JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 14, NO. 5, MAY 1996

# An All-Optical Network Architecture

Thilo Gipser, *Student Member, IEEE,* and Ming-Seng Kao, *Member, IEEE*

***Abstract*— A novel photonic network, MATRIX (for Multi-wavelength All-optical TRansparent Information eXchange), is proposed in this paper. The all-optical multihop network supports wavelength continuity and provides a very high network capacity. Spatial reuse of wavelengths as well as the multiplicity of fibers in optical fiber cables are exploited and enable the interconnection of $N^2$ network nodes with merely $N$ wavelengths. The node structure is simple since neither tunable devices nor wavelength converters are required. Packets are routed through the network by photonic fast packet switching as well as by wavelength and experience a maximum hop number of two. Multiple optical paths between any pair of nodes provide a good network survivability.**

## I. INTRODUCTION

THE EVOLUTION of network technology has led to three generations of networks so far, distinguishable according to their underlying physical-level technology. The first one, based on copper-wire or microwave-radio technology, was developed before fiberoptic technology emerged and fully employs electronic equipments. Examples include well-known networks as Ethernet, Token Bus, Token Ring, and ARPANET. The second generation already exploits some advantages of the fiberoptic transmission technology such as the high speed of laser diodes and photonic components, or the low attenuation and dispersion of optical fibers. The partial or complete replacement of copper point-to-point links with silica fibers is characteristic for these networks and leads to better performance parameters such as higher data rates, lower error probabilities, and reduced electromagnetic sensitivity. However, changes in topology and protocol are negligible compared to networks of the first generation. Examples for second generation networks include FDDI and DQDB. Third generation or photonic networks make use of the probably most important property of optical fibers: the vast optical bandwidth. Exploitation of this bandwidth can be achieved in applying multiplex techniques with regard to wavelength (WDM), time (TDM), or waveshape (CDMA). The current favorite choice seems to be WDM because of the lower system complexity. Examples for WDM packet networks include LambdaNet [1], ShuffleNet [2], [3], and multihop networks based on de Bruijn graphs [4]. A large number of networks of all three generations are excellently surveyed in [5].

All-optical networks provide further advantages. In these networks, the information is kept permanently in the optical domain during the whole transmission from the source to the destination node. This prevents limiting electronic bottlenecks in intermediate nodes, caused by a detour of the transmitted information over the electrical domain. The data rate of the packet payload can exceed substantially the processing speed of intermediate nodes since fast electronic buffering of transit packets is avoided.

Photonic networks rely on a fiber infrastructure from which wide parts are already installed today [6]. A strong trend is going toward an increase in the number of fibers per optical fiber cable, for three reasons. First, future data traffic is expected to grow rapidly and the capacity of future optical networks will exceed that of todays networks by several orders of magnitude [7]. Second, installation cost, the main factor in the total cable cost, is not affected by the number of fibers inside a cable [8]. And third, contrary to electrical cables optical fiber cables may contain thousands of fibers [9]. It is therefore advantageous to consider the multiplicity of fibers per cable in the network design.

Network failures in high capacity networks can result in enormous loss of traffic and revenue. Desirable for the future are therefore lightwave networks with a good survivability. Mesh networks with their potentially high connectivity are very well suited to satisfy high survivability demands. Different mesh networks for all three network generations have therefore been proposed so far [10]–[13]. The well-known Manhattan street network (MSN) [10] is based on a regular mesh structure with an even number of rows and columns. All nodes located in the same row or column are connected in unidirectional rings and the connection grid of all nodes forms a toroidal surface. Packets are routed by deflection and according to the consistent number of incoming and outgoing links; node buffering is not necessary. The slotted interconnected grid network (SIGnet) [11] is a photonic network with a piecewise-regular topology. The nodes are bidirectionally interconnected with WDM channels and packets are routed by deflection. The RookNet [12] is a regular mesh network. Each node is equipped with two transmitters for the transmission of packets to the nodes in the same row and column, respectively, and can receive packets simultaneously from all row and column nodes. In the intermediate nodes, transit packets are received, concentrated, buffered and retransmitted.

A novel photonic multihop network, called MATRIX [14], [15], is proposed in this paper. The all-optical high-speed network has a mean hop number less than two and supports wavelength continuity, i.e., no wavelength translation is required in the network. Wavelength reuse as well as concurrency in wavelength and space enable the interconnection of $N^2$ nodes with merely $N$ wavelengths. The node structure

---

Manuscript received February 1, 1995; revised February 12, 1996.

T. Gipser is with the Swiss Federal Institute of Technology, Communication Technology Laboratory, ETH Zentrum, 8092 Zurich, Switzerland.

M.-S. Kao is with the Department of Communication Engineering and Center for Telecommunications Research, National Chiao Tung University, Hsinchu, Taiwan.

Publisher Item Identifier S 0733-8724(96)03877-7.

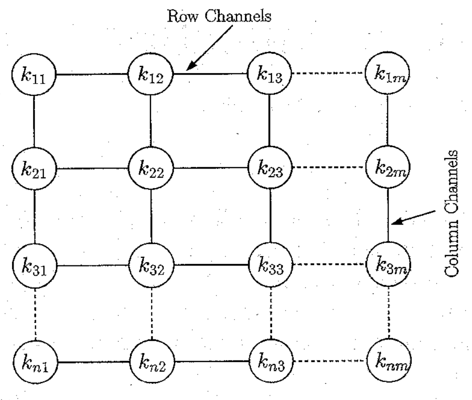


Fig. 1. Grid structure of an $nxm$MATRIX ($k_{ij}$: Network node located in row $i$and column $j$).

is simple and the nodes are equipped with integrated arrays of fixed-wavelength laser diodes and optical receivers. The lightwave network is intended to be used as a very high capacity metropolitan area network (MAN) or a wide area network (WAN).

The paper is structured as follows: The architecture of MATRIX and the structure of its nodes are introduced in Section II, followed by a detailed performance analysis in Section III. Network simulations are presented in Section IV and Section V is concerned with the network survivability. Section VI outlines the behavior of the network under nonuniform traffic conditions. A centralized carrier supply of the network nodes is discussed in Section VII and some concluding remarks are given in Section VIII.

## II. System Description

### A. Network Architecture

The network nodes, having an arbitrary geographical distribution and being interconnected with optical fiber cables, are virtually organized in a matrix structure as shown in Fig. 1. It is not necessary to realize every single link between two nodes with a separate cable since several links can be integrated into one optical fiber cable instead. The

$$M = n \cdot m;\ n, m \in \{1, 2, \ldots\}, nm > 1 \qquad (1)$$

network nodes are organized in $n$ different rows and $m$ different columns. In the following, all $m$ nodes located in a specific row are referred to as *row nodes* and all $n$ nodes located in the same column as *column nodes*, respectively. The fiber cable links between the network nodes consist of multiple fibers. Every row fiber cable interconnects all row nodes and contains $m$ fibers. Analogously, every column fiber cable interconnects all column nodes and contains $n$ fibers. One fiber from both the row and column fiber cable is assigned to each node.

Two communication schemes are used in the network.

- *Communication via Row Channels:* All row nodes communicate with each other via direct paths located in a row fiber cable.

