WDM-Based Local Lightwave Networks Part I: Single-Hop Systems

Next-generation lightwave networks must be completely optical, serve thousands of end users, and support innovative concurrency mechanisms.

Biswanath Mukherjee

BISWANATH MUKHERJEE is Assistant Professor of Computer Science at UC Davis.

This article reviews the characteristics and alternative architectures for single-hop systems; multihop systems are discussed in a companion article [34]. *Reseearch supported in part by the US Air Force Office of Scientific Research (AFOSR) under rant No. 89-0292.*

One can identify three generations of networks based on the underlying physical-level technology employed [1]. Networks built before the emergence of fiber optic technology (i.e., those based on copper-wire or microwave-radio technology) are referred to as first-generation networks. Their examples include Ethernet (and the subsequent IEEE 802.3), IEEE 802.4 token bus, IEEE 802.5 token ring, Cambridge ring, ARPANET, IBM's Systems Network Architecture (SNA), and Digital's Digital Network Architecture (DNA). Second-generation networks employ fiber in traditional architectures. An excellent example of this generation is the upgrade of long-haul trunks in a wide area network (WAN) from copper or microwave-radio to fiber connections. Other examples include new designs such as the Fiber Distributed Data Interface (FDDI) ring network and IEEE 802.6 distributed queue dual bus (DQDB) for local area network/metropolitan area network (LAN/MAN) environments and Broadband Integrated Services Digital Network (BISDN) for WANs. Although some improved performance can be achieved by employing fiber (e.g., higher data rates, lower error rates, and reduced electromagnetic emissions from the cabling), the limitation of this generation is due to the electronic front ends employed at the network nodes. In third-generation networks, fiber is used because of its unique properties. These networks employ totally new approaches to exploit the unique properties of fibers (such as an information-carrying capacity nearly four orders of magnitude greater than peak electronic processing speeds) in order to meet the needs of emerging high-bandwidth applications. Specifically, these networks are "all optical" in nature, in the sense that once information enters the network, it may remain in the optical domain (and may not face any electronic bottlenecks) until it is delivered to its destination. The focus of this article is on the emerging third-generation networks.

The investigation of third-generation networks is driven by new emerging applications [1, 2]. Services such as "fiber to the office" and "fiber to the home" [3] are expected to provide multiple connections of high-definition television (HDTV), digital audio, and other new applications. As universities and research environments continually upgrade their computing facilities, one can envision their need for networking support to interconnect hundreds or thousands of high-performance full-motion color-graphics workstations. Medical imaging, which does not trust image-compression techniques and hence requires the transmission of uncompressed images, is another bandwidth-intensive application. Added to the above are a strong need for high-speed mechanisms to interconnect supercomputers, LANs, and MANs, an increasing proliferation of client-server-paradigm-based implementations in distributed software designs, and a trend towards more graphics-oriented software. With the above applications, it is envisioned that the future end user will generate a sustained bandwidth demand of approximately 1 Gb/s [2, 4]. Second-generation networks, because of their electronic limits, are not expected to be able to support a network of approximately a few hundred or a few thousand such end users.

The above gigabit-per-second-per-user network applications are first appearing in the LAN and MAN environments (e.g., workstation interconnections in university campuses and medical imaging in hospitals). The feasibility of all-optical communication has been demonstrated at LAN/MAN distances [5-15], and they should emerge for WANs soon. Another reason third-generation LANs and MANs are likely to be deployed before corresponding WANs is that it is relatively inexpensive (at least under U.S. regulations) to deploy or to get access to dark fiber within a metropolitan area, while long distance fibers are typically more expensive and harder to access. This article is concerned with third-generation all-optical LANs and MANs, with the expectation that these studies will eventually lead to a better understanding of the issues in, and better designs of, all-optical WANs.

Realizing that the maximum rate at which each end user can access the network is limited by electronic speed (to a few gigabits per second), the key in designing lightwave networks in order to exploit the huge bandwidth is to introduce

 0890-8044/92/$03.00 1992 ©IEEE

concurrency among multiple-user transmissions into the network architectures and protocols. In an all-optical network, concurrency may be provided according to either wavelength or frequency (wavelength division multiple access—WDMA), time slots (time division multiple access—TDMA), or wave shape (spread spectrum or code division multiple access—CDMA). All-optical TDMA and CDMA have been investigated by a number of authors (e.g., [16-19]); however, their basic need to have nodes synchronize to within one time slot (for TDMA) and one chip time (for CDMA) make them relatively less attractive than WDMA. WDMA, on the other hand, employs mostly existing technologies associated with intensity-modulation direct-detection systems [10, 20], and it is the current favorite since all of the end-user equipment need operate only at the bit rate of a wavelength division multiplex (WDM) channel, which can be chosen to be the peak electronic processing speed. This survey will concentrate on WDMA.

Although most of the architectures and protocols discussed here are meant for LANs and MANs, some of them can also be very applicable to situations where the propagation between the end users is negligible, i.e., the end users may be collocated. Therefore, such systems can also be applicable as WDM photonic switches [21, 22] or multiprocessor interconnects [23, 24].

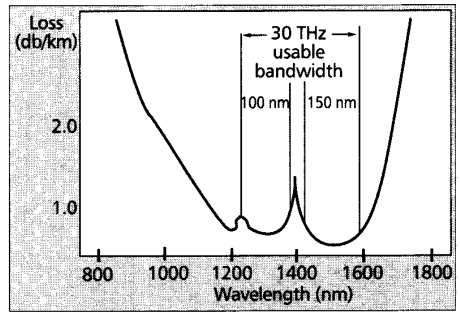
In the next section, we review the characteristics of lightwave technology that facilitate the design of WDM networks, and then discuss how WDM local networks can be built based on the single-hop and multihop approaches. In subsequent sections, various categories of single-hop systems are discussed: experimental systems, systems based on no pretransmission coordination, and systems based on pretransmission coordination, which also require a separate control channel. Finally, the article concludes with a simple classification for single-hop systems.

Characteristics of Lightwave

Technology and WDM Networks The low-loss region of a single-mode fiber is shown in Fig. 1. Note that the fiber has two low-loss wavebands (of approximate widths 100 nm and 150 nm). A direct conversion from wavelengths to frequencies yields an aggregate bandwidth of the low-loss region to be approximately 30 THz. Using a modulation rate of 1 b/Hz, the above bandwidth translates into nearly four orders of magnitude higher than peak electronic data rate of a few bits per second.

Parallel and concurrently operating WDM channels can be derived by having end users transmitting into and receiving from non-overlapping portions of the fiber's low-loss wavelength spectrum. The spacing between adjacent channels can be reduced if the end-user transmitters (lasers) are of good quality, i.e., they are stable and do not drift too far from their nominal operating wavelength range—the minimum channel spacing is limited by crosstalk. By employing distributed feedback (DFB), distributed Bragg reflector (DBR), and other narrow linewidth lasers, a channel spacing of 1 nm or less can be achieved. Systems with 1 nm or lower channel spacing are generally referred to as dense WDM systems [10].

In order to develop an effective WDM network, it may be worthwhile to enable an end user to



■ **Figure 1.** *The low-loss region of an optical fiber.*

transmit into multiple WDM channels, receive from multiple WDM channels, or both. Accordingly, a great amount of effort is being devoted to designing tunable transmitters (lasers) and tunable receivers (filters), which can operate over a wide range of WDM channels and the channel tuning (switching) operations of which can also be performed very fast (ideally in a few nanoseconds or less) [20].

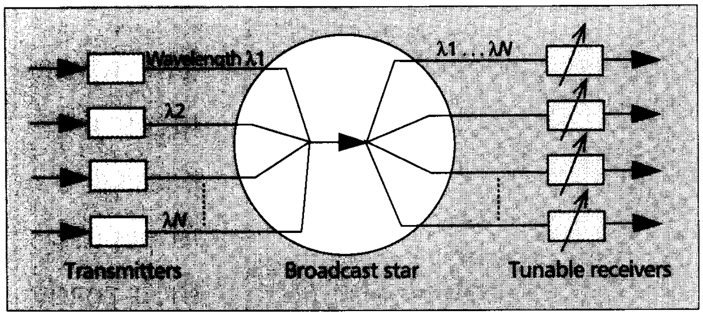
The status of tunable laser technology is documented in [10]. Thermally tuned and external-cavity mechanically tuned lasers require channel switching times equal to a few milliseconds. Hence, they are considered to be very slow in comparison to a typical packet transmission time because, for example, at a 1-Gb/s channel rate, the time to transmit a 1000-bit packet is only 1 μs. Also, thermally tuned lasers have an operating wavelength range of about 1 nm, but mechanical lasers can be tuned over a broader wavelength range (nearly 100 nm) [10]. An external-cavity semiconductor laser with an acousto-optic tunable filter within the cavity has better characteristics than the previous two. A tuning range of 83 nm and a few microseconds tuning time have been demonstrated [25]. Faster tuning speeds (i.e., shorter tuning times) can be achieved by a semiconductor laser, which is tuned by adjusting the injection current in one or more sections of the laser, but the wavelength tuning range of these devices is not very wide. For example, a tunable three-section DBR laser has been used to make a fast-tuning coherent receiver, and measurements demonstrated wavelength switching times of 1.8 ns over a wavelength range of 12.5 nm [26].

The state of the art in tunable filters is also summarized in [10]. The advantages of passive filters are their ready availability and the very fine resolution that they can achieve (especially the Fabry-Perot filter). Their main disadvantages are slow tuning speeds (in the milliseconds range), due to mechanical inertia, and losses. Active and laser-diode-amplifier-based tunable filters can perform faster tuning (in the nanoseconds range), but the number of channels that they can handle is limited. Research in tunable laser and filter technology is progressing [25, 27].

An additional development facilitating WDM technology is the recent progress in erbium-doped optical amplifiers, which provide the extra power budget needed to compensate for the insertion loss of multiplexing components and also for power splitting or tapping [28]. In addition, these optical amplifiers are sufficiently broadband that all optical signals can be amplified simultaneously.

Another new concept in WDM networks is the notion of subcarrier multiplexing [29].

The advantages of passive filters are their ready availability and the very fine resolution that they can achieve.



■ **Figure 2.** *A broadcast-and-select WDM network.*

Although rapid tuning between wavelength channels may not be feasible today, systems can be built that employ rapid tuning between subcarriers within the same wavelength.

Yet another new device developed recently is the Protection-Against-Collision (PAC) switch [22]. The PAC switch simply functions like collision-avoidance switches developed for other networks [30, 31]. A PAC-based network is discussed later in this survey.

Characteristics of WDM Networks

The simplistic nature of WDMA systems opens up a wealth of new applications in point-to-point and multiple access communication systems. In MANs, WDMA can considerably simplify the physical interconnection between the network nodes and permit the construction of centralized hubs. By employing high-speed tunable transmitters (lasers) and receivers (filters), dense WDM may also be used for circuit and packet switching, as well as for wavelength routing. In general, nearly arbitrary (regular or irregular) virtual network topologies can be constructed for any given physical topology of the network. Obviously, the optimality of a virtual topology will be governed by the offered pattern of network loading (the traffic matrix). The problem becomes especially attractive now, since a change in the offered traffic pattern may mean that a new virtual topology is more optimal, and the underlying network architecture's ability to retune its transmitters and receivers to arrive at this new topology (without disrupting the normal network operations, including its connectivity) is becoming an important avenue of research.

