operation, and we provide multiple bias lines across the array to tune the modulators for best performance. These techniques have been verified by modeling with measured data and through transmission experiments. We have demonstrated 24–channel dense WDM transmission over SMF across a 15-nm window around 850 nm with 3-dB contrast and across a 35-nm window with over 2-dB contrast. The latter is sufficient for digital transmission with differential receivers.

Appendix
Switch Operation

To illustrate the operation of the AMOEBA switch, we consider an example scenario where input port one (#1) requires connection to output port three (#3). Assuming that the input node one is not busy, the packet header is shifted into Input_FIFO(1). The IREADY(1) will indicate that input Input_FIFO(1) is ready to accept new incoming data. Only when IREADY(1) is high, the SIN(1) signal is applied and

incoming data is accepted by Input_FIFO(1). The destination address for this packet is contained in the packet header at the beginning of the packet. It is examined by the routing controller and checked against the register files that store the current connection configuration. If the register files indicate that output port three (#3) is busy on another (unrelated) connection, then the new connection cannot be established at this time. In this case, the input_FIFO(1) continues to receive incoming packet data until the FIFO memory is full (because data does not leave the FIFO). When the FIFO memory becomes full, the IREADY(1) signal is deactivated, thus indicating to the originator of the packet that Input_FIFO(1) is busy and unable to accept any more incoming data.

The routing controller is the "brain" of the AMOEBA switch chip. Due to its logic complexity, the routing controller was designed using behavioral VHDL (as opposed to manual logic design), synthesized with logic synthesis tools, and placed and routed using automated placement and routing tools (as opposed to manual layout). The final result of this is a "sea of standard cells" containing over 20 000 MOS transistors.

The simplified operation of the routing controller is as follows: the controller logic continuously polls the 16 input ports (1 port is polled per clock cycle) to determine if the input port is requesting a new connection to be established. If the input port being examined is idle or already has an active connection, then the controller skips this input port. Otherwise, the controller attempts to establish a connection, inserting two extra clock cycles to negotiate a new connection. Within these two clock cycles the controller checks that the output port being requested is available (1st clock cycle), updates the routing tables and reconfigures the crossbar (second clock cycle). Thereafter, the new connection can start to transmit data and the routing controller resumes its scan of input ports. Whenever the requested output connection cannot be established (due to output port being active on another connection), the routing controller does not process the request and proceeds to poll the next input port. In this scheme, between 15 and 45 clock cycles later the routing controller will reattempt to establish the requested connection.

In this manner, the routing controller continuously polls the routing table to determine if output port three (3) has become available. When this condition is finally met, the routing controller sets up the new crossbar connection (between input port #1 and output port #3) and updates the routing table. Thereafter, the packet data is asynchronously transmitted between the Input_FIFO(1) and Output_FIFO(3). During this transmission, the two control crossbars implement the handshaking, stopping the data-flow whenever congestion is encountered (such as overflow of input buffer on the receiving end of the connection). At the end of the data packet there is a special end-of-packet bit (not used for data coding) transmitted on a separate line. When this symbol is transmitted by Input_FIFO(1) it activates connection tear down logic and closes the connection. This occurs without the aid of the routing controller, hence facilitating parallel and asynchronous connection tear-down capability. The routing controller was tested by increasing the clock frequency and determining the maximum clock frequency at which the inputs were routed to the correct outputs. This occurred at a clock frequency of 50 MHz.

REFERENCES

[1] K. W. Goossen, J. A. Walker, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. D. Bacon, D. Dahringer, L. M. F. Chirovsky, A. L. Lentine, and D. A. B. Miller,, "GaAs MQW modulators integrated with silicon CMOS," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 360–62, Apr. 1995.

[2] A. V. Krishnamoorthy, A. Lentine, K. Goossen, J. Walker, T. Woodward, J. Ford, G. Aplin, L. D'Asaro,, S. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, L. Chirovsky, and D. Miller, "3-D integration of MQW modulators over active sub-micron CMOS circuits: 375 Mb/s transimpedance receiver-transmitter circuit," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 1288–1290, Nov. 1995.