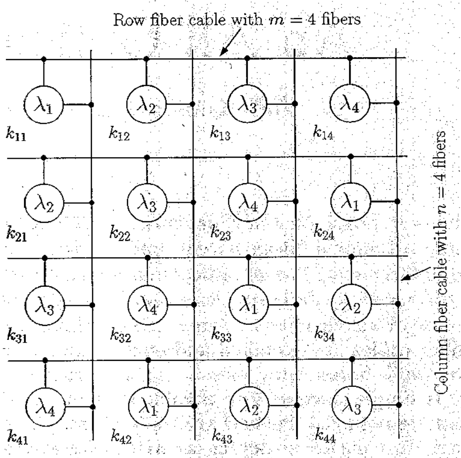


Fig. 2. Recurrent wavelength assignment in a 4 x 4 network ($\lambda_i$: node identification wavelength (NIW)).

- *Communication via Column Channels:* All column nodes communicate with each other via direct paths located in a column fiber cable.

Every node has therefore direct links to the $(m-1)$ nodes in the same row and to the $(n-1)$ nodes in the same column. No direct paths exist to the remaining nodes in the network.

A node identification wavelength (NIW) is assigned to each network node $k_{ij}$ located in the $i$th row and the $j$th column. All NIW's used in the network are elements of the set

$$\Lambda = \{\lambda_l \mid l \in \{1, 2, \cdots, N\}\} \qquad (2)$$

with

$$N = \max(n, m). \qquad (3)$$

All $M$ network nodes can be interconnected with these $N$ wavelengths. The assignment of a specific wavelength to a network node follows a certain wavelength assignment scheme. This scheme has to fulfill the following requirement:

> *A certain wavelength $\lambda_l$ may appear repeatedly in the network but at maximum once in each column and once in each row.*

This requirement prevents wavelength conflicts on row and column channels since every NIW occurs at most once in each row and each column. The relation

$$l \equiv (i + j - 1)(\mathrm{mod} N),\ i \in \{1, 2, \cdots, n\},\ j \in \{1, 2, \cdots, m\} \qquad (4)$$

satisfies the given requirement. The appropriate wavelength assignment to the network nodes is depicted in Fig. 2.

In order to avoid wavelength translation during packet transmission through the network, two different media access strategies are applied to communicate over row and column channels. The following strategy is employed for the communication of a network node with other row nodes:

> A node *transmits* packets to the other $(m-1)$ row nodes via its row fiber and on the wavelength that is identical with the NIW of the desired destination node in this row. It *receives* packets on its NIW from the remaining row fibers assigned to other row nodes.

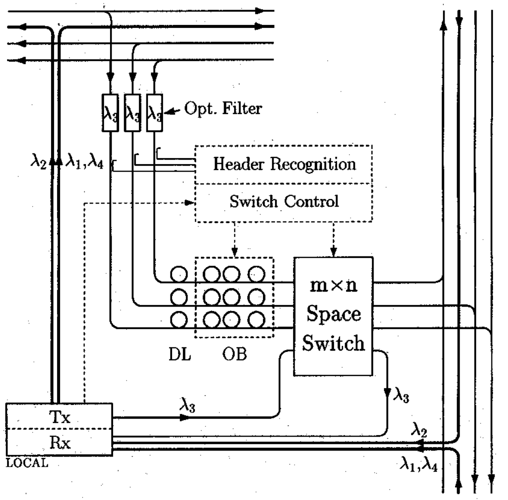


Fig. 3. Node $k_{22}$ of a 4 x 4 network (Tx: transmitter, Rx: receiver, DL: delay line, OB: optical buffers).

An example is shown in Fig. 3 where the second row fiber (from the top) and the NIW $\lambda_3$ have been assigned to the node $k_{22}$ of a $4 \times 4$ MATRIX. The node transmits packets to other row nodes via its row fiber and on all wavelengths except its NIW. The wavelength chosen for a packet determines the destination node in this row. The node receives all packets coming from $(m-1)$ row fibers on the NIW $\lambda_3$.

Exactly the opposite access strategy is applied for the communication of a network node with other column nodes:

> A node *receives* packets from the other $(n-1)$ column nodes via its column fiber and on the wavelength that is identical with the NIW of the source node in this column.
> It *transmits* packets on the NIW and via the remaining column fibers assigned to other column nodes.

In the example in Fig. 3, the second column fiber (from the left) has been assigned to the node $k_{22}$. The node transmits packets on its NIW to other column nodes via appropriate column fibers. The fiber chosen for a packet determines the destination node in this column. On the other hand, the node receives packets on $(n-1)$ different wavelengths from its column fiber.

The use of two different media access strategies allows a simple all-optical packet transfer from row to column channels. All packets received from other row nodes are on the same wavelength (the NIW) as are the packets transmitted to other column nodes. The processing of optical packets in a node can therefore be done on one wavelength only and without any wavelength translation. This leads to a simple and uniform structure of the network nodes.

### *B. Routing*

Packets are routed through the network according to the wavelength and packet header information (containing the destination address). Since the average number of hops in a multihop network significantly influences both network throughput and delay, the routing algorithm has been chosen such that the average hop number in the grid topology is kept as small as possible. Packets in MATRIX reach the intended destination in one hop via the appropriate row (column) channels if the source node $k_{ij}$ is located in the same row (column) as the destination node $k_{kl}$. Otherwise, two hops are necessary and packets are transmitted first on row channels to an intermediate node $k_{il}$ and then on column channels to the desired destination node:

$$\text{if} \begin{cases} i = k & : k_{ij} \rightarrow k_{il} \\ j = l & : k_{ij} \rightarrow k_{kj} \\ i \neq k, j \neq l & : k_{ij} \rightarrow k_{il} \rightarrow k_{kl}. \end{cases} \tag{5}$$

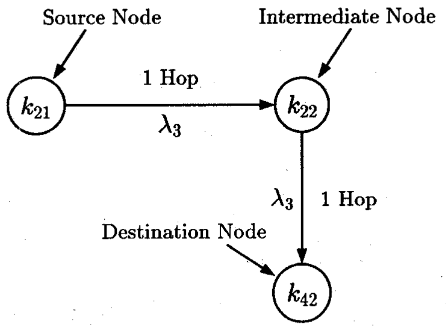


Fig. 4. Routing in MATRIX (Example: All-optical packet transmission from node$k_{21}$to node$k_{42}$).

An example is shown in Fig. 4 where packets are transmitted from the source node $k_{21}$ to the destination node $k_{42}$. Packets are forwarded first on the row channel from node $k_{21}$ to the intermediate node $k_{22}$. This node is located in the same row as the source node $k_{21}$ and the same column as the destination node $k_{42}$. Then, packets are transmitted on the column channel from the intermediate node to the destination node $k_{42}$.

Routing in the columns of the network is accomplished by selection of appropriate fibers for transit packets. Routing in the rows of the network is done by wavelength. No wavelength translation is necessary what reveals also the spanning tree of a network node as shown in Fig. 5. The spanning tree represents a graphical representation of all paths outgoing from a specific network node and leading to all other nodes. Fig. 5 shows that each network node can be reached with one or two hops.