A local lightwave network can be constructed by exploiting the capabilities of emerging optical technology, e.g., dense WDM and tunable optical transceivers, as follows. The vast optical bandwidth of a fiber is carved up into smaller-capacity channels, each of which can operate at peak electronic processing speeds (over a small wavelength range) of, say, a few bits per second. By tuning its transmitter(s) to one or more wavelength channels, a node can transmit into those channel(s); similarly, a node can tune its receiver(s) to receive from the appropriate channels. The system can be configured as a broadcast-and- select network in which all of the inputs from various nodes are combined in a WDM passive star coupler, and the mixed optical information is broadcast to all outputs [10, 25] (see Fig. 2). An $N \times N$ star coupler can be considered to consist of an $N \times 1$ combiner followed by a $1 \times N$ splitter; thus, the signal strength incident from any input can be (approximately) equally divided among all the N outputs. The passive star topology is attractive, first because of its logarithmic splitting loss in the coupler (since the splitter portion of the coupler circuit is essentially a binary tree type structure), and second because of no tapping or insertion loss (as in a linear bus) [32]. In addition, the passive property of the optical star coupler is important for network reliability, since no power is needed to operate the coupler; also, it allows fast information relaying without the bottleneck of electro-optic conversion.

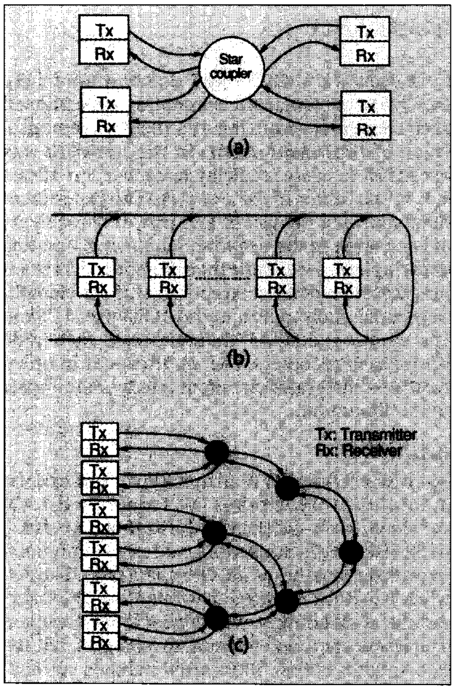
In general, the physical topology, instead of being a star, can be a linear bus or a tree [1, 33] (see Fig. 3); also, the topology of the data-collection part of the network and that of the data-distribution portion need not be identical [33]. The passive star network can typically support a larger number of users than a linear bus topology because power loss and tapping loss in linear buses limit the number of users that can be attached without adding broadband optical amplifiers. However, interest in the linear and tree structures has been revitalized due to the recent development of relatively inexpensive erbium-doped broadband optical amplifiers [28]. In general, the optimal physical topology design problem may be referred to as the cable plant design problem, and a number of studies dealing with this problem to determine that the necessary power budget is satisfied have been reported [33]. From an architectural perspective, given any physical network topology, the fact that the input lasers (transmitters), the output filters (receivers), or both can be made tunable opens up a multitude of possible virtual network configurations.

The present goal is to exploit the properties of lightwave technology to develop terabit networking technology that can support a large number of end users, ranging from a few hundred to a few thousand connection points, with each user receiving peak electronic transmission speeds of a few bits per second.

In a WDM network, each network node is typically equipped with a small number of transmitters and receivers, with some of these transceivers (transmitters or receivers) being dynamically tunable to different wavelengths.

The tunable transceivers are used differently depending on the type of network architecture chosen. In multihop networks, a node is assigned one or more channels to which its transceivers and receivers are to be tuned. These assignments are only rarely changed, usually to improve network performance. Connectivity between any arbitrary pair of nodes is achieved by having all nodes also act as intermediate routing nodes. The intermediate nodes are responsible for routing data among lightwave channels such that the data sent out on one of the sender's transmit channels finally gets to the destination on one of the destination's receive channels, possibly after multihopping through a number of intermediate nodes. A number of different multihop architectures are possible, with a range of operational properties (e.g., ease of routing) and performance characteristics (e.g., average packet delay, number of hops that must be traversed, and efficient use of links).

In single-hop networks, there are no intermediate nodes. As a result, a significant amount of dynamic coordination between nodes is required. For a packet transmission to occur, one of the transmitters of the sending node and one of the



■ **Figure 3.** *Alternative physical topologies for a WDM local lightwave network: (a) star, (b) bus, and (c) tree.*

receivers of the destination node must be tuned to the same wavelength for the duration of the packet's transmission. In the single-hop environment, it is also important that transmitters and receivers be able to tune to different channels quickly, so that packets may be sent or received in quick succession. Currently, the tuning time for transceivers is relatively long in comparison to packet transmission times, and the tunable range of these tranceivers (the number of channels they can scan) is small. Thus, the key challenge in single-hop architectures is to develop protocols for efficiently coordinating the data transmissions.

From a performance perspective, single-hop and multihop networks are currently equally attractive. Accordingly, any survey of WDM networks must properly review both approaches. This article examines single-hop networks. A companion article, to appear in a later issue of *IEEE Network*, will survey multihop lightwave networks [34].

For both single-hop and multihop networks, it is important to keep in mind that any design must be not only simple and implementable (i.e., based on realistic assumptions about the properties of optical components), but also scalable to large user populations. We will focus our attention on such realizable approaches in this survey.

Single-Hop Systems

For a single-hop system to be efficient, the bandwidth allocation among the contending nodes must be dynamically managed. Such systems can be classified into two categories: those employing pretransmission coordination, and those not requiring any pretransmission coordination. Pretransmission coordination systems employ a single shared control channel through which nodes arbitrate their transmission requirements, and the actual data transfers take place through a number of data channels [35-42]. Idle nodes may be required to monitor the control channel. Only during data packet transmission or data packet reception does a node tune its transmitter or its receiver, respectively, to the proper data channel. No such control channel exists in systems that do not require any pretransmission coordination, and arbitration of transmission rights is performed either in a preassigned fashion or through contention-based data transmissions on the regular data channels (e.g., requiring nodes to either transmit on or receive from pre-determined channels) [8-11, 23, 43, 44]. As a result, for a large user population whose size may be time-varying, pretransmission coordination may be the preferred choice, although most experimental demonstrations and prototype WDM systems belong to the non-pretransmission-coordination category.

An alternative classification of WDM systems can be developed based on whether the nodal transceivers are tunable or not. A node's network interface unit (NIU) can employ one of the following four structures:

- Fixed Transmitter(s) and Fixed Receiver(s) (FT-FR)
- Tunable Transmitter(s) and Fixed Receiver(s) (TT-FR)
- Fixed Transmitter(s) and Tunable Receiver(s) (FT-TR)
- Tunable Transmitter(s) and Tunable Receiver(s) (TT- TR)

Fixed transceivers, which can only access some predetermined channels, are readily available in the market, but cost considerations often restrict the installation of a large number of such transceivers at each node. The FT-FR structure is generally suitable for constructing multihop systems in which no dynamic system reconfiguration may be necessary, although a single-hop FT-FR system (LAMBDANET) with a small number of nodes has been demonstrated. FT-FR and TT-FR systems, because they employ fixed receivers, may not require any coordination in control channel selection between two communicating parties, while such coordination is usually necessary in systems employing FT-TR and TT-TR structures. If each node is assigned a different channel under the FT-FR or FT-TR structures, then no channel collisions will occur and simple medium access protocols can be employed, but the maximum number of nodes will be limited by the number of available channels. Systems based on the TT-TR structure are probably the most flexible in accommodating a scalable user population, but they also have to deal with the channel-switching overhead of the transceivers.

In addition, some systems require that a node be equipped with multiple transmitters or receivers. Accordingly, the following general classification for single-hop systems can be developed:

$$\begin{cases} FT^iTT^j\text{-}FR^mTR^n & \textit{No pretransmission coordination} \\ CC\text{-}FT^iTT^j\text{-}FR^mTR^n & \textit{control-channel (CC)-based system} \end{cases}$$

where a node has i fixed transmitters, j tunable transmitters, m fixed receivers, and n tunable receivers. In this classification, the default values of i, j, m, and n, if not specified, will be unity. Also, wherever possible, the number of network nodes, if finite,

■■■■■ For a single-hop system to be efficient, the bandwidth allocation among the contending nodes must be dynamically managed.

Most experimental WDM network prototypes belong to the single-hop category, and they do not employ any control channel for pre-transmission coordination.

will be denoted by M. Thus, Bellcore's LAMBDANET [8] is an FT-FR^M system, since each of the M nodes in the system requires one fixed transmitter and an array of M fixed receivers. The TT and TR portions of the classification are suppressed, since the system requires no tunable transmitter or tunable receiver.

Most experimental WDM network prototypes belong to the single-hop category,[1] and they do not employ any control channel for pretransmission coordination [9 - 11, 43]. Below, first such experimental WDM systems are outlined; then, other non-pretransmission-coordination systems and pretransmission-coordination protocols. While the above mechanisms can generally be employed on any physical topology, we go on to discuss the specific case of a linear-bus physical topology, and protocols that can employ an attempt-and-defer mechanism are outlined.

Experimental WDM Systems

New experimental WDM systems are rapidly being developed. While [10] documented the state of experimental systems in 1990, several new systems, such as IBM's RAINBOW and Columbia's TeraNet, have since been developed. Initial work in this field was done by the British Telecom Research Laboratory (BTRL), whose experiment introduced the concept of a multiwavelength network operating in the broadcast mode and using mechanically tunable filters at each receiver [5]. The AT&T Bell Laboroatories experiment was the first to demonstrate channel spacings on the order of 1 nm [6]. The Heinrich Hertz Institute (HHI) reported the first broadcast star demonstration of video distribution using coherent lightwave technology [7].

Other recent demonstrations include those by AT&T [13, 15] and NTT [14]. Results from a number of demonstrations are reported in [13], initially with two 45-Mb/s channels and employing tunable receivers, later with two 600-Mb/s channels, and followed by two 1.2-Gb/s channels and employing tunable transmitters as well. An experimental system employing six 200-Mb/s channels, spaced by 2.2 Hz, is reported in [15]. The work in [14] demonstrates a system operating with a 5-Hz-spaced (equivalent wavelength spacing 0.04 nm) 16-channel system based on tunable receivers, where each channel can carry 622 Mb/s (enough to accommodate an HDTV channel).