[3] A. V. Krishnamoorthy and K. W. Goossen, "Progress in optoelectronic-VLSI smart pixel technology based on GaAs/AlGaAs MQW modulators," *Int. J. Optoelectron.*, vol. 11, no. 3, pp. 181–198, May–June 1997.

[4] A. V. Krishnamoorthy and F. E. Kiamilev, "Fanout, replication, and buffer sizing for a class of self-routing packet-switched multistage photonic switch fabrics," in *Tech. Dig. OSA Topical Meeting Photonic Switching*, Salt Lake City, UT, Mar. 1995, pp. 87–89.

[5] A. L. Lentine, K. W. Goossen, J. A. Walker, J. E. Cunningham, W. Y Jan, T K. Woodward, A. V. Krishnamoorthy, B. J. Tseng, S. P. Hui, R. E. Leibenguth, L, M. F. Chirovsky, R. A. Novotny, D. B. Buchholz, and R. L. Morrison, "Optoelectronic VLSI switching chip with over 1Tbit/s potential optical I/O bandwidth," *Electron. Lett.*, vol. 33, no. 10, pp. 894–895, May 1997.

[6] T. H. Szymanski, A. Au, M. Lafrenière-Roula, V. Tyan, B. Supmonchai, J. Wong, B. Zerrouk, and S. T. Obenaus, "Terabit optical local area network for multiprocessing systems," *Appl. Opt.*, vol. 37, no. 2, pp. 264–275, Jan. 1998.

[7] P. Lukowicz, S. Sinzinger, K. Dunkel, and H.-D. Bauer, "Design of an opto-electronic VLSI parallel fiber bus," *Proc. OSA Topical Meeting Optics in Computing (OC'98)*, Brugges, Belgium, June 1998, pp. 289–292.

[8] T. M. Pinkston, M. Raksapatcharawong, and Y. Choi, "WARRP core: Optoelectronic implementation of network-router deadlock-handling mechanisms," *Appl. Opt.*, vol. 37, no. 2, pp. 276–283, Jan. 1998.

[9] D. A. B. Miller and H. M. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture," *J. Parallel and Distrib. Comput.*, vol. 41, pp. 42–52, 1997.

[10] A. V. Krishnamoorthy and D. A. B. Miller "Firehose architectures for free-space optically interconnected VLSI circuits," *J. Parallel Distrib. Computing*, vol. 41, no. 1, pp. 109–114, Feb. 1997.

[11] A. Krishnamoorthy, J. Ford, K. Goossen, J. Walker, B. Tseng, S. Hui, J. Cunningham, W. Jan, T. Woodward, M. Nuss, R. Rozier, F. Kiamilev, and D. Miller, "The AMOEBA switch: An opto-electronic switch for multiprocessor networking using dense-WDM," in *Proc. IEEE Workshop Massively Parallel Processing Using Optical Interconnections*, Maui, HI, Oct. 1996, pp. 94–100.

[12] A. V. Krishnamoorthy, J. Ford, K. Goossen, J. Walker, A. Lentine, S. Hui, B. Tseng, L. Chirovsky, R. Leibenguth, D. Kossives, D. Dahringer, L. D'Asaro, F. Kiamilev, G. Aplin, R. Rozier, and D. Miller, "Photonic page buffer based on GaAs MQW modulators bonded directly over active silicon CMOS circuits," *Appl. Opt.*, vol. 35, no. 14, pp. 2439–2348, May 1996.

[13] T. K. Woodward, A. V. Krishnamoorthy, A. L. Lentine, and L. M. F. Chirovsky, "Optical receivers for optoelectronic VLSI," *IEEE J. Select. Topics Quantum Electron.*, vol. 2, pp. 106–116, Apr. 1996.

[14] Y. Li, T. Wang, H. Kosaka, S. Kawai, and K. Kasahara, "Fiber-image-guide-based bit-parallel optical interconnects," *Appl. Opt.*, vol. 35, no. 35, pp. 6920–6933, Dec. 1996.

[15] E. De Souza *et al.*, "Wavelength-division multiplexing with femtosecond pulses," *Opt. Lett.*, vol. 20, no. 10, pp. 1166–1168, 1995.