### *C. Node Configuration*

The configuration of a network node is depicted in Fig. 3. Each node has $(m-1)$ incoming row channels. Optical filters (or frequency selective couplers) ensure that only those packets on the same wavelength as the NIW of the node are received. Incoming packets are either destined for the node or have to be switched to outgoing channels. All headers of incoming packets are read with low-speed optical receivers in order to get the destination addresses [16]. This header recognition of packets passing by is accomplished by coupling out and detecting a small fraction of the lightwave power or, alternatively, directly by a photodetection with a semiconductor laser amplifier [17]. Before they reach the optical buffers, the incoming packets are delayed for one time slot in order to enable the switch control to consider them in the next service decision.

The heart of each node is an $m \times n$ space switch. This space switch routes the incoming packets without wavelength

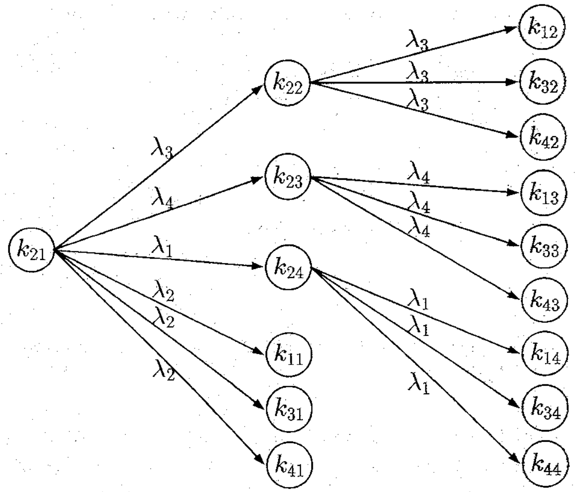


Fig. 5. Directed spanning tree rooted at node $k_{21}$ of a 4 x 4 MATRIX.

translation from row to appropriate column channels or to the local unit of the network node. An inherent problem in packet-switched networks is contention, i.e., the situation that several packets at different input ports of a switch simultaneously demand access to the same output port. Input queueing along with a contention resolution algorithm may be used to resolve the conflict of output request. Different promising approaches have been reported recently concerning the buffering of optical packets in delay lines [18], [19] or recirculating delay lines [20]–[23]. In the present paper, two arrangements are investigated further with respect to an employment in the all-optical network. The first one is a quite simple arrangement [19] consisting of delay lines alternating with optical switches as shown in Fig. 6(a). In each delay line a packet suffers a delay of one time slot. A packet can be delayed for $b$ time slots at maximum and gets lost if it could not be read out within this time. The other arrangement is composed of a set of recirculating optical delay lines as shown in Fig. 6(b). Optical losses in the memory loops are compensated by optical amplifiers. The round trip time in a loop equals one time slot and a specific packet can be stored for multiple slots. New packets arriving at the memory loops get lost if all input buffers are full and if there is no possibility to forward them in the same time slot through the optical space switch. The switch control employs a simple and fast service discipline in order to read out the packets from the optical buffers. The feasibility of such buffers has been investigated intensively. It has been shown that optical packets can be stored for tens of round trips in recirculating delay lines maintaining a BER well below $10^{-9}$ [21] and that in principle infinite numbers of round trips are possible when optical signal regeneration is applied [25]. Furthermore, it is possible to store more than one packet in such an optical memory loop [15], [22]–[24]. However, optical loss is still high in these arrangements and optical buffering should be kept as small as possible. In MATRIX, buffering and switching of optical packets is needed merely once. Note that optical buffering can be avoided if time multiplexing is applied additionally.

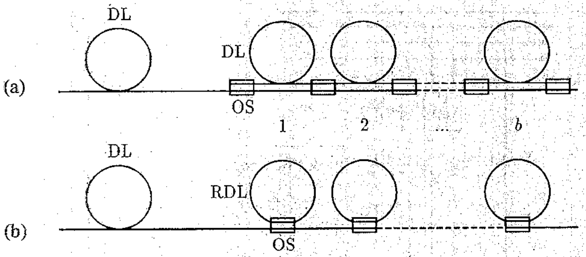


Fig. 6. Optical buffering with delay lines (a) and recirculating delay lines (b) (DL: delay line, RDL: recirculating delay lline, OS: switched directional coupler, $b$: number of buffers).

In the local unit, which consists of the transmitter and receiver unit, newly generated packets are offered in parallel from $m$ transmitters on $m$ different wavelengths to the optical network, and relayed optical packets destined for the node are removed from $n$ receivers on $n$ wavelengths from the network. The wavelengths of $(m-1)$ transmitters are combined in a wavelength multiplexer and fed into the row fiber of the node. Packets generated on the NIW are sent via the space switch to the column channels. Analogously, packets are received on the NIW from the row channels and on $(n-1)$ wavelengths from the column fiber of the node. Both the fixed-wavelength array of $m$ transmitters and the fixed-wavelength array of $n$ receivers show a high potential for monolithical integration with other components [26]–[28]. Such integrated and compact devices have an improved reliability at reduced per-wavelength component cost and are therefore well suited for an application in lightwave networks. Additionally, they enable every network node a simultaneous access to all wavelengths. The development of such devices is an important issue in the current research on all-optical networks [29].

## III. PERFORMANCE EVALUATION

A uniform traffic pattern is assumed for the following calculations, i.e., a message generated at any network node is equally likely to be destined for any other node.

### A. Average Hop Number

The connectivity graph of a network is a graphical representation of all logical links in the network. A regular connectivity graph indicates that each node is in a similar way connected to all other network nodes. The average hop number can then be expressed as

$$\bar{h} = \frac{1}{M-1} \sum_{h=1}^{h_{\max}} h M_h \tag{6}$$

where $M_h$ stands for the number of nodes being reachable from a given source node with $h$ hops. The connectivity graph of MATRIX is regular since the number of incoming channels equals the number of outgoing channels at each node and the same number of nodes can be reached from each node. From each network node $M_1 = n + m - 2$ nodes can be reached in one hop and the remaining $M_2 = nm - n - m + 1$ nodes in two hops. For $nm > 1$ the average hop number is thus

$$\bar{h}(n,m) = \frac{2nm - n - m}{nm - 1} = 2 - \frac{n + m - 2}{nm - 1}. \tag{7}$$

Note, that for $nm = \max(n,m)$, i.e., the network consists of one row or column only, the mean hop number equals one. It can be seen from (7) that

$$1 \leq \bar{h}(n,m) < 2 \tag{8}$$

and that the diameter $h_{\max}$ of the network, defined as the maximum hop distance of the shortest paths between any pair of nodes, equals two. A small average hop number is important in any multihop network not only to minimize the mean propagation delay but also to maximize the number of new packets which can enter the network per time slot and, thus, the network capacity.

### *B. Network Delay*

A packet traversing MATRIX experiences an end-to-end delay

$$\bar{D} = \bar{h} \cdot \bar{\tau}_{\text{prop}} + \bar{\tau}_{\text{node}} + \tau_{\text{trans}} \tag{9}$$

where $\bar{\tau}_{\text{prop}}$ denotes the mean propagation delay per hop, $\bar{\tau}_{\text{node}}$ the mean waiting time in the optical buffers and $\tau_{\text{trans}}$ the transmission delay (equal to one time slot). Since packets need to be buffered only once during their transmission and the average hop number $\bar{h}$ given in (7) is moderate, small end-to-end delays can be expected.