Experimental demonstrations of subcarrier-multiplexing-based systems have also been recently reported [45, 46]. In [45], for example, a system with one control node and two data nodes is demonstrated. Fixed wavelength transmitters are used at each node, and a 1-Mb/s control information is subcarrier-multiplexed at frequencies f_c = 710 MHz, f_1 = 720 MHz, and f_2 = 730 MHz at the central controller, node 1, and node 2 (which are operated at three distinct wavelengths: 1526 nm, 1543 nm, and 1553 nm, respectively). A receiving node employs a tunable filter. It has a passband in the 1526-nm wavelength region to continuously receive from the control channel, and another tunable passband to receive from either of the data nodes (depending on the control information in the f_c subcarrier).

The following subsections elaborate on the various experimental demonstrations and prototypes of single-hop WDM networks.

[1] *An experimental multihop WDM network is TeraNet; it is reviewed in the companion paper on WDM multihop networks [34].*

LAMBDANET

In Bellcore's LAMBDANET demonstration [43], an FT-FR^M system, each transmitter was equipped with a laser transmitting at a fixed wavelength. Each of the wavelengths in the network was broadcast to every receiving node via a broadcast star at the center of the network. The experiment demonstrated the use of an array of N receivers at each node in the network, employing a grating demultiplexer to separate the different optical channels. Recent experiments report that 18 wavelengths were successfully transmitted at 2 Gb/s over 57.5 km. It is mentioned in [43] that, although each node requires N receivers, recent advances in opto-electronic integrated circuits may reduce the impact of this limitation.

RAINBOW

In IBM's RAINBOW project [11], the experimental prototype takes the form of a direct-detection circuit-switched MAN backbone consisting of 32 IBM PS/2s as network nodes, communicating with one another at 200-Mb/s data rates. The network structure is a broadcast-star, but the lasers and filters are housed centrally near the star coupler. The lasers are tuned to fixed wavelengths, but the Fabry-Perot etalon filters are tunable in sub-millisecond switching times, i.e., this is a FT-TR system. An in-band receiver polling mechanism is employed under which each idle receiver is required to continuously scan the various channels to determine if a transmitter wants to communicate with it. The transmitting node continuously transmits a setup request (a packet containing the destination node's address) and has its own receiver tuned to the intended destination's transmitting channel to listen for an acknowledgment from the destination for circuit establishment. The destination node, after reading the setup request, will send such an acknowledgment on its transmitter channel, thereby establishing the circuit. Because of its long setup-acknowledgement delay, this mechanism may not be very suitable for packet-switched traffic, although it would work well for circuit-switched applications with long holding times.

A follow-up report [12] provides further information on the Rainbow-I prototype, which was demonstrated at Telecom '91 in Geneva. The lessons learned from the corresponding prototype development and demonstration are provided in [47]. Information on Rainbow prototypes planned for the future can be found in [12]. It is expected that Rainbow- II(a) will accommodate packet switching with 160 nodes, each operating at 200 Mb/s, by employing rapidly tunable microwave subcarriers. Rainbow-II(b) is expected to support 128 nodes and 1 Gb/s/node, by providing packet switching and a multihop approach. The ultimate goal is Rainbow III, which should support 100 packet-switched nodes, each operating at 1 Gb/s [12].

Fiber-Optic Crossconnect

The goal of the Fiber Optic Crossconnect (FOX) demonstration [9] was to investigate the potential of using fast tunable lasers in a parallel processing environment (with fixed receivers), i.e., this is a TT-FR system. The architecture employed two star networks, one for signals traveling from the processors to the memory banks and the other for information flowing in the reverse direction. Since the utilization of the memory accesses is relatively slow, a binary exponential backoff algorithm was used for resolving contentions, and it was shown to achieve sufficiently good performance. Since the transmitters are tunable, for applications in which data packet transmission times are in the range of 100 ns to 1 μs, transmitter tuning times less than a few tens of nanoseconds will ensure reasonable efficiency.

Other Experimental WDM Systems

In HYPASS [48], an extension of FOX, the receivers were made tunable as well (i.e., a TT-TR system), resulting in vastly improved throughputs. Other recent experiments include BHYPASS, STAR-TRACK, passive photonic loop (PPL), and broadcast video distribution systems. Characteristics of these systems are discussed in [10], and are not repeated here to conserve space.

Other Protocols Based on No Pretransmission Coordination

Several protocols that do not require any pretransmission coordination have recently been reported [22, 23, 43, 44]. Some of these protocols are based on fixed assignment of the channel bandwidth, whereas others are based on demand assignment. We categorize these protocols accordingly in the following subsections.

Fixed Assignment

A simple technique that allows one-hop communication is based on a fixed assignment technique, time division multiplexing (TDM) extended over a multi-channel environment [43]. Each node is equipped with one tunable transmitter and one tunable receiver; hence, these systems are classified as TT-TR systems. The tuning times are assumed to be zero, and the transceiver tuning ranges are the entire set of N available channels. Time is divided into cycles, and it is predetermined at what point in a cycle and over what channel a pair of nodes is allowed to communicate.

For example, for the case of three nodes (numbered 1, 2, and 3) and two channels (numbered Ø and 1), one can formulate the following channel allocation matrix, which indicates a periodic assignment of the channel bandwidth and in which $t = 3i$ where $i = 1, 2, 3, \ldots$:

Channel No.	t	$t+1$	$t+2$
0	(1,2)	(1,3)	(2,1)
1	(2,3)	(3,1)	(3,2)

An entry (i,j) for channel k in slot l means that node i has exclusive permission to transmit a packet to node j over channel k during slot l. The allocation matrix can be generalized for an arbitrary number of nodes M and an arbitrary number of channels N, based on some straightforward properties [43]. This scheme has the usual limitations of any fixed assignment technique, i.e., it is insensitive to the dynamic bandwidth requirements of the network nodes and not easily scalable in terms of the number of nodes. Also, the packet delay at light loads can be significantly high.

The above work has been extended to a versatile time- wavelength assignment algorithm [49] in which node i is equipped with t^i transmitters and r^i receivers, all of which are tunable over all available channels. The scheduling algorithm is mindful of the fact that the transceiver tuning times are non-negligible. Specifically, the algorithm is designed such that, given a traffic demand matrix, it will minimize the tuning times in the schedule, while also minimizing the packet transmission duration.

The work in [49] is further extended in [50] so that users, based on their traffic flow pattems with other users, can be grouped into separate communities. Users within a community are connected by their own local WDM star, but users can communicate with users in other communities (also in a single-hop fashion) via a remote WDM star. Again, given a traffic demand matrix, the algorithm determines the proper time-wavelength schedule.

Partial Fixed Assignment Protocols

The above fixed assignment protocol is too pessimistic because its main goal is to avoid both channel collision and receiver collisions. (A receiver collision occurs when a collision-free data packet transmission cannot be picked up by the intended destination since the destination's receiver may be tuned to some other channel for receiving data from some other source.) However, alternative protocols can be defined in which the channel allocation procedures are less restrictive. A number of such protocols have also been studied in [43].

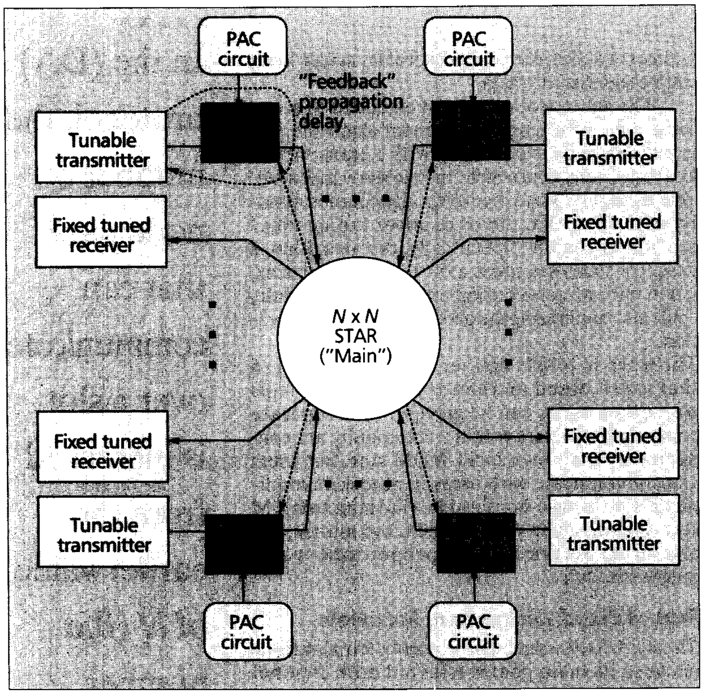
In the Destination Allocation (DA) protocol, the number of node pairs that can communicate over a slot is increased from the earlier value of N (the number of channels) to M. During a slot, a destination is still required to receive from a fixed channel, but more than one source can transmit to it in this slot. Thus, even though receiver collisions are avoided, the possibility of channel collision is introduced. For the three-node two-channel case, an example slot allocation may be the following:

Channel No.	t	$t+1$
0	(1,2) (3,2)	(1,3)
1	(2,3)	(3,1) (3,1)

In the above protocol, the destination receiver is tunable, although it is known *apriori* which channel it will be tuned to during various slots in a cycle. The time taken by the receiver to switch from channel to channel over consecutive slots may be large. Therefore, to avoid this penalty, protocols have been developed in which a node's receiver is always tuned to a fixed channel. Such schemes are discussed below.

A Source Allocation (SA) protocol has also been defined in which the control of access to the channels is further reduced. Now, over a slot duration, N source nodes are allowed to transmit,

> In the (DA) protocol, the number of node pairs that can communicate over a slot is increased from the earlier value of N (the number of channels) to M.



■ **Figure 4.** *Architecture of the PAC optical packet network (the dashed lines indicate energy on the various channels from the "main" star).*

each over a different channel. Since a node can transmit to each of the remaining (M - 1) nodes, the possibility of receiver collisions is introduced. An example periodic slot allocation policy for the three-node two-channel case may now be the following: Finally, an Allocation Free (AF) protocol can be defined in which all source-destination pairs have full rights to transmit on any channel over any slot duration.

Due to the possibility of receiver collisions, the latter two protocols (SA and AF) may not have much practical importance.

Channel No.	t	t=1	t=2
0	(1,2) (1,3)	(1,3) (1,3)	(2,1) (2,3)
1	(2,3) (2,3)	(3,1) (3,2)	(3,2) (3,2)

Random Access Protocols I

This scheme requires that each node be equipped with one tunable transmitter and one fixed receiver (i.e., it is a TT-FR system). The channel on which a node will receive is directly determined by the node's address, e.g., by employing interleaving on the low-order bits of the node's address [51, 52]. The channel a receiver receives from is referred to as that node's home channel.

Two slotted ALOHA protocols were proposed in [23], and it was shown how they outperform the control-channel-based slotted-ALOHA/ALOHA protocol in [35] and its improved version in [36]. (The latter two protocols will be discussed later.) Under one of the protocols, time is slotted on all the channels, and these slots are synchronized across all channels. A slot length equals a packet's transmission time. In the second protocol, each packet is considered to be of L minislots, and time across is synchronized across all channels over minislots. Throughput calculations were performed for these two schemes, and slotting across the entire packet length was found to perform better than minislotting, since the latter scheme increases the vulnerability period of a data packet transmission (just like pure ALOHA has poorer performance than slotted ALOHA). Also, not surprisingly, the maximum throughput on each data channel is found to be $1/e$, the value for the single-channel case.

