HONET: AN INTEGRATED SERVICES WAVELENGTH DIVISION OPTICAL NETWORK

M. Kovačević

UNIVERSITY OF CALIFORNIA

Los Angeles

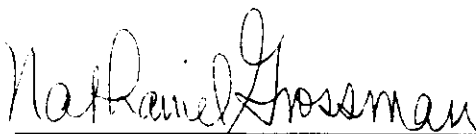# HONET: An Integrated Services Wavelength Division Optical Network

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Computer Science

by

Milan Mihailo Kovačević
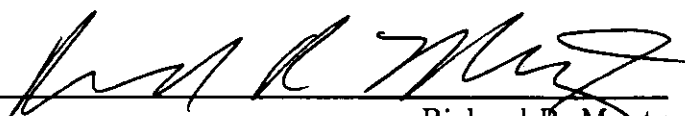
1993

The dissertation of Milan Mihailo Kovačević is approved.
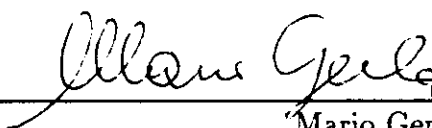
_____
Nathaniel Grossman

_____
Nicholas Bambos

_____
Richard R. Muntz

_____
Leonard Kleinrock

_____
Mario Gerla, Committee Chair

University of California, Los Angeles

1993

In memory of my mother Olga

and to my father Mihailo

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my appreciation to my Doctoral Committee, which consists of Professors Mario Gerla, Leonard Kleinrock, Richard Muntz, Nicholas Bambos and Nathaniel Grossman, for their interest and efforts in reviewing this work. I am particularly grateful to my advisor Mario Gerla for his stimulating and knowledgeable guidance during the course of my research. Mario helped me originate many of the ideas developed in this work through interesting and insightful discussions. I am also grateful to Professor Leonard Kleinrock, whose outstanding lectures attracted me to the area of computer networks.

Joseph Bannister of The Aerospace Corporation deserves credit for his participation in this research. His contributions and suggestions have been instrumental in improving the quality of my research. I am also grateful to Joe for his patience and effort in proofreading the manuscript.

I owe much to all of the fellow graduate students, staff and faculty for making my stay at UCLA enjoyable and stimulating. My officemates Ying-Dar Lin, Tsung-Yuan Tai, Siddhartha Devadhar, Mitchell Tsai and Chun-Hung Lin have been interesting and supportive colleagues. I would especially like to thank Verra Morgan and Professor Miloš Ercegovac for everything they have done for me throughout all these years.

Finally, the deepest gratitude goes to my parents Olga and Mihailo, and my love Milanka for their encouragement, patience and support.

VITA

1960	Born in Belgrade, Yugoslavia.
1985	Electrical Engineering Diploma, School of Electrical Engineering, University of Belgrade, Yugoslavia.
1985–1987	Research Engineer, Dept. of Electronics and Computers, School of Electrical Engineering, University of Belgrade, Yugoslavia.
1987	M.Sc in Computer Engineering, School of Electrical Engineering, University of Belgrade, Yugoslavia.
1987-1988	Research Assistant, UCLA Computer Science Department, Los Angeles, California.
1988	Computer Consultant, Jet Propulsion Laboratory, Pasadena, California.
1989	Summer Resarcher, IBM Corporation, OSI Network Development Division, Palo Alto, California.
1990-1992	Teaching Associate/ Teaching Fellow, UCLA Computer Science Department, Los Angeles, California.
1990-1993	Research Assistant, UCLA Computer Science Department, Los Angeles, California.
1992-1993	UCLA Chancellor's Dissertation Year Fellowship Award
1993	M.Sc. in Computer Science, University of California, Los Angeles.

PUBLICATIONS AND PRESENTATIONS

J. Bannister, M. Gerla and M. Kovačević, An All-Optical Multifiber Tree Network, *Technical Report CSD-920026*, Computer Science Department, University of California, Los Angeles, June 1992.

J. Bannister, M. Gerla and M. Kovačević, An All-Optical Multifiber Tree Network, *IEEE LEOS Summer Topical Meeting on Optical Multiple Access Networks*, Santa Barbara, California, August 1992.

J. Bannister, M. Gerla and M. Kovačević, An All-Optical Multifiber Tree Network, *Proc. of IEEE INFOCOM '93*, San Francisco, California, vol. 1, pp. 282-292, March 1993.

J. Bannister, M. Gerla and M. Kovačević, An All-Optical Multifiber Tree Network, to appear in *IEEE/OSA Journal of Lightwave Technology*, special issue on Broadband Optical Networks, 1993.

M. Bojović, Z. Konstantinović and M. Kovačević, Operating System Kernel for Fault-Tolerant Multiprocessor System, *Proc. of the ISMM Cairo International Conference*, March 1987.

A. Borella, F. Chiaraluce, F. Meschini, M. Gerla and M. Kovačević, An Analytical Approach to the Performance Evaluation of MONETs, *Proc. of International Conference on Computer Communications and Networks (ICCCN)*, pp. 257-261, San Diego, California, June 1993.