### *C. Mean Channel Efficiency*

The channel efficiency, a measure of how efficiently channels are used in a network, depends on the average hop number and the traffic load in the network.

Each node has $(m-1)$ incoming row channels (on the NIW) and $(n-1)$ incoming column channels (on different wavelengths). Similarly, $(n-1)$ column channels and $(m-1)$ row channels are originating from each node. The connectivity degree $c$, defined as the number of channels leading to or originating from a node [30] equals, therefore,

$$c = n + m - 2. \tag{10}$$

The total number of channels in the network is then

$$w_T = c \cdot M. \tag{11}$$

The mean channel efficiency $\bar{\eta}$ in a multihop network can be expressed as

$$\bar{\eta} = \frac{1}{w_T} \sum_{w=1}^{w_T} S^{(w)} \sum_{h=1}^{h_{\max}} \frac{p_h^{(w)}}{h} \tag{12}$$

where $S^{(w)}$ denotes the normalized successful load on channel $w$ and $p_h^{(w)}$ the probability that a packet on channel $w$ needs $h$ hops to reach its destination.

In MATRIX, the normalized successful loads on all row channels (index $R$) are identical and so are the loads on the column channels (index $C$). For $nm > \max(n,m)$, i.e., the network consists of at least two rows and two columns, one has

$$\bar{\eta} = \frac{1}{w_T}\left(w^{(R)} \cdot S^{(R)} \sum_{h=1}^{h_{\max}} \frac{p_h^{(R)}}{h} + w^{(C)} \cdot S^{(C)} \sum_{h=1}^{h_{\max}} \frac{p_h^{(C)}}{h}\right). \tag{13}$$

With the total number of channels $w_T = w^{(R)} + w^{(C)} = nm(m-1) + nm(n-1)$, the load on a row channel $S^{(R)} = \frac{n}{\max(n,m)}$, the load on a column channel $S^{(C)} = \frac{m}{\max(n,m)}$ and the packet probabilities $p_1^{(R)} = \frac{1}{n}$, $p_2^{(R)} = \frac{n-1}{n}$, $p_1^{(C)} = \frac{1}{m}$, $p_2^{(C)} = \frac{m-1}{m}$ follows the mean channel efficiency as

$$\bar{\eta}(n,m) = \frac{nm - 1}{\max(n,m) \cdot (n + m - 2)}. \tag{14}$$

The channel efficiency for $nm = \max(n,m)$ can be calculated analogously. Hence, one has

$$\bar{\eta}(n,m) = \begin{cases} \frac{nm-1}{\max(n,m)\cdot(n+m-2)} & \text{if } nm > \max(n,m), \\ \frac{1}{\max(n,m)-1} & \text{if } nm = \max(n,m). \end{cases} \tag{15}$$

From (15) it can be seen that

$$\bar{\eta}(n,m) = \bar{\eta}(m,n) \tag{16}$$

and for $n = m = N$ we get the identity

$$\bar{\eta}(N,N) = \frac{1}{\bar{h}(N,N)} = \frac{N+1}{2N}. \tag{17}$$

With the parameter

$$\beta = \begin{Bmatrix} \frac{n}{m} & \text{if } m \geq n \\ \frac{m}{n} & \text{otherwise} \end{Bmatrix} = \frac{\min(n,m)}{\max(n,m)}, \quad \beta \leq 1 \tag{18}$$

indicating the ratio between the number of rows and columns in the network, the channel efficiency $\bar{\eta}$ in (15) can be expressed as

$$\bar{\eta}(\beta, M) = \begin{cases} \frac{M-1}{M + \frac{M}{\beta} - 2\sqrt{\frac{M}{\beta}}} & \text{if } nm > \max(n,m), \\ \frac{1}{\sqrt{\frac{M}{\beta}} - 1} & \text{if } nm = \max(n,m) \end{cases} \tag{19}$$

and becomes for large $M$

$$\lim_{M \to \infty} \bar{\eta}(\beta, M) = \begin{cases} \frac{1}{1+\beta^{-1}} & \text{if } nm > \max(n,m), \\ 0 & \text{if } nm = \max(n,m). \end{cases} \tag{20}$$

The mean channel efficiency $\bar{\eta}$ as a function of $M$ for different values of $\beta$ is shown in Fig. 7. Note, that a high channel efficiency is achievable in the network which reaches the maximum for $\beta = 1.0$ with a given $M$. The limit of the channel efficiency given in (20) is also plotted for different values of $\beta$.

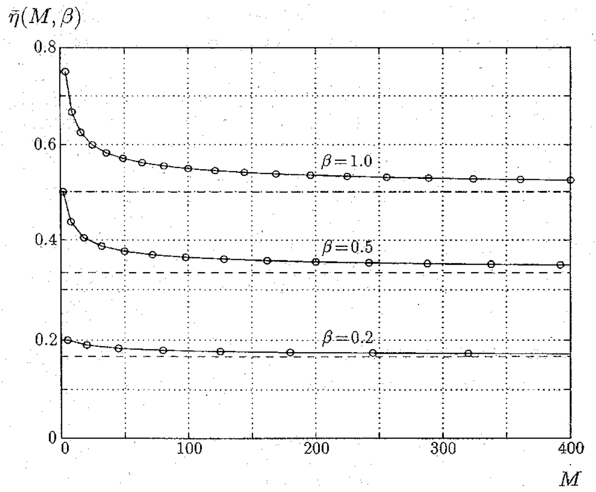


Fig. 7. Mean channel efficiency $\bar{\eta}$as a function of the number of nodes $M$for different values of $\beta$.

### D. Network Capacity

Channel capacity in multihop networks has to be shared between newly generated traffic and relayed traffic. The maximum new traffic generated by all nodes which may be transmitted through the network defines the network capacity.

In general, the network capacity $C$ of a multihop network can be expressed as the sum of all link capacities $C^{(w)}$, i.e.,

$$C = \sum_{w=1}^{w_T} C^{(w)} = \sum_{w=1}^{w_T} S^{(w)} \sum_{h=1}^{h_{\max}} \frac{p_h^{(w)}}{h}. \tag{21}$$

For equal hop-probabilities $p_h^{(w)} = p_h$ on every channel $w$, (21) changes to

$$C = \sum_{h=1}^{h_{\max}} \frac{p_h}{h} \sum_{w=1}^{w_T} S^{(w)}. \tag{22}$$

With[1]

$$\sum_{h=1}^{h_{\max}} \frac{p_h}{h} = \frac{1}{\bar{h}} \tag{23}$$

the network capacity follows as

$$C = \frac{1}{\bar{h}} \sum_{w=1}^{w_T} S^{(w)}. \tag{24}$$

A small average hop number $\bar{h}$ is therefore crucial in order to achieve a high network capacity. From (12) and (21) one has

$$C = \bar{\eta} \cdot w_T. \tag{25}$$

The network capacity of MATRIX under uniform traffic conditions can be expressed with (15), (11), (10), and (1) as

$$C(n,m) = \begin{cases} \min(n,m) \cdot (nm-1) & \text{if } nm > \max(n,m), \\ \max(n,m) & \text{if } nm = \max(n,m). \end{cases} \tag{26}$$

From (26) the expression

$$C(n,m) = C(m,n) \tag{27}$$

holds, i.e., an $n \times m$ MATRIX has the same capacity as an $m \times n$ network.