Random Access Protocols II

A slotted ALOHA and a random TDM protocol were proposed and investigated in [44]. Unlike previous work, these protocols assume limited tuning range, but the tuning speeds are still assumed to be infinitely fast. Again, both of these protocols are based on slotted architectures. Any node, say node i, is equipped with a single tunable transmitter and a number of fixed receivers (i.e., this is a $TT\text{-}FR^x$ system where x is a system parameter). Let $T(i)$ and $R(i)$ be the set of wavelengths over which node i can transmit and receive, respectively. The assignment of transmitters and receivers to various nodes is performed such that the intersection of $T(i)$ and $R(j)$ is always non-null for all i and j, i.e., any two nodes can communicate with one another via one hop. The optimal node/transceiver assignment problem is a challenging but open problem.

Under the slotted ALOHA scheme, if node i wants to transmit to node j, it arbitrarily selects a channel from the set $T(i) \cap R(j)$, and transmits its packet on the selected channel with probability $p(i)$.

The random TDMA scheme operates under the presumption that all network nodes, even though they are distributed, are capable of generating the same random number to perform the arbitration decision in a slot. It is indicated in [44] that this can be done by equipping all nodes with the same random number generator starting with the same seed. Thus, for every slot, and for each channel at a time, the distributed nodes generate the same random number, which indicates the identity of the node with the corresponding transmission right. The problem associated with variable propagation delays between nodes and the hub is not addressed.

Analytical Markov chain models for the slotted ALOHA and random TDMA schemes are formulated to determine the system's delay and throughput performances.

The PAC Optical Network

This is a TT-FR system based on a star configuration [22]. Packet collisions can be avoided by employing protection-against-collision (PAC) switches at each node's interface with the network's star coupler. These collisions are avoided by allowing a node access to a channel only if the channel is available. Also, packets simultaneously accessing the same channel are denied access. The concept is similar to that in collision-avoidance stars [30, 31], except that it is now extended to a multi-channel environment.

The PAC circuit probes the state of the selected channel (i.e., it performs carrier sensing) by using an n-bit burst that precedes the packet. The carrier burst is switched through a second $N \times N$ "control" star coupler, where it is combined with a fraction of all the packets coming out of the "main" star plus all carrier bursts trying to gain access to the "control" star (see Fig. 4). The resulting electrical signal controls the optical switch, which connects the input to the network. The switch is closed only if no energy is detected on the selected channel from other nodes. When two or more nodes try to access the channel simultaneously, all of them detect the "carrier," and their access to the network is blocked. Blocked packets are reflected back to the sender.

Note that the length of the carrier burst n will influence the characteristics of this mechanism. Also, it would probably be preferable to collocate the individual PAC circuits as close to the hub as possible.

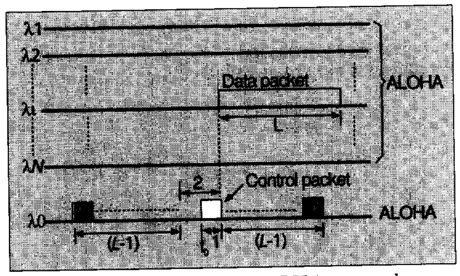
Protocols Based on Pretransmission Coordination

Partial Random Access Protocols

The simplest requirement for single-hop communication is that each node be equipped with a single tunable transmitter and a single tunable receiver, and that the system employ a control channel, i.e., a CC-TT-TR system. Such a class of systems was first studied in [35]. The tuning times were assumed to be zero, and transceivers were assumed to be tunable over the entire wavelength range under consideration. A number of protocols and their performance capabilities were outlined in this work. Access to the control channel was provided via three random access protocols: ALOHA, slotted ALOHA, and carrier sense multiple access (CSMA). ALOHA, CSMA, and an N-server switch mechanism were studied as sub-protocols for the data channels. We elaborate on the specific protocols below.

Assume that time is normalized to the duration of a control packet transmission, which is fixed and is of size one unit. For analytical comparisons, the work in [35] assumed data packets to be of fixed length, of size L units. Also, assume that there are N data channels. A control packet contains three pieces of information—the source address, the destination address, and a data channel wavelength number, which may be chosen at random by the source and on which the corresponding data packet is to be transmitted.

Under the ALOHA/ALOHA protocol, a node transmits a control packet over the control channel at a randomly selected time, after which it immediately transmits the data packet on data channel i, $1 < i < N$, which was specified in its control packet. For a better understanding of the protocol's operation and its performance characteristics, see Fig. 5. Note that the "vulnerable period" of the control packet equals two time units, extending from t_o-L to t_o+L where t_o is the instant the control packet's transmission is started. That is, any other control packet transmitted during the tagged packet's "vulnerable period" would "collide" with (and destroy) the tagged packet. (Since different nodes can be at different distances from the hub, these times are specified relative to the activity seen at the hub.) However,



■ Figure 5. *The ALOHA/ALOHA protocol.*

even if the control packet transmission is successful, the corresponding data packet may still encounter a collision. This may happen if there is another successful control packet transmission over the period t_o-L to t_o+L, and the data channel chosen by that control packet is also i. Using such arguments, the throughput performance of this protocol can be analytically obtained. However, what this and the other protocols in [35] ignore is the possibility of "receiver collisions." Even if the control and data packet transmissions occur without collision, the intended receiver of the destination node might not always be able to read either the control packet or the data packet if it is tuned to some other data channel for receiving data from some other source. For a large or infinite population system, the effect of receiver collisions on the system's performance can be ignored [35, 53].

The slotted-ALOHA/ALOHA protocol is similar, except that access to the control channel is via the slotted-ALOHA protocol. Other schemes outlined in [35] include ALOHA/CSMA, CSMA/ALOHA, and CSMA/N-Server protocols. However, the main limitation of this scheme is that carrier sensing is based on near-immediate feedback, which may not be a practical feature of high-speed systems.

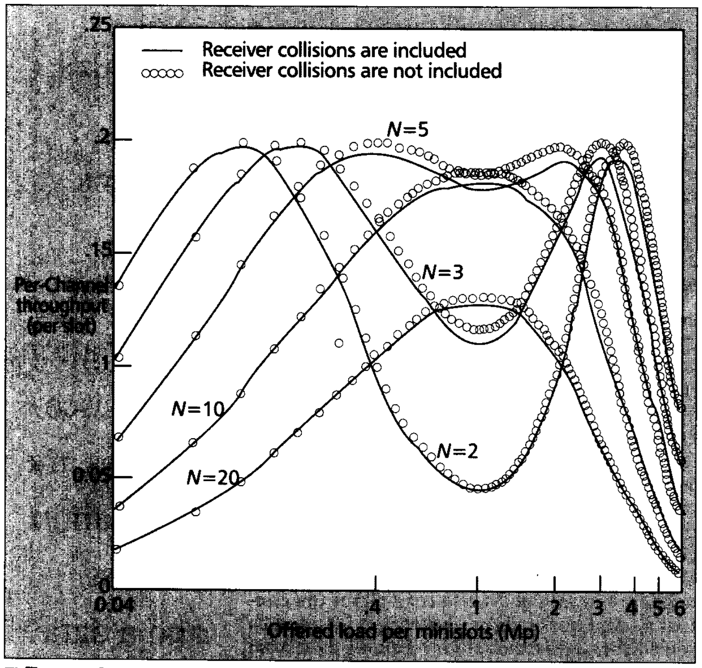
Improved Random Access Protocols

The work in [35] has been extended to obtain improved protocols and performance predictions [36]. The focus in [36] is on realistic protocols that do not require any carrier sensing, since the normalized channel propagation delay in a high-speed environment would exceed unity. Hence, slotted ALOHA for the control channel and ALOHA and the N-server mechanism for the data channels are examined. Another improvement in the protocols is also studied. Specifically, it is required that a node delay its access to a data channel until after it learns that its transmission on the control channel has been successful. As a result, better throughput performance can be achieved, and such results for the improved protocols have been analytically demonstrated in [36].

Bimodal Throughput, Nonmonotonic Delay, and Receiver Collisions—Both the original set of protocols in [35] and the improvements in [36] ignored "receiver collisions," indicating that the probability of receiver collisions is small for large population systems and that they would be taken care of by higher-level protocols. A receiver collision occurs when a source transmits to a destination without any channel collision; however, the destination may be tuned to some other channel receiving information from some other source. The study in [53] first shows that the slotted-

■ ■ ■ ■ ■

A control packet contains three pieces of information: the source address, the destination address, and a data channel wavelength number.



■ **Figure 6** *Bimodal throughput characteristics of the slotted-ALOHA / delayed-ALOHA protocol for* L = *10 slots per data packet and* N = *number of data channels.*

ALOHA/delayed-ALOHA protocol in [36] can have a bimodal throughput characteristic. Basically, if the number of data channels is small, the data channel bandwidth is under-dimensioned, and the data channels are the bottleneck; and if there is a large number of data channels, the control channel's bandwidth is under- dimensioned, and it is the bottleneck. See Fig. 6 for some representative throughput results. The work in [53] finds a useful relationship for optimally dimensioning the available bandwidth (properly selecting the number of data channels) so that neither is the bottleneck. Specifically, it is required that under the slotted-ALOHA/delayed-ALOHA protocol with L-slot data packets, the number of data channels should be given by:

$$N = \left\lfloor \frac{2L-1}{e} \right\rfloor$$

Investigations in [53] reveal that the system has an interesting delay characteristic, i.e., that the mean packet delay is not necessarily monotonic. For example, for some sets of system parameters, such as short backoff period, the mean packet delay can actually be reduced even though the offered load is increased. Of course, this can only happen when the throughput also decreases due to the data channel bandwidth being under-dimensioned. Figure 7 shows some representative delay results from [53].

The work in [53] also studies the slotted-ALOHA/delayed-ALOHA protocol's performance degradation due to receiver collisions for finite population systems, and the corresponding throughput results are also shown in Fig. 6. As expected, the throughput reduction due to receiver collisions is more prominent when the system population is smaller [53].

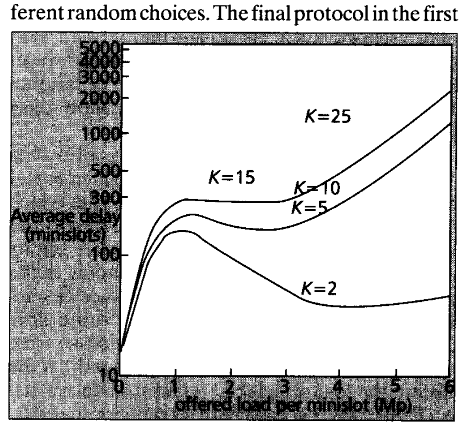
Extended Slotted ALOHA and Reservation ALOHA Protocols

Two sets of slotted ALOHA protocols and a set of Reservation ALOHA protocols are defined in [38], for the same setting as before (one tunable transmitter and one tunable receiver per node). The first set of slotted ALOHA protocols parallel those in [35] in the sense that a data packet is transmitted after a control packet transmission, independent of whether the control packet transmission is successful or not.