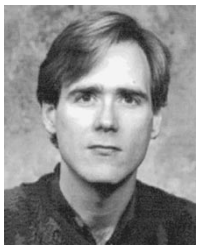
[16] M. Ramaswamy-Paye and J. G. Fujimoto, "Compact dispersion-compensating geometry for Kerr-lens mode-locked femtosecond lasers," *Opt. Lett.*, vol. 19, no. 21, pp. 1756–1758, 1994.

[17] S. Tsuda *et al.*, "Low-loss intracavity AlAs/AlGaAs saturable Bragg reflector for femtosecond mode locking in solid-state lasers," *Opt. Lett.*, vol. 20, no. 12, pp. 1406–1408, 1995.

[18] J.-M. Hopkins, G. J. Valentine, W. Sibbett, and U. Keller, "Highly compact femtosecond Cr^{3+} : LiSAF laser with a simplified cavity geometry," presented at the 1999 Conf. Lasers and Electro-Optics, Baltimore, MD, May 23–28, 1999, paper CTuA4.

[19] M. Zirngibl, C. H. Joyner, and C. R. Doerr, *et al.*, "An 18-channel multifrequency laser," *IEEE Photon. Technol. Lett.*, vol. 8, no. 7, pp. 870–872, 1996.
[20] M. Stern, W. I. Way, V. Shah, M. B. Romeiser, and W. C. Young, "800 nm digital transmission in 1300 nm optimized single-mode fiber," in *Proc. Optical Fiber Communications Conf. (OFC'87)*, 1987, Paper MD2.
[21] J. E. Ford, A. Krishnamoorthy, S. Tsuda, W. Knox, M. Nuss, and D. Miller, "WDM/SDM fiber network design for the AMOEBA optoelectronic switch," in *Proc. LEOS Topical Meeting Smart Pixels*, Keystone, Aug. 1996, pp. 97–98.
[22] K. W. Goossen, J. E. Cunningham, W. Y. Jan, and R. Leibenguth, "On the operational and manufacturing tolerances of GaAs-AlAs MQW modulators," *IEEE J. Quantum Electron.*, vol. 34, pp. 431–438, Mar. 1998.

Ashok V. Krishnamoorthy (S'90–M'93), for photograph and biography, see this issue, p. 156.



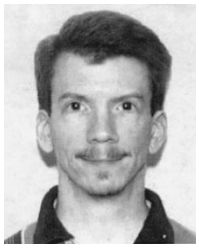
Joseph E. Ford (M'94) received the B.Sc. degree in physics from the University of California at Los Angeles (UCLA) in 1983, the Masters degrees in laser physics and optical engineering from the University of British Columbia and Rochester's Institute of Optics, respectively, and the Ph.D. degree in electrical engineering/applied physics from the University of California at San Diego in 1992 for studies in parallel optical interconnection by wavelength and random-phase addressed photorefractive holograms.

At Call/Recall, a 3-D data storage startup company, he demonstrated image storage in a doped-polymer cube using two-wavelength two-photon recording. In 1994, he became a member of technical staff in the Advanced Photonics Research Department of Bell Laboratories where he works on integrating surface-normal optoelectronic and micromechanical devices with SMF for optical modulation, switching, and data transmission.

Fouad E. Kiamilev (M'92) received the B.A.C.S. and the B.S.E.E. degrees (honors) in 1988, the M.S.E.E. degree in 1990, and the Ph.D. degree in electrical engineering in 1992, all from the University of California at San Diego.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering at the University of North Carolina at Charlotte (UNCC). Since joining UNCC in 1992, he has established a research group focused on optoelectronic VLSI IC development, architecture for free-space optically interconnected systems, and high-speed CMOS IC design. He has published over 20 journal papers, 70 conference papers, and files five patents. He has served on several program committees for the IEEE Lasers and Electro-Optics Society (LEOS), the IEEE Computer Society (CS), and the Optical Society of America.

Dr. Kiamilev was the recipient of the Office of Naval Research Graduate Fellowship.