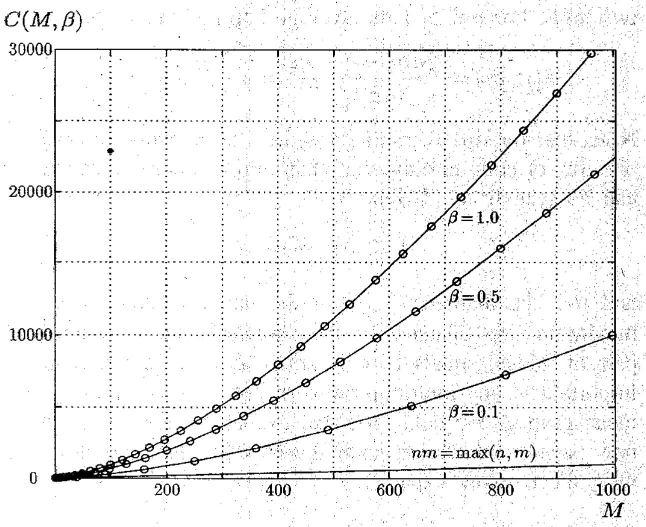


Fig. 8. Network capacity $C$ as a function of the number of nodes $M$for different values of the parameter $\beta$.

The combination of (18) and (26) yields the network capacity as a function of $\beta$ and $M$

$$C(\beta,M) = \begin{cases} (M-1)\sqrt{\beta M} & \text{if } nm > \max(n,m), \\ \sqrt{\frac{M}{\beta}} & \text{if } nm = \max(n,m). \end{cases} \tag{28}$$

Since each node has access to $c$ channels, the throughput per node equals

$$\gamma = c \cdot \bar{\eta} \tag{29}$$

and with (15) and (10) follows:

$$\gamma(n,m) = \begin{cases} \frac{nm-1}{\max(n,m)} & \text{if } nm > \max(n,m), \\ 1 & \text{if } nm = \max(n,m). \end{cases} \tag{30}$$

In Fig. 8 the overall capacity $C$ of MATRIX is plotted as a function of $M$ and for different values of $\beta$. Given $M$, the best performance can be obtained for $\beta = 1.0$.

The maximum data rate through the network results from

$$R(n,m) = R_C \cdot C(n,m) \tag{31}$$

where $R_C$ represents the maximum data rate on the network channels. The high connectivity degree $c$ together with the good channel efficiency $\bar{\eta}$ provide a huge capacity in the network. As an example, in a 30 × 30 core network operating at a data rate of $R_C = 10$ Gbit/s, the aggregate network throughput is approximately 270 Tb/s. With an average data rate of 10 Mb/s for each subscriber, 30 000 end users per node or $27 \times 10^6$ subscribers can totally be supported.

## IV. NETWORK SIMULATIONS

The analytical results obtained in the previous section have been compared with network simulations. The network has been implemented as a slotted system and for $N = n = m = 4$ in the simulation environment OPNET$^{\text{TM}}$. Packets generated in the network follow a uniform traffic pattern

---

[1] Note, that $p_h$ is *not* the probability of a $h$-hop packet leaving the network; therefore $\bar{h} \neq \sum_{h=1}^{h_{\max}} h p_h$.

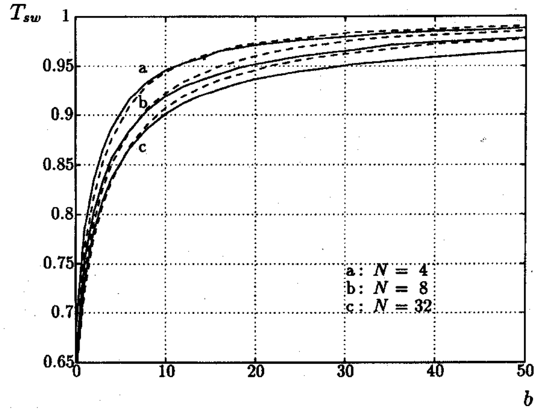


Fig. 9. Throughput performance $T_{sw}$ of an $NxN$space switch with optical input buffering as a function of the number of buffers $b$ for different switch sizes $N$ (— arrangement with recirculating delay lines, - - arrangement with delay lines).

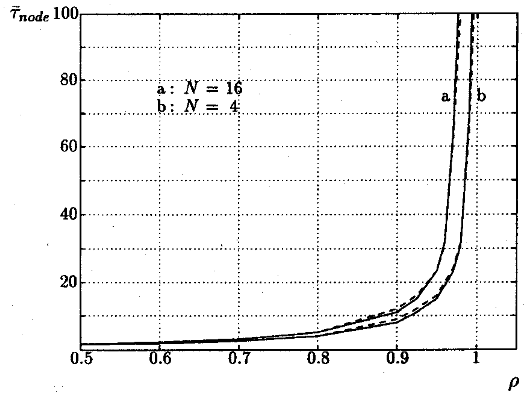


Fig. 10. Mean waiting time $\bar{\tau}_{node}$ (in slots) in the optical buffer arrangements as a function of the load $p$ for different switch sizes $N$ (— arrangement with recirculating delay lines, - - arrangement with delay lines).

and are transmitted simultaneously on the network links. Furthermore, the mean propagation delay per hop $\bar{\tau}_{\rm prop}$ in the network is equal to one time slot. A simple and therefore fast service discipline (described in [31, p. 1588]) was applied for service decisions in the network nodes.

Figs. 9 and 10 show the throughput $T_{\rm sw}$ and the mean waiting time $\bar{\tau}_{\rm node}$ of an $N \times N$ space switch with optical input buffers consisting of delay lines and recirculating delay lines as given in Fig. 6. Note, that despite the different physical configurations the performance of both buffering arrangements is very similar. The slightly higher throughput of the buffering arrangement with optical delay lines compared to that one with recirculating delay lines is due to the fact that packet losses in the first arrangement occur at the end of the buffering stage. Since the contention resolution algorithm considers the packets at the end of the buffering stage first, "problem" packets are automatically chosen with higher priority in the service decisions. However, the arrangement with recirculating delay lines shows a higher flexibility compared to the arrangement with optical delay lines.

Optical input buffering with recirculating delay lines was chosen for the simulations of a 4 × 4 network. Fig. 11 shows the aggregate network throughput for different numbers of optical buffers $b$ per input queue. The maximum achievable throughput $T$ coincides with the network capacity analytically derived in (26). The results shown in Figs. 12 and 13 are obtained when packet acknowledgment (ACK) is taken into consideration. It can be seen that the throughput equals the maximum achievable throughput with increasing $\rho$ as long as the links are not saturated due to retransmitted packets. The more buffers are used the lower is the probability that a packet has to be retransmitted. Note, that the delay increases significantly when the network links get saturated.