One of the protocols can be understood by considering Fig. 8a. A cycle is defined to be a contiguous set of $N + L$ minislots, where N is the number of data channels (and there is an additional control channel) and L is the data packet length. A node that has a data packet to send will arbitrarily choose one of the N control minislots and transmit a control packet in it. If it chooses the ith control minislot in a cycle, it will transmit its data packet in the ith data channel during the same cycle. This fixed assignment of a control minislot to each data channel ensures that if a control packet is successful, then the corresponding data packet will also be successful.

The first set of slotted-ALOHA-based protocols in [38] also includes five variations of the above access mechanism. Note that the mechanism in Fig. 8a is quite wasteful, since during a cycle the data channels are unused during the first N minislots, while the control channel is idle during the last L minislots of the cycle. To improve the protocol's efficiency, the variation in Fig. 8b can be employed. Now, each cycle consists of L minislots, where $L > N$, but the control channel preassignment mechanism spans consecutive cycles. That is, a node transmitting a control packet in the ith control minislot of the Kth cycle will transmit its corresponding data packet in the ith data channel of the $(K+1)$th cycle. Note that now the wastage is reduced to only the last $(L-N)$ minislots on the control channel in each cycle.

Variations of the above two schemes are possible, e.g., one need not preassign a control minislot to a data channel, i.e., a node can select a control minislot and a data channel via two different random choices. The final protocol in the first



■ **Figure 7.** *Non-monotonic delay characteristics of the slotted-ALOHA/delayed-ALOHA protocol for* L = *10 slots per data packet,* N = *3 data channels, zero propagation delay, and different values of the backoff parameter* K.

set of slotted-ALOHA-based protocols in [38] employs asynchronous cycles on the different data channels. The control channel still consists of periods of N minislots followed by $(L-N)$ idle minislots, and data channels are preassigned to the control minislots, but now a data transmission on channel i starts immediately after the control packet transmission in minislot i.

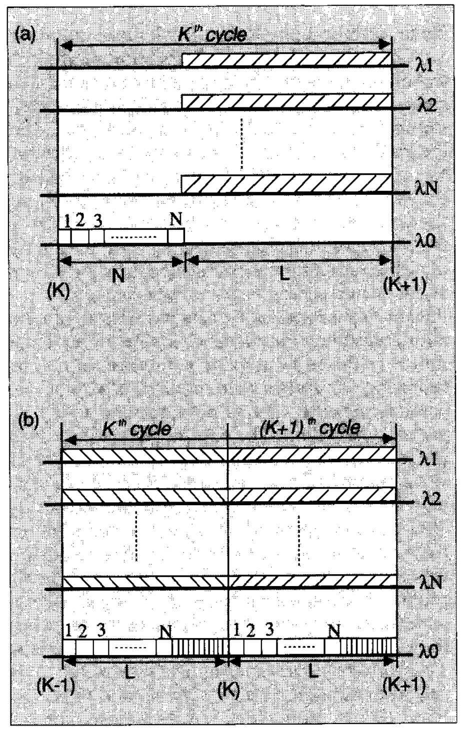
The second set of slotted ALOHA protocols in [38] parallel the improved approaches in [36] (employing delayed feedback)—that a data packet transmission is initiated only after the node learns that its control packet transmission is successful.

In order to accommodate circuit-switched traffic or traffic with long holding times, e.g., file transfers, two Reservation-ALOHA-based protocols are also outlined in [38]. The first builds up on the improved slotted ALOHA technique in Fig. 8a. The data channels are preassigned to control minislots, and a node transmits its data packet, which is the first data packet of a multi-packet message, in the same cycle only if its control packet is successful (as in [36]) (feedback from the channel is assumed to be immediate). If both the control and data packet transmissions succeed, then the node essentially reserves the same data channel in all subsequent cycles until its use of the data channel is completed. It does so by transmitting a jam signal in the corresponding minislot of subsequent cycles, while transmitting its data packets continuously on the corresponding data channel. The second Reservation ALOHA protocol builds up on the improved protocol with asynchronous cycles on the different data channels.

Some further improvements to the above architectures have also been recently reported [21, 39]. Specifically, for the architecture in Fig. 8a, note that instead of wasting the data channels during the first N minislots of a cycle, one can make use of these minislots for control operation as well. Thus, the multi-control-channel protocol in [39] extends the control operation over all channels. Now, one can define communities of interest, and all nodes belonging to a community must tune their receiver during the control part of a cycle to a particular predetermined channel. Any node wishing to send a data packet must therefore transmit its control packet to the destination on the above predefined channel. Data transmission can occur over any of the available data channels (randomly selected by the source before its control packet transmission and so indicated in the control packet). Also, data channel collisions may occur.

Data channel collisions are avoided in the protocols in [21] by requiring that a cycle consist of a control slot (which is further divided into M minislots, one preassigned to each node), an information slot (of X minislots) through which a receiver can notify the senders whether their control packet transmissions are successful or not, followed by a data slot through which collision-free data transmissions occur. This scheme will have value only for small propagation delay systems, and when fast tunable devices become available (because of the multiple channel tuning times involved over a cycle for a data packet transmission).

A multiple-channel-based reservation scheme is also studied by the same authors in [54], where, in a cycle, all channels have a contention slot, followed by a control slot, followed by data slots.



■ **Figure 8.** *The slotted ALOHA protocols.*

Receiver Collision Avoidance (RCA) Protocol

In the previous protocols, the main difficulty in detecting receiver collisions arose due to the simplicity of the systems, i.e., the availability of only one tunable receiver per node to track both the control channel and the data channel activities. However, even for such a simple system, it turns out that by adding some intelligence to the receivers, receiver collisions can be avoided and resolved at the data link (medium access control) layer. Thus, the Receiver Collision Avoidance (RCA) protocol[37] operates under the same basic system parameters as before, i.e., one tunable transmitter and one tunable receiver per node and a contention-based control channel. In addition, the protocol accommodates the fact that transceiver tuning times can be nonzero (of duration T slots, say). For simplicity of presentation, all nodes are assumed to be D slots away from the hub and $N = L$, but these conditions can be generalized. The protocol is briefly outlined below.

Channel selection: Before a control packet is sent, the sender should decide which channel will be used to transmit the corresponding data packet. In order to avoid data channel collision, the RCA protocol proposes a simple and fixed data channel assignment policy. For the case $N = L$, each control slot is numbered 1 through N, periodically, as in a TDM system. Specifically, each control slot is assigned a fixed wavelength, which will be the channel number on which a data packet will be transmitted if the corresponding control packet is successfully sent in that slot. Not only is this assignment scheme simple, but it also guarantees that the corresponding data channel transmission will be collision-free, as in some of the protocols in [38]. The cases $N \neq L$ are treated in [37].

■ ■ ■ ■ ■

For some sets of system parameters, such as short backoff period, the mean packet delay can actually be reduced even though the offered load is increased.

It is observed that the FIFO policy (of giving higher priority to older collided packets) results in better performance.

Node Activity List (NAL): Each node maintains an NAL, which contains information on the control channel history during the most recent $2T+L$ slots. Each entry contains the slot number and a status (Active or Quiet). If the status is Active (which means that a successful control packet is received), the corresponding NAL entry will also contain the source address, the destination address, and the wavelength selected, which are copied from the corresponding control packet. NAL may not be available (or its information outdated) if the local receiver has been receiving on some data channel.

Packet Transmission: Consider a packet generated at Transmitter i and destined for Receiver j. Transmitter i will send out a control packet only if the following conditions hold: node i's NAL does not contain any entry with either node i or node j as a packet destination. The control packet thus transmitted will be received back at node i after 2D slots, during which time node i's receiver must also be on the control channel. Based on the NAL updated by node i's receiver, if a successful control packet to node i (without receiver collision) is received during the $2T+L$ slots prior to the return of the control packet, then a receiver collision is detected and the current transmission procedure has to be aborted and restarted. Otherwise, Transmitter i starts to tune its transmitter to the selected channel at time $t+2D+1$, and the tuning takes T slots, after which L slots are used for data packet transmission, which is followed by another T-slot duration, during which the transmitter tunes back to the control channel.

Packet Reception: The packet reception procedure is quite straightforward and is not elaborated on here to conserve space.

Dynamic Time-Wavelength Division Multiple Access Protocol

A single transmitter and a single receiver per node are the minimal requirements for a single-hop system, but the protocol and the system's performance can be improved by equipping nodes with multiple transceivers. The dynamic time-wavelength division multiple access (DTWDMA) protocol [40] requires that each node be equipped with two transmitters and two receivers: one transmitter and one receiver at each node are always tuned to the control channel, each node has exclusive transmission rights on a data channel to which its other transmitter is always tuned, and the second receiver at each node is tunable over the entire wavelength range, i.e., this is a CC-FT^2-FRTR system. Hence, if there are N nodes, the system requires $N+1$ channels, N for data transmission and the $(N+1)$th for control. Access to the control channel is TDM-based. The system is slotted with slots synchronized over all channels at the passive star (hub). A slot on the control channel consists of N minislots, one for each of the N nodes. Each minislot contains a source address field, a destination field, and an additional field by which the source node can signal the priority of the packet it has queued up for transmission, e.g., the priority information could be the delay the packet would experience from its arrival instant until the time it would reach the hub when it is transmitted. Note that control information is transmitted collision-free, and after transmitting in a control minislot, the node transmits the data packet in the following slot over its own dedicated data channel. By monitoring the control channel over a slot, a node determines if it is to receive any data over the following data slot. If a receiver finds that there are more than one node transmitting data to it over the next data slot, it checks the priority fields of the corresponding minislots and selects the one with highest priority. To receive the data packet, the node simply tunes its receiver to the source node's dedicated transmission channel. Figure 9 elaborates on the protocol's operation.

A novelty of this mechanism is that even though there may be a "collision," in the sense that two or more nodes might have transmitted data packets to the same destination over a data slot duration, exactly one of these transmissions will always be successfully received. Also, this mechanism has an imbedded acknowledgment feature, since all other nodes can learn about successful data packet transmissions by following the same distributed arbitration protocol. In addition, the mechanism supports arbitrary propagation delay between the various nodes and the passive hub. The main limitation of the system is its scalability property, since its control channel is TDM-based and it also requires that each node's transmitter have its own dedicated data channel. An additional issue is that this mechanism requires either infinitely fast receivers or that the receiver tuning time be part of the slot duration, which may lead to a reduction of the protocol's efficiency. Without this limitation, it is shown in [40] that, for a large user population, the peak throughput of the system is $1 - 1/e \approx 0.63$.

A number of extensions to the DTWDMA protocol have been reported [41, 55-57].