Richard G. Rozier (S'91–M'92) received the B.S.E.E., M.S.E.E., and Ph.D. degrees from the University of North Carolina at Charlotte, all in electrical engineering, in 1992, 1996, and 1998, respectively.

He is currently employed as a Postdoctoral Member of Technical Staff in the Optical LAN's Research Department at Bell Laboratories, Lucent Technologies, Holmdel, NJ. His work focuses on the research and development of high performance optoelectronic VLSI ASIC design for computer networking applications. This includes custom printed circuit board design that provides testing capabilities as well as integration of the ASIC into a system.

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

Keith W. Goossen (S'82–M'83) received the B.S. degree in electrical engineering from the University of California at Santa Barbara, in 1983, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1988. In his doctoral thesis, he invented the first photovoltaic 10-μm detector based on photoemission from quantum wells.

In 1988, he joined AT&T Bell Laboratories, Holmdel, NJ, now Lucent Technologies, Bell Laboratories. While there his work has been primarily on surface-normal optical modulators and their integration with electronics for optical interconnects. This includes the first demonstration of high-speed, VLSI-density optical transceivers utilizing silicon CMOS and GaAs-based modulators via flip-chip bonding. He has also investigated integration of quantum well modulators on silicon via direct epitaxy, and produced extensive research on quantum well modulator technology, including the discovery of strong excitons in shallow quantum wells, and the development of stacked diode modulators and wafer-scale Fabry–Perot modulators. He has also invented the optical principle of the first broad-band micromechanical modulator for local access fiber networks. In addition he has developed modeling techniques for electrical propagation on normal and superconducting microstrip transmission lines.

Dr. Goossen is a member of the IEEE Lasers and Electro-Optics Society.

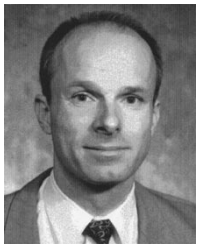
B. Tseng, photograph and biography not available at the time of publication.

James A. Walker received the B.S. and M.S. degrees in electrical engineering from Rutgers University, New Brunswick, NJ, in 1984 and 1989, respectively.

He joined Bell Laboratories in 1984. He is a Member of Technical Staff in the Advanced Photonics Research Department at Lucent Technologies, Bell Laboratories. His primary research interest is technology development for WDM telecommunications systems, with principle focus on MEMS components such as the MARS data transmitter, MARS variable attenuator, and membrane-based WDM equalizer. He is also actively involved in research on high-speed CMOS interconnects using optical I/O and contributed to the development of the optoelectronic VLSI technology currently being developed at Bell Labs. He was a member of the MEMS research group at Bell Labs from 1986 to 1990 and has contributed to a number of MEMS research areas including thin-film shape memory alloys, material science of polysilicon, and friction, wear, and reliability of micromechanical systems. He currently holds over 30 patents and has authored more than 90 articles and talks in the MEMS, Photonics, and Packaging areas.

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

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



Martin C. Nuss (M'90) was born in Lrrach, Germany. In 1982, he received the Diploma and in 1985 the Dr. rer. nat. degree in physics, both from the Technical University Munich, Germany. He joined Bell Laboratories, Murray Hill, NJ, in 1985, as a Post-Doctoral member of the Technical Staff. Since 1987, he has been located in Holmdel, NJ, first as a Member, and later as a Distinguished Member of the Technical Staff. Since 1998, he is the Head of the Optical Enterprise Networks Research Department in Holmdel. His work in ultrafast optoelectronics led to the use of terahertz bandwidth electromagnetic transients for time-resolved spectroscopy at far-infrared frequencies and the study of time-dependent phenomena in semiconductors and superconductors. More recently, he brought terahertz spectroscopy to the commercial arena through the invention of terahertz "T-ray" imaging. He is also the coinventor of the Chirped-Pulse wavelength-division-multiplexing technique. His current work is centered on optical networking and next-generation enterprise networking technologies such as 10-Gb Ethernet.

Dr. Nuss is a Fellow of the Optical Society of America and a member of the IEEE Lasers and Electro-Optics Society.