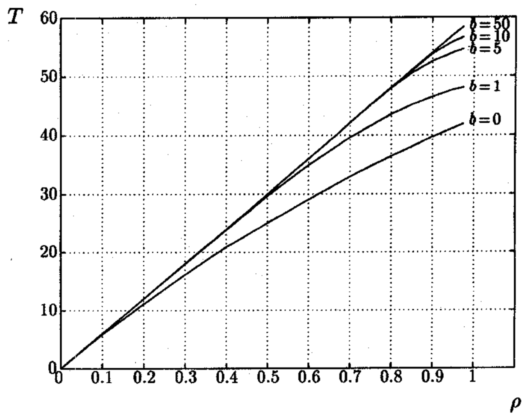


Fig. 11. Aggregate network throughput $T$ as a function of the load $p$ for different buffer sizes $b$.

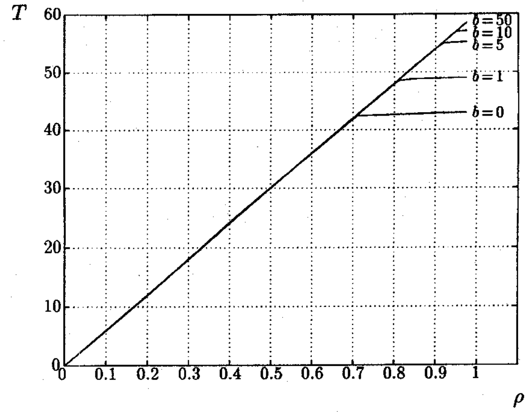


Fig. 12. Aggregate network throughput $T$ in the network with ACK as a function of the load $p$ for different buffer sizes $b$.

## V. NETWORK SURVIVABILITY

Survivability is a key concern in networking and will become increasingly important in future optical networks as they carry a steadily growing traffic. Any failures in high capacity networks could result in an enormous loss of traffic and revenue [32]. In case of a link or node failure the network has to retain its connectivity, i.e., as few nodes as possible should be affected by the failure. According to their topology mesh networks are well suited to satisfy high reliability demands.

The simplest way to overcome a node or link failure is to force any node in case of a failure to send packets first to those network nodes, whose links to the destination node

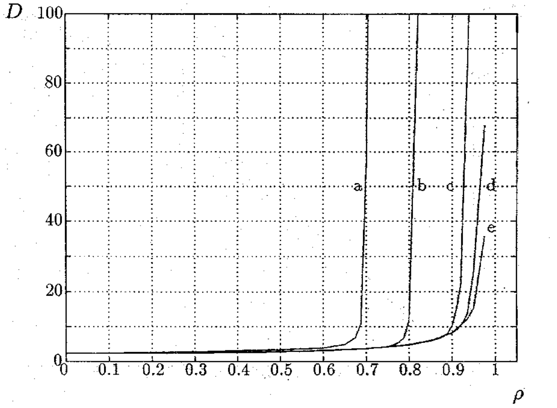


Fig. 13. End-to-end delay $D$ (in slots) in a network with ACK as a function of the load $p$ for different buffer sizes $b$(a: $b = 0$,b: $b = 1$,c: $b = 5$,d: $b = 10$, e: $b = 50$.

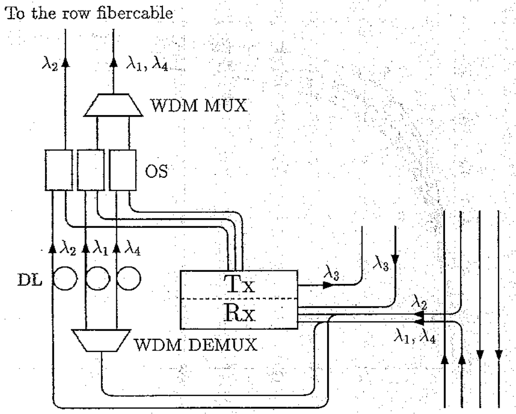


Fig. 14. Enhanced network node $k_{22}$ of a 4 x 4 MATRIX (MUX: wavelength multiplexer, DEMUX: wavelength demultiplier, DL: delay line, OS: optical switch).

are not affected by the failure. These additional intermediate nodes receive the packets and retransmit them like newly generated packets to the destination nodes. However, this failure treatment implies a detour via the electrical domain. In contrast, a slight enhancement of the node structure provides all-optical paths even in the case of a network failure.

MATRIX is composed of different subnets consisting of space switches as well as the channels leading to each switch from other row nodes and those leading from the switch to other column nodes. Each of these subnets works on a specific NIW and forward only traffic on this wavelength. Different subnets are independent and have no links between them. In the whole network exactly

$$\frac{M}{N} = \min(n, m) \tag{32}$$

space switches work on a given wavelength $\lambda_l$. A small enhancement of the nodes enables the interconnection of the $\min(n, m)$ space switches working on the same wavelength and, therefore, many additional all-optical paths in the network.

According to Fig. 3, a node receives packets exactly on the same wavelengths from the column nodes as it uses for its communications with other row nodes. This property is exploited in the enhanced node structure as shown in Fig. 14. The enhancement enables the nodes to route packets not only from row to column channels but also from column to row channels. All incoming packets from column channels are received as discussed in Section II, but some of them might now be destined for another node. In that case, a copy of the transit packet in the appropriate delay line is forwarded to the row fiber in the next time slot. New packets generated at the node and requiring simultaneous access to the same row fiber channel are retained and transmitted in the next free time slot. This backpressure mechanism automatically throttles sources as network failures occur and fits continuously the number of packets offered to the optical network to the available transmission capacity. Note, that the modification of the node is very simple since neither additional receivers (for the packet header recognition) wavelength converters nor additional optical buffers are required.

Many additional optical paths are provided in a network with enhanced network nodes. The transmitter units (receiver units) in an $n \times m$ network with $n \leq m$ ($n > m$) need the whole set $\Lambda$ for their communications. Since the receiver units (transmitter units) dispose of a wavelength set $\Lambda'$ with $\Lambda' \subseteq \Lambda$, the following relation holds:

> *In an arbitrary* $n \times m$ *network are always* $\min(n, m)$ *optical paths on different wavelengths available between any pair of enhanced network nodes.*

An example is shown in Fig. 15. If in the $4 \times 4$ network the link between node $k_{13}$ and node $k_{23}$ fails, the direct optical channel from source node $k_{13}$ to destination node $k_{43}$ is disrupted. However, other optical paths on different wavelengths can be used in order to send packets all-optically to their destinations. It is therefore possible to transmit information on different paths between two nodes and the network remains fully connected after any link or node failure.

The enhancement of a node shown in Fig. 14 is not sufficient for $n \times m$ networks with $n \neq m$ in certain failure cases, because the ability to receive a given wavelength may be not available in some rows or columns. This problem can be solved if appropriate nodes are provided with the ability to forward packets also on other wavelengths or if wavelength translation is applied to the packet copies in enhanced network nodes.