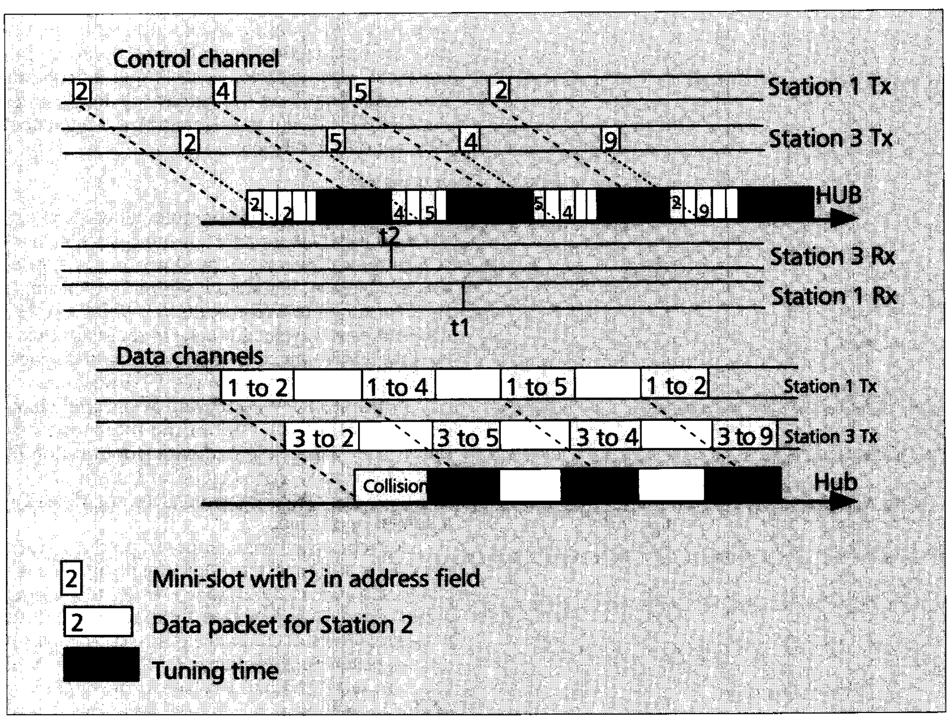
Note that in DTWDMA, when two or more nodes transmit data to the same destination, the destination will read one of the data packets while ignoring the others. The work in [55] proposes the use of an optical delay line to essentially buffer one or more of the collided packets, which would have otherwise been lost. If the destination is not going to receive packets from any source over the next data slot, then it can read a previously buffered packet. The larger the capacity of the optical delay line, the lower the fraction of lost packets. Note that packet loss can still occur if a large number of successive slots have collisions for the same destination. Different policies (first-in first-out—FIFO—vs. last-in first-out—LIFO) exist, depending on the order in which previously collided packets vs. new packets are presented to the destination by the delay line buffer. Simulation results in [55] indicate that with a delay line buffer of 10 or so, the network's throughput can be raised to approximately 0.95 compared to approximately 0.63 for no delay line buffer (the case in [40]). The above results are obtained for the asymptotic case of a large (infinite) user population. Also, it is observed that the FIFO policy (of giving higher priority to older collided packets) results in better performance (higher throughput and lower mean packet delay).

Avoiding the rebroadcast of control packets corresponding to previously collided data packets is also proposed in [56]. Nodes are now made more intelligent so that they can remember information from previous control packet transmissions and participate in a distributed algorithm for efficient scheduling of data packet transmissions. Specifically, each node is required to maintain an

$N \times N$ backlog matrix B, the element b_{ij} of which indicates how many packets at node i are available for transmission to node j. Since all packet arrivals are announced through the broadcast channel, all nodes can maintain identical copies of B locally, by assuming that all nodes are equidistant from the hub. All nodes use B and the same scheduling algorithm to compute the same transmission schedule T, which is an $N \times N$ matrix of binary entries such that $t_{ij} = 1$ indicates that node i should transmit a packet to node j over the next data slot (on node i's dedicated data channel), and $t_{ij} = \emptyset$ otherwise. A proper T matrix must have at most one non-zero element in all rows and one non-zero element in all columns. The work in [56] outlines two algorithms to compute the best possible T. One of them is the Maximum Remaining Sum (MRS) algorithm, which can find a suboptimal T in a small number of operations ($O(N2)$, while the other is the System of Distinct Representative (SDR) algorithm, which can find the optimal schedule T but is compute-intensive (it requires $O(N4)$ operations). Typical numerical examples in [56] indicate that the loss of scheduling efficiency of MRS (as compared with SDR) is not very significant. Numerical examples also indicate that the maximum utilization of a data channel under the improved scheme can approach unity (as compared with 0.63 for the original DTWDMA protocol). However, the scheduling algorithms in [56] also assume that all nodes are equidistant from the WDM star coupler; an extension that eliminates this restriction will be very desirable.

Two additional protocols are outlined in [57], which are also aimed at improving the DTWDMA protocol. Just like in [56], each node maintains a backlog matrix (referred to as S_{ij} in [57]). The first algorithm, called Dynamic Allocation Scheme (DAS), requires that each node execute an identical algorithm based on a common random number generator with the same seed at all nodes. (Recall that this common random number generator requirement was also present in the schemes in [44].) Thus, all transmitters mutually arrive at the same conclusion. First, a transmitter is selected randomly, and among the destinations for which it has queued packets (as determined by identical copies of S_{ij} at all nodes), one receiver is chosen randomly (the same at all nodes). The chosen transmitter will transmit to the chosen receiver over the next slot. The process is repeated to select the other transmitters and the corresponding receivers they will be transmitting to in a similar fashion, except that transmitters and receivers that have already been scheduled are excluded (since each node has one data transmitter and one data receiver). The second protocol in [57], called Hybrid TDM (HTDM), requires that time on the data channels be divided into frames consisting of $M + X$ slots where M is the number of nodes and X is a positive integer. After every $[M/X]$ slots, one slot is left "open," into which a node may transmit to any receiver. Who gets to transmit into such open slots is again determined by the same random-number-based mechanism as in the DAS protocol. Obviously, since HTDM has these open slots, it can provide lower delays than TDM, especially for non-uniform traffic. It is also shown in [57] that HDTM has lower signaling needs than DAS, but DAS also performs close to optimally when the channel propagation delay is zero.

Finally, we note that these schemes can become useful if mechanisms can be developed that can synchronize random- number generators at distributed network locations.

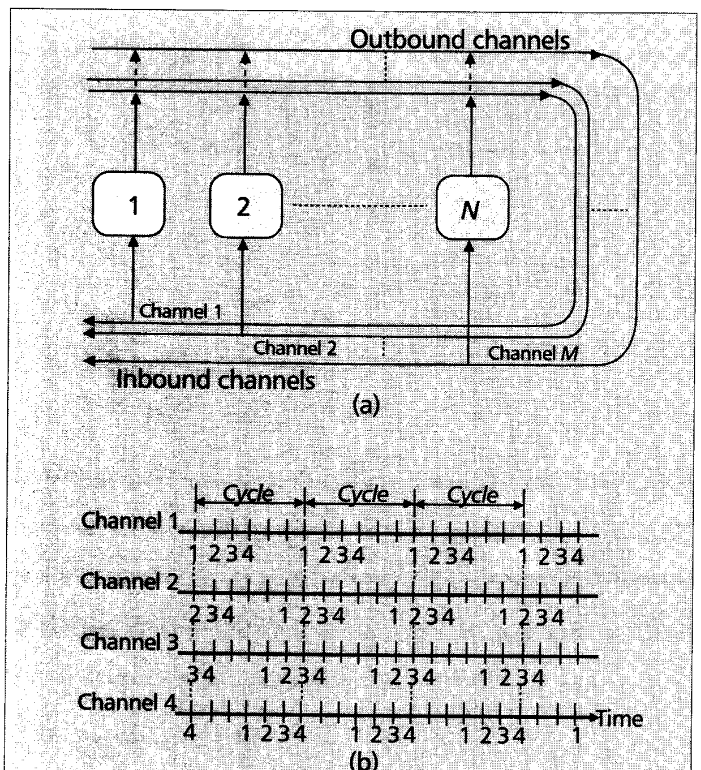


■ **Figure 9.** *The DTWDMA protocol. Note that* t_1 *and* t_2 *are instants when Stations 1 and 3, respectively, learn about the status of their first control packet transmissions to Station 2.*

The TDMA/N-Server Protocol

This protocol employs a TDMA control channel, but it improves upon the DTWDMA protocol [40] by allowing a user population larger than the number of data channels, variable-length data packets, and possibly faster access to the data channel (by not requiring that a node delay its data packet transmission until the beginning of the next control cycle). Each node has its own control slot on the control channel, allocated in TDM fashion, but the number of data channels, M, is limited (possibly less than the number of nodes). Also, each node is equipped with only one tunable transmitter for both control and data packet transmissions. However, as in [40], each node has two receivers, one fixed on the control channel and another tunable over all data channels. Thus, this system can be classified as (CC-TT-FRTR). A control packet consists of the source address (possibly redundant due to the TDM-type control channel), destination address, selected channel number, and length of the data packet (variable). After a control packet transmission, a data packet transmission is initiated on the previously selected data channel α slots later, where $\alpha = \max(t_s, t_r)$; t_s is the transmitter's tuning time, and t_r is the tuning time of the intended receiver.

Collisionless data transmissions are ensured by requiring that each node maintain two status tables. One of them keeps track of the status of the data channels (to eliminate data channel collisions), while the other keeps track of the status of the tunable receiver at each node (in order to avoid receiver collisions on the data channel). The status of these two tables, which are identical at all nodes, is maintained by having the nodes monitor the control packet transmissions on the control channel and update the tables accordingly.



■ **Figure 10.** *The multichannel bus network,: (a) network structure with tunable transmitters and fixed receivers; (b) a cycle in AMTRAC for* N *=4 and* M *= 4.*

Special Case: Linear Bus with Attempt-and-Defer Nodes

The previous protocols were based on a centralized hub architecture, independent of the network's physical topology. However, for a linear-topology physical network (see Fig. 3b), nodes can also employ a multiple access mechanism for sharing their transmissions on a common data channel. In particular, a node on the linear multichannel bus can be equipped with a sense tap, using which it can access a channel via an attempt-and-defer policy. By tuning a sense tap to a channel, a node can determine if there is any activity on that channel from its upstream nodes. If there is activity, the node will defer until the activity has ended. Otherwise, the node can transmit on the channel until its transmission is completed or it senses activity at its sense tap, at which time it aborts its packet transmission, deferring to the upstream transmission.

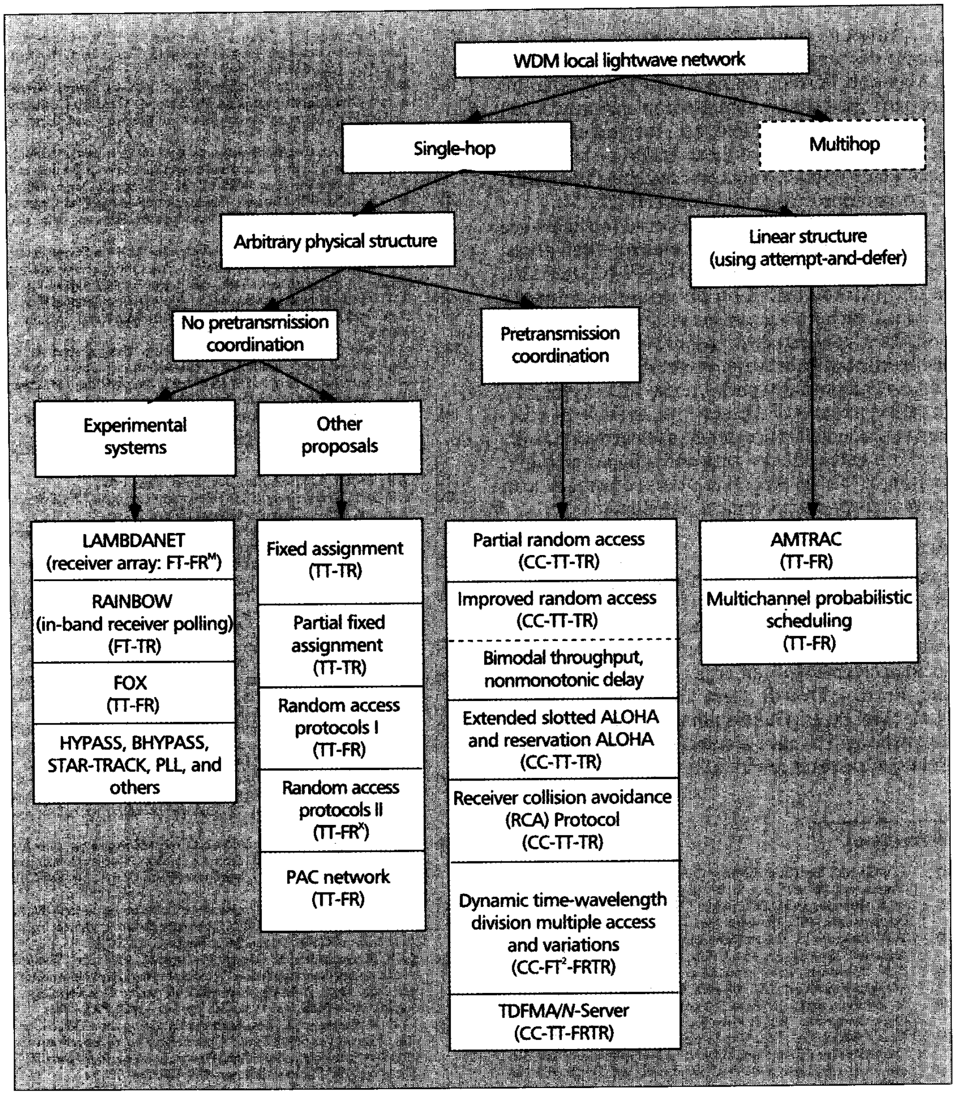
Until recently, protocols based on the above attempt-and-defer mechanism were not considered serious contenders for implementation in a local lightwave network because each tap on the linear bus results in a 3-dB power loss (due to halving of the power at each splitter) so that, with a power margin of a few tens of dBs, only a small number of nodes could be accommodated (without amplification of the optical signal). However, the recent emergence of the erbium-doped fiber amplifier [28] should rekindle interest in the linear- bus-based lightwave networks, because optical signal amplification is no longer a problem. Protocols based on the linear topology, e. g.; AMTRAC [51] and multichannel $p(i)$-persistent [58], are outlined below.

AMTRAC

The AMTRAC approach simultaneously exploits the combined performance advantages of both multichannel and train-oriented protocols. Under this mechanism, a node is equipped with a single receiver tuned to a fixed channel, while a node's transmitter is wavelength-agile in the sense that it can tune itself to any channel, depending on the packet's destination; hence, this is a TT-FR system.

Let there be N nodes and M parallel channels (see Fig. 10a), and let the internode distance be D. Time is slotted with slot duration D, and a cycle is defined to be of length $2(N\text{-}1)$ slots (see Fig. 10b). Transceiver tuning times are assumed negligible.

A cyclic structure is imposed on the system. A node is assigned a scheduling point on each channel in each cycle. If node i wishes to transmit a packet to node j, which happens to receive packets only from channel c (since its receiver is always tuned to this channel), then node i will check to see if channel c is free at the beginning of slot i-1 in a cycle. If it is free, node i will transmit its packet; otherwise, it will wait for the next cycle to repeat the same procedure.



■ **Figure 11.** *Classification of single-hop WDM network architectures.*

The AMTRAC approach simultaneously exploits the combined performance advantages of both multi-channel and train-oriented protocols.

The protocol has subsequently been refined to incorporate multihopping when it is known that single-hopping would result in longer delay [52].

MultiChannel Probabilistic Scheduling

The network structure is the same as before, i. e., a single tunable transmitter and a fixed-tuned receiver at each node. All of the M parallel channels are slotted and synchronized, and the slot length equals the packet transmission time. At the beginning of every slot, out of the M channels, node i chooses a channel c with probability P_{ic}. After probabilistically selecting a channel (say channel c), node i transmits the packet (if any) at the top of its buffer for channel c if the slot arriving on that channel is sensed to be empty.

In [58], the proper transmission probabilities (P_{ic}) for various fairness criteria (e.g., equal average packet delay at each node) have been analytically obtained using both exact and approximate approaches. Also, to keep pace with the fluctuations (if any) of the offered traffic at the various nodes, the transmission probabilities can be dynamically adjusted.

Conclusion

Various contributions made to date on the research and development of WDM-based local lightwave networks were discussed in this article. The characteristics of single-hop systems were studied; various architectures, protocols, and experimental prototypes belonging to single networks were reviewed. Multihop network architectures will be reviewed in a companion paper [34]. Figure 11 provides a summarized classification of single-hop systems.

Much more exciting research and development on single-hop systems is expected. Single-hop systems that can accommodate a large and time-varying user population, and properly utilize the available channel capacity based on wavelength-agile transceivers (with non-zero tuning times and limited tuning ranges) are prime candidates for further study. It is expected that reservation-based protocols as well as protocols that can simultaneously accommodate circuit-switched traffic, narrowband packet traffic (short packets), and bulk data transfers (wideband packet traffic) will be developed.

The recent emergence of the erbium-doped fiber amplifier should rekindle interest in the linear-bus-based lightwave networks.

Although most of the architectures and protocols discussed in this article are meant for LANs and MANs, some of them can also be applied to situations where the propagation between the end users is negligible, e.g., the end users may be collocated. Some of these WDM systems have also been proposed as photonic switching fabrics [21, 22] or multiprocessor interconnects [23, 24].

While this tutorial/survey focused on lightwave networks spanning local and metropolitan areas, work on optical WDM WANs [59-63] has also been initiated based on photonic wavelength routing/switching mechanisms at intermediate nodes [10]. These networks do not need to have a centralized hub as in local lightwave networks. As a result, WDM WANs can employ spatial reuse of wavelengths in different parts of the network, thereby increasing the amount of concurrency in the network. Some work on embedding a ring and a hypercube [59, 60] on WDM WANs has been reported, and more activity on this topic is anticipated.

Acknowledgment

I thank Craig Partridge and the anonymous referees for their constructive criticism, which served to improve the article. I also thank the following individuals for pointing out several references and for their comments on an earlier draft: Imrich Chlamtac, Patrick Dowd, Aura Ganz, Mark Karol, Nader Mehravari, Ganti Sudhakar, Terry Todd, and Zhensheng Zhang. I also acknowledge Subrata Banerjee's assistance with several of the illustrations.