## VI. NETWORK BEHAVIOR UNDER NONUNIFORM TRAFFIC

The calculations in Section III have been carried out for uniform traffic conditions. However, in most cases real traffic is nonuniform. Unfortunately and contrary to uniform traffic patterns, simple expressions for the network parameters can not be derived for the case of nonuniform traffic [33]. However, an upper bound for the capacity in MATRIX can be expressed. The maximum network capacity for any traffic pattern in the network is

$$C_{\max}(n, m) = \min(n, m) \cdot nm \tag{33}$$

where the first term describes the maximum transmission capacity of a network node limited by the minimum number of its transmitters or receivers, and the second one equals the number of nodes in the network. The accompanying

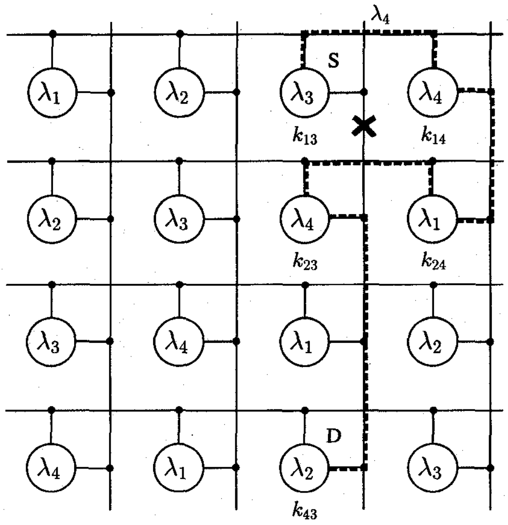


Fig. 15. Survivability routing in the optical domain in a 4 x 4 MATRIX (Example: Failure between the source node $k_{13}$and the destination $k_{43}$ and detour via nodes $k_{14}$ and $k_{24}$).

traffic pattern has the property that any node has a slightly higher traffic with the row and column nodes than with the remaining nodes of the network. Note, that $C_{\max}$ is only insignificantly higher than the network capacity $C(n,m)$ in (26) under uniform traffic conditions. For more realistic nonuniform traffic patterns it has been shown that in multihop networks the throughput per node is reduced by a factor of about 0.3 to 0.5 compared to the values predicted for the uniform traffic case [34].

MATRIX is highly flexible if certain network nodes need additional transmission capacity. The average transmission capacity available for unidirectional traffic between two nodes and for uniformly distributed traffic is with (30) for $nm > \max(n,m)$

$$C_{pp_{\text{norm}}} = \frac{\gamma(n,m)}{nm-1} = \frac{\min(n,m)}{nm}. \tag{34}$$

The transmission capacity $C_{pp}$ between two network nodes can be increased significantly. As stated in the previous section, $\min(n,m)$ optical paths are always available between any pair of network nodes. If a transmitter of the source node transmits packets only to one destination node, a capacity $C_{pp} = 1$ is available between the two nodes. If all available optical paths are used simultaneously and provided that the traffic demand of the other network nodes is low enough, a maximum transmission capacity

$$C_{pp_{\max}} = \min(n,m) \tag{35}$$

results. Therefore, in case of increased traffic demands the transmission capacity available for unidirectional traffic between individual network nodes can be varied in a wide range. Note, that this additional transmission capacity is provided without wavelength translation or detour over the electrical domain.

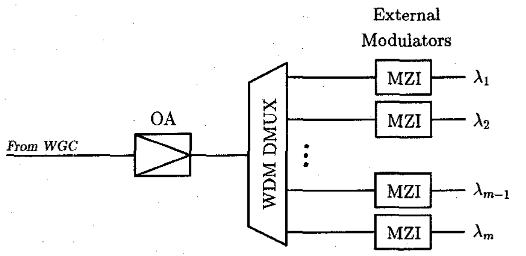


Fig. 16. Multiwavelength transmitter unit supporting centralized carrier supply (OA: optical amplifier, DMUX: demultiplier, MZI: Mach-Zehnder interferometer).

## VII. CENTRALIZED CARRIER SUPPLY

As discussed in Section III, $c$ channels are originating from each network node. Each node disposes of $m$ transmitters working on different wavelengths from the set $\Lambda$ in order to offer newly generated packets to the network. These transmitters can be realized with narrow-linewidth semiconductor lasers being well-suited for long-haul transmission. In this case each node has to provide local control for the laser transmitters in order to guarantee wavelength stability. However, since $\Lambda$ is the same for all network nodes, a centralized carrier supply can be used instead of generating the carriers separately at each node. A set $\Lambda$ of $N$ stable wavelengths exhibiting narrow linewidths and little wavelength drifts is delivered from a common wavelength generation center (WGC) to all network nodes on a carrier distribution fiber via the row and column fiber cables. At each node, a multiwavelength transmitter unit is used to implement the $m$ transmitters as shown in Fig. 16. The device consists of an optical amplifier at the input, followed by a WDM demultiplexer to separate the $N$ wavelengths. The required $m$ wavelengths are modulated by $m$ external modulators.

The centralized carrier supply provides several advantages. First, merely one set $\Lambda$ of stable and narrow linewidth carriers has to be generated at the WGC instead of generating the carriers at each node individually. Hence, the carrier set used in each node is identical, and therefore no wavelength ambiguity appears among the network nodes. Second, the narrow linewidth carriers together with an external modulation represent an excellent solution for dispersion-limited long-haul transmissions. In addition, the multiwavelength transmitter unit can be monolithically integrated in order to be a reliable and simple device. On the other hand, the centralized carrier supply has to fulfill hard requirements with regard to reliability and survivability of the network. However, it is possible to operate a spare WGC in case of a failure.

## VIII. CONCLUSION

MATRIX is a network in which the number of wavelengths, the number of fibers, and the average number of hops are well balanced. As a main property, the all-optical network supports an information exchange with wavelength continuity. Packets are routed according to their wavelengths and with space-division switching. Provided that no network failures occur, buffering and switching of optical packets are needed merely once between source and destination node. This is advantageous since buffering and switching still involve high optical losses. In case of network failures, packets can be routed on alternative all-optical paths through the network. The node structure is uniform and simple in the sense that each node has to process only packets on one wavelength. Finally, it has been shown that a certain regularity of the network architecture may imply significant simplifications in the node structure.

## ACKNOWLEDGMENT

The authors would like to extend their appreciation to P. Leuthold, Director of the Communication Technology Laboratory, ETH Zurich, for the support of this research.

## REFERENCES

[1] M. S. Goodman, H. Kobrinski, M. P. Vecchi, R. M. Bulley, and J. L. Gimlett, "The LambdaNet multiwavelength network: Architecture, applications and demonstration," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 995–1004, Aug. 1990.

[2] A. S. Acampora, "A multichannel multihop local lightwave network," in *Proc. GLOBECOM'87*, Nov. 1987, vol. 3, pp. 1459–1467.

[3] M. G. Hluchyj and M. J. Karol, "ShuffleNet: An application of generalized perfect shuffles to multihop lightwave networks," *J. Lightwave Technol.*, vol. 9, pp. 1386–1397, Oct. 1991.

[4] K. Sivarajan and R. Ramaswami, "Lightwave networks based on de Bruijn graphs," *IEEE/ACM Trans. Networking*, vol. 2, pp. 70–79, Feb. 1994.

[5] H. R van As, "Media access techniques: The evolution toward terabit/s LAN's and MAN's," *Comput. Networks ISDN Syst.*, vol. 26, pp. 603–656, Mar. 1994.

[6] D. W. Smith, "The road to superhighways," in *Proc. ECOC'94*, Sept. 1994, vol. 2, pp. 903–906.