References

[1] P. E. Green, "The Future of Fiber-Optic Computer Networks," *IEEE Comp.*, vol. 24, no. 9, pp. 78-87, Sept. 1991.
[2] M. N. Ransom and D. R. Spears, "Applications of Public Gigabit Networks," *IEEE Network*, Mar. 1992.
[3] P. W. Shumate, Jr., "Optical Fibers Reach into Homes," *IEEE Spectrum*, vol. 26, no. 2, pp. 43-47, Feb. 1989.
[4] B. E. Carpenter, L. H. Landweber, and R. Tirler, "Where Are We with Gigabits?" *IEEE Network*, Guest Editorial, Mar. 1992.
[5] D. B. Payne and J. R. Stem, "Transparent Single Mode Fiber Optical Networks," *J. Lightwave Tech.*, vol. LT-4, pp. 864-869, 1986.
[6] N. A. Olsson *et al.*, "68.3 km Transmission with 1.37 Tbit km/s Capacity Using Wavelength Division Multiplexing of Ten Single-Frequency Lasers at 1.5 μ," *Elect. Lett.*, vol. 21, pp. 105-106, 1985.
[7] E.-J. Bachus *et al.*, "Ten-Channel Coherent Optical Fiber Transmission," *Elect. Lett.*, vol. 22, pp. 1002-1003, 1986.
[8] M. S. Goodman *et al.*, "The LAMBDANET Multiwavelength Network: Architecture, Applications, and Demonstrations," *IEEE J. Sel. Areas in Commun*, vol. 8, pp. 995-1004, Aug. 1990.
[9] E. Arthurs *et al.*, "Multiwavelength Optical Crossconnect for Parallel-Processing Computers," *Elect. Lett.*, vol. 24, pp. 119-120, 1986.
[10] C. A. Brackett, "Dense Wavelength Division Multiplexing Networks: Principles and Applications," *IEEE J. Sel. Areas in Commun.*, vol. 8, pp. 947-964, Aug. 1990.
[11] N. R. Dono *et al.*, "A Wavelength Division Multiple Access Network for Computer Communication," *IEEE J. Sel. Areas in Commun.*, vol. 8, pp. 983-994, Aug. 1990.
[12] F. J. Janneillo, R. Ramaswami, and D. G. Steinberg, "A Prototype Circuit-Switched Multiwavelength Optical Metropolitan-Area Network," submitted to ICC '92.
[13] I. P. Kaminow, "FSK with Direct Detection in Optical Multiple-Access FDM Networks," *IEEE J. Sel. Areas in Commun.*, vol. 6, pp. 1005-1014, Aug. 1990.
[14] H. Toba, K. Oda, K. Nosu, and N. Takato, "Factors Affecting the Design of Optical FDM Information Distribution Systems," *IEEE J. Sel. Areas in Commun.*, vol. 6, pp. 965-972, Aug. 1990.
[15] B. S. Glance and O. Scaramucci, "High-Performance Dense FDM Coherent Optical Network," *IEEE J. Sel. Areas in Commun.*, vol. 6, pp. 1043-1047, Aug. 1990.
[16] P. R. Prucnal, M. A. Santoro, and T. R. Fan, "Spread Spectrum Fiber Optic Local Area Network Using Optical Processing," *J. Lightwave Tech.*, vol. LT-4, pp. 547-554, May 1986.
[17] P. R. Prucnal, M. A. Santoro, and S. K. Sehgal, "Ultrafast All-Optical Synchronous Multiple Access Fiber Networks," *IEEE J. Sel. Areas in Commun.*, vol. SAC-4, pp. 1484-1493, Dec. 1986.
[18] A. M. Weiner, J. P. Heritage, and E. M. Kirschner, "High-Resolution Femtosecond Pulse Shaping," *J. Opt. Soc. Am.*, vol. BS, p. 1563, 1988.
[19] L. A. Bergman, P. L. Chua, J. L. Lambert, and J. M. Morookian, "Optical Protocols for Terabit Networks," High Perf. Network Res. Rep., Info. Sci. Inst., Feb. 1992.
[20] *IEEE J. Sel. Areas in Commun*, Spec. Issue on Dense Wavelength Division Multiplexing Techniques for High Capacity and Multiple Access Communication Systems, vol. 8, no. 6, Aug. 1990.
[21] G. N. M. Sudhakar, N. Georganas, and M. Kavehrad, "A Multi-Channel Optical Star LAN and Its Application as a Broadband Switch," to appear, *Proc. IEEE ICC '92*, Chicago, IL, June 1992.
[22] M. J. Karol and B. Glance, "Performance of the PAC Optical Packet Network," *Proc. IEEE Globecom '91*, Phoenix, Ariz., pp. 1258-1263, Dec. 1991.
[23] P. W. Dowd, "Random Access Protocols for High Speed Interprocessor Communication Based on an Optical Passive Star Topology," *J. Lightwave Tech.*, vol. 9, pp. 799-808, June 1991.
[24] P. W. Dowd, "Wavelength Division Multiple Access Channel Hypercube Processor Interconnection," to appear, *IEEE Trans. on Comp.*, 1992.
[25] G. Coquin, K. W. Cheung, and M. Choy, "Single- and Multiple-Wavelength Operation of Acousto-Optically Tuned Lasers at 1.3 microns," *Proc. 11th IEEE Int'l. Semiconductor Laser Conf.*, Boston, MA, pp. 130-131, 1988.
[26] N. Shimosaka *et al.*, "Photonic Wavelength-Division and Time-Division Hybrid Switching System Utilizing Coherent Optical Detection," *Proc. ECOC '89*, Gothenburg, Sweden, pp. 292-295, Sept. 1989.
[27] *Tech. Digest, OFC '92*, San Jose, Calif., Feb. 1992.
[28] T. E. Bell, "Technology 1991: Telecommunications," *IEEE Spectrum*, vol. 28, no. 1, pp. 44-47, Jan. 1991.
[29] T. E. Darcie, "Subcarrier Multiplexing for Multiple-Access Lightwave Networks," *J. Lightwave Tech.*, vol. LT-5, pp. 1103-1110, 1987.
[30] A. Albanese, "Star Network with Collision Avoidance Circuits," *Bell Sys. Tech. J.*, vol. 62, pp. 631-638, Mar. 1983.
[31] T. Suda and S. Morris, "Tree LANs with Collision Avoidance: Station and Switch Protocols," *Comp. Networks and ISDN Sys.*, vol. 17, pp. 101-110, 1989.
[32] P. S. Henry, "High-Capacity Lightwave Local Area Networks," *IEEE Commun. Mag.*, vol. 27, no. 10, pp. 20-26, Oct. 1989.
[33] M. M. Nassehi, F. A. Tobagi, and M. E. Marhic, "Fiber Optic Configurations for Local Area Networks," *IEEE J. Sel. Areas in Commun.*, vol. SAC-3, pp. 941-949, Nov. 1985.
[34] B. Mukherjee, "Architectures and Protocols for WDM-Based Local Lightwave Networks, Part II: Multihop Systems," (tentative title), to appear, *IEEE Network*, July 1992.
[35] I. M. I. Habbab, M. Kavehrad, and C.-E. W. Sundberg, "Protocols for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology," *J. Lightwave Tech.*, vol. LT-S, pp. 1782-1794, Dec. 1987.
[36] N. Mehravari, "Performance and Protocol Improvements for Very High Speed Optical Fiber Local Area Networks Using a Passive Star Topology," *J. Lightwave Tech.*, vol. 8, pp. 520-530, Apr. 1990.
[37] F. Jia and B. Mukherjee, "The Receiver Collision Avoidance (RCA) Protocol for a Single-Hop WDM Lightwave Network," Div. of Comp. Sci., Univ. of Calif., Davis, Tech. Report, Dec. 1991.
[38] G. N. M. Sudhakar, N. Georganas, and M. Kavehrad, "Slotted Aloha and Reservation Aloha Protocols for Very High-Speed Optical Fiber Local Area Networks Using Passive Star Topology," *J. Lightwave Tech.*, vol. 9, no. 10, pp. 1411-1422, Oct. 1991.
[39] G. N. M. Sudhakar, M. Kavehrad, and N. Georganas, "Multi-Control Channel Very High Speed Optical Fiber Local Area Networks and Their Interconnections Using Passive Star Topology," *Proc. IEEE Globecom '91*, Phoenix, Ariz., pp. 624-628, Dec. 1991.
[40] M.-S. Chen, N. R. Dono, and R. Ramaswami, "A Media-Access Protocol for Packet-Switched Wavelength Division Multi-Access Metropolitan Area Networks," *IEEE J. Sel. Areas in Commun.*, vol. 8, pp. 1048-1057, Aug. 1990.
[41] K. Bogineni and P. W. Dowd, "Collisionless Media Access Protocols for High Speed Communication in Optically Interconnected Parallel Computers," *Proc. SPIE*, vol. 1577, pp. 276-287, Sept. 1991.
[42] P. W. Dowd and K. Bogineni, "Simulation Analysis of a Collisionless Multiple Access Protocol for a Wavelength Division Multiplexed Star-Coupled Configuration," *25th Annual Simulation Symp.*, Orlando, Fla., Apr. 1992.
[43] I. Chlamtac and A. Ganz, "Channel Allocation Protocols in Frequency-Time Controlled High Speed Networks," *IEEE Trans. Commun.*, vol. 36, pp. 430-440, Apr. 1988.
[44] A. Ganz and Z. Koren, "WDM Passive Star Protocols and Perfommance Analysis," *Proc. IEEE INFOCOM '91*, Bal Harbor, Fla., pp. 991-1000, Apr. 1991.
[45] M. W. Maeda *et al.*, "Wavelength-Division Multiple Access Network Based on Centralized Common-Wavelength Control," *Tech. Digest, OFC '92*, San Jose, Calif., pp. 85, Feb. 1992.
[46] A. Budman *et al.*, "Multi-Gigabit Optical Packet Switch for Self-Routing Networks with Subcarrier Addressing," *Tech. Digest, OFC '92*, San Jose, Calif., pp. 90-91, Feb. 1992.
[47] P. E. Green. "An All-Optical Computer Network: Lessons Learned," *IEEE Network*, Mar. 1992.
[48] E. Arthurs, M. S. Goodman, H. Kobrinski, and M. P. Vecchi, "HYPASS: An Optoelectronic Hybrid Packet-Switching System," *IEEE J. Sel. Areas in Commun.*, vol. 6, pp. 1500-1510, Dec. 1988.
[49] A. Ganz and Y. Gao, "A Time-Wavelength Assignment Algorithm for a WDM Star Network," to appear, *Proc. IEEE INFOCOM '92*, Florence, Italy, May 1992.
[50] A. Ganz and Y. Gao, "Traffic Scheduling in Multiple WDM Star Systems," to appear, *Proc. IEEE ICC '92*, Chicago, Ill., June 1992.
[51] I. Chlamtac and A. Ganz, "A Multibus Train Communication Architecture for High-Speed Fiber Optic Networks," *IEEE J. Sel. Areas in Commun.*, vol. SAC-6, pp. 903-912, July 1988.

[52] I. Chlamtac and A. Ganz, "Towards Alternative High Speed Network Concepts: The SWIFT Architecture," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 431-439, Apr. 1990.
[53] B. Mukherjee and F. Jia, "Bimodal Throughput Characteristics of a Single-Hop WDM System," Div. of Comp. Sci., Univ. of Calif., Davis, Tech. Rep. No. 9133, Aug. 1991.
[54] G. N. M. Sudhakar, M. Kavehrad, and N. Georganas, "A Simple Contention-Based Reservation Scheme for Optical Star LAN Using a Passive Bus Topology," to appear, *Proc. Photonics '92*.
[55] I. Chlamtac and A. Fumagalli, "QUADRO-Stars: High Performance Optical WDM Star Networks," *Proc. IEEE Globecom '91*, Phoenix, Ariz., pp. 1224-1229, Dec. 1991.
[56] M. Chen and T.-S. Yum, "A Conflict Protocol for Optical WDM Networks," *Proc. IEEE Globecom '91*, Phoenix, Ariz., pp. 1276-1291, Dec. 1991.
[57] R. Chipalkatti, Z. Zhang, and A. S. Acampora, "High-Speed Communication Protocols for Optical Star Networks Using WDM," to appear, *Proc. IEEE INFOCOM '92*, Florence, Italy.
[58] S. Banerjee and B. Mukherjee, "An Efficient and Fair Protocol for a Multichannel Lightwave Network," *Tech. Digest, OFC '92*, San Jose, Calif., p. 183, Feb. 1992.
[59] I. Chlamtac, A. Ganz, and G. Karmi, "Lightnet: Lightpath Based Solutions for Wide Bandwidth WANs," *Proc. IEEE INFOCOM '90*, San Francisco, Calif., pp. 1014-1021, June 1990.
[60] I. Chlamtac, A. Ganz, and G. Kanni, "Transport Optimization in Broadband Networks," *Proc. IEEE INFOCOM '91*, Bal Harbour, Fla., pp. 49-58, Apr. 1991.
[61] G. R. Hill, "Wavelength Domain Optical Network Techniques," *Proc. IEEE*, vol. 77, pp. 121-132, Jan. 1990.
[62] T. E. Stern, "Linear Lightwave Networks: How Far Can They Go?" *Proc. IEEE Globecom '90*, San Diego, Calif., pp. 1866-1872, Dec. 1990.
[63] T. E. Stem, K. Bala, and S. Jiang, "Purely Optical Wide Area Networks," *Tech. Digest, OFC '92*, San Jose, Calif., pp. 94, Feb. 1992.

Biography

Biswanath Mukherjee received Bachelors and doctoral degrees from IIT Kharagpur (India) and Univ. of Washington, Seattle, respectively. Since 1987, he has been Assistant Professor of Computer Science at UC Davis. He has served on the Technical Program Committee of several INFOCOM conferences, IEEE Globecom '92, and 1992 IEEE/LEOS/COMSOC Summer Topical Meeting on Optical Multiple Access Networks. He is Co-organizer /Co-chair of the NSA/USAF/DOE Workshop on "Future Directions in Computer Security and Anomaly Detection."