[7] P. E Green, "The future of fiber-optic computer networks," *IEEE Comput.*, vol. 24, pp. 78–87, Sept. 1991.

[8] J. Bannister, M. Gerla, and M. Kovačević, "An all-optical multifiber tree network," *J. Lightwave Technol.*, vol. 11, pp. 997–1008, May/June 1993.

[9] K. Hogari, K. Nakayama, and F. Ashiya, "Design and performance of 2000-fiber cable," *J. Opt. Commun.*, vol. 14, pp. 202–207, Dec. 1993.

[10] N. F. Maxemchuk, "Routing in the Manhattan street network," *IEEE Trans. Commun.*, vol. 35, pp. 503–512, May 1987.

[11] T. D. Todd and A. M. Bignell, "Traffic processing algorithms for the SIGnet metropolitan area network," *IEEE Trans. Commun.*, vol. 40, pp. 568–576, Mar. 1992.

[12] Y. Oie, Y. Sasaki, and H. Miyahara, "RookNet: A switching network for high speed communication," in *Proc. ICC'93*, May 1993, pp. 1225–1230.

[13] H. A. Jäger, "WDM-Gridconnect as a transport structure," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 576–578, May 1995.

[14] T. Gipser and M. S. Kao, "MATRIX: A new network for multiwavelength all-optical transparent information exchange," in *Proc. EFOC&N'94*, June 1994, pp. 146–150.

[15] T. Gipser and M. S. Kao, "MATRIX: An all-optical high-speed network architecture," in *Proc. ECOC'94*, Sept. 1994, vol. 1, pp. 499–502.

[16] Z. Haas and R. D. Gitlin, "Field coding: A high-speed 'almost-all' optical interconnect," in *Proc. 25th Annu. Conf. Inform. Sci. Syst.*, Mar. 1991, pp. 834–839.

[17] H. Nakajima, "High-speed and high-gain optical amplifying photodetection in a semiconductor laser amplifier," *Appl. Phys. Lett.*, vol. 54, pp. 984–986, Mar. 1989.

[18] Z. Haas, "The staggering switch: An electronically controlled packet switch," *J. Lightwave Technol.*, vol. 11, pp. 925–936, May/June 1993.

[19] I. Chlamtac and A. Fumagalli, "An optical switch architecture for Manhattan networks," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 550–559, May 1993.

[20] J. Spring and R. S. Tucker, "Photonic 2 × 2 packet switch with input buffers," *Electron. Lett.*, vol. 29, pp. 284–285, Feb. 1993.

[21] W. Pieper, M. Eiselt, G. Grosskopf, R. Langenhorst, A. Ehrhardt, and H. G. Weber, "Investigation of crosstalk interference in a fiber loop optical buffer," *Electron. Lett.*, vol. 30, pp. 435–436, Mar. 1994.

[22] M. Calzavara, P. Gambini, M. Puleo, B. Bostica, P. Cinato, and E. Vezzoni, "Simultaneous buffering of ATM packets in a multiwavelength optical-fiber-loop memory," in *Proc. OFC'94*, Feb. 1994, pp. 195–96.

[23] Y. Yamada, K. Sasayama, and K. Habara, "BER measurements in frequency-division-multiplexed optical loop buffer at data rate of 2.5 Gbit/s," *Electron. Lett.*, vol. 30, pp. 1869–1871, Oct. 1994.

[24] T. G. Hodgkinson and P. Coppin, "Optical domain frequency translation of a modulated optical carrier," *Electron. Lett.*, vol. 26, pp. 1262–1264, Aug. 1990.

[25] K. Smith, "Methods and applications of all-optical signal regeneration in high speed networks," in *Proc. EFOC&N'95*, June 1995, vol. 1, pp. 34–36.

[26] C. E. Zah, F. J. Favire, B. Pathak, R. Bhat, C. Caneau, P. S. D. Lin, A. S. Gozdz, N. C. Andreadakis, M. A. Koza, and T. P. Lee, "Monolithic integration of multiwavelength compressive-strained multiquantum-well distributed-feedback laser array with star coupler and optical amplifiers," *Electron. Lett.*, vol. 28, pp. 2361–2362, Dec. 1992.

[27] J. B. Soole, A. Scherer, Y. Silberberg, H. P. LeBlanc, N. C. Andreadakis, C. Caneau, and K. R. Poguntke, "Integrated grating multiplexer and PIN Array for high-density wavelength division multiplexed detection at 1.5 $\mu$m," *Electron. Lett.*, vol. 29, pp. 558–560, Mar. 1993.

[28] S. Chandrasekhar, L. M. Lunardi, R. A. Hamm, and G. J. Qua, "Eight-channel *p-i-n*/HBT monolithic receiver array at 2.5 Gb/s per channel for WDM applications," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 1216–1218, Oct. 1994.

[29] A. S. Acampora, "Toward an all-optical telecommunications infrastructure," in *Proc. EFOC&N'94*, June 1994, pp. 1–4.

[30] M. Decina, V. Trecordi, and G. Zanolini, "Throughput and packet loss in deflection routing multichannel-metropolitan area networks," in *Proc. GLOBECOM'91*, Dec. 1991, vol. 2, pp. 1200–1208.

[31] M. G. Hluchyj and M. J. Karol, "Queueing in high-performance packet switching," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 1587–1597, Dec. 1988.

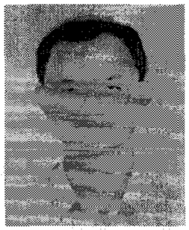
[32] M. To and P. Neusy, "Unavailability analysis of long-haul networks," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 100–109, Jan. 1994.

[33] D. Guo and A. S. Acampora, "Entropy of traffic distribution and multihop lightwave network capacity under nonuniform traffic," in *Proc. ICC'95*, June 1995, vol. 3, pp. 1282–1287.

[34] M. Eisenberg and N. Mehravari, "Performance of the multichannel multihop lightwave network under nonuniform traffic," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 1063–1078, Aug. 1988.

**Thilo Gipser** (S'95) was born in Lupfig, Switzerland on November 21, 1964. He received the Dipl. Ing. degree in electrical engineering in 1989 from the Swiss Federal Institute of Technology in Zurich, Switzerland. In 1993 he finished his postgraduate studies in communications and is currently a Ph.D. candidate at the Communication Technology Laboratory of the Swiss Federal Institute of Technology in Zurich.

Since 1989 he has been with the Communication Technology Laboratory where he designed and tested a coherent optical transmission system. His current research interests include optical switching, high-speed communications, and future network architectures.



**Ming-Seng Kao** (S'89–M'90) was born in Taipei, Taiwan, R.O.C., in 1959. He received the B.S.E.E. degree from the National Taiwan University, in 1982, the M.S. degree in optoelectronics from the National Chiao-Tung University, in 1986, and the Ph.D. degree in Electrical Engineering from the National Taiwan University, in 1990.

From 1986 to 1987 he was an Assistant Researcher at the Telecommunications Laboratories, Chung-Li, Taiwan. In 1990, he joined the faculty of National Chiao-Tung University, Hsinchu, Taiwan, where he is now a Professor of Communication Engineering Department. Between 1993–1994, he was a Visiting Professor at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, where he worked in the area of optical communications. His current interests are in high-speed optical networks and CATV networks.