M. Gerla, M. Kovačević and J. Bannister, Multilevel Optical Networks, *Proc. of IEEE ICC '92*, Chicago, Illinois, pp. 1168-1172, June 1992.

M. Gerla, M. Kovačević and J. Bannister, Optical Tree Topologies: Access Control and Wavelength Assignment, to appear in *Computer Networks and ISDN Systems*, **1993**.

M. Kovačević, Z. Konstantinović, M. Bojović and V. Majetić, MIP - The Software System for Interactive Control and Testing of Programs in a Multiprocessor Environment, *Proc. of the 31. Yugoslav Conference - ETAN*, Bled, June 1987

M. Kovačević, Basic Software Support for a Fault-Tolerant Multimicroprocessor System, *M.Sc Thesis*, School of Electrical Engineering, University of Belgrade,

Yugoslavia, June 1987.

M. Kovačević and M. Gerla, Efficient Multiaccess Scheme for Linear Lightwave Networks, *Proc. of SPIE International Symposium OE/Fibers '91 - Advanced Fiber Communication Technologies*, Boston, Massachusetts, vol. 1579, pp. 74-83, September 1991.

M. Kovačević and M. Gerla, Rooted Routing in Linear Lightwave Networks, *Proc. of IEEE INFOCOM '92*, Florence, Italy, vol. 1, pp. 39-48, May 1992.

M. Kovačević and M. Gerla, HONET - A New Hybrid Optical Network Architecture, *Proc. of International Conference on Computer Communications and Networks (ICCCN)*, pp. 92-96, San Diego, California, June 1992.

M. Kovačević and M. Gerla, Analysis of a T/WDMA Scheme with Subframe Tuning, *Proc. of IEEE ICC '93*, Geneva, Switzerland, vol. 2, pp. 1239-1244, May 1993.

M. Kovačević, M. Gerla and J. Bannister, Time and Wavelength Division Multiaccess with Acoustooptic Tunable Filters, *Journal of Fiber and Integrated Optics*, special issue on Optical Networking, vol. 12, no. 2, pp. 113-132, 1993.

M. Kovačević, M. Gerla and J. Bannister, On the Performance of Shared-Channel Multihop Lightwave Networks, accepted for *6th IEEE Workshop on Local and Metropolitan Area Networks*, October 1993.

M. Kovačević and M. Gerla, A New Optical Signal Routing Scheme for Linear Lightwave Networks, submitted to *IEEE Transactions on Communications*.

Z. Konstantinović, M. Bojović and M. Kovačević, Software Support for a Fault-Tolerant System - FTS16, *Proc. of the 30. Yugoslav Conference - ETAN*, Herceg Novi, June 1986.

Z. Konstantinović, M. Cvetinović, M. Bojović and M. Kovačević, Reliable Multiprocessor Architecture - FTS16, *Proc. of the ISMM Cairo International Conference*, March 1987.

V. Majetić, M. Bojović, Z. Konstantinović and M. Kovačević, Interactive Program for Development and Testing of the Basic System Support for the FTS-16 Multiprocessor System, *Proc. of the 31. Yugoslav Conference - ETAN*, Bled, June 1987.

ABSTRACT OF THE DISSERTATION

HONET: An Integrated Services Wavelength Division Optical Network

by

Milan Mihailo Kovačević

Doctor of Philosophy in Computer Science

University of California, Los Angeles, 1993

Professor Mario Gerla, Chair

We investigate a class of high-speed fiber-optic local and metropolitan area network architectures that can support a wide range of applications, namely: datagram (e.g., file transfer); real-time, connection oriented communications (e.g., voice and video), and; multicast/broadcast (e.g., video broadcasting and conferencing). Successful results have been reported by using WDM (wavelength division multiplexing) in two types of optical network architectures: single-hop (i.e., direct path between each source/destination pair), and multihop (store-and-forward processing at intermediate nodes along the path). It has been observed that, while neither of these schemes can adequately satisfy all requirements, yet single-hop and multihop networks have positive and negative properties that tend to complement and compensate each other. We exploit this situation by proposing a hybrid architecture - HONET (Hybrid Optical Network) which combines the concepts of both schemes retaining their advantages and avoiding their limitations. HONET uses a multihop network for packet service and a single-hop

network for circuit-switched service.

We first study possible virtual topologies for HONET's multihop network that can be established with channel sharing using a time division multiple access technique (TDMA) and determine the optimal degree of channel sharing which maximizes throughput. We then discuss some of the possible HONET implementations and compare them with other related optical network architectures.

Next, we propose a novel time and wavelength division multiaccess scheme (T/WDMA) that supports circuit-switched traffic over HONET's single-hop network. The scheme employs subframe tuning and pipelining in order to allow an implementation with off-the-shelf, relatively slow tunable receivers. Thus, tuning time can be much longer than the TDM slot reserved for the connection.

We then consider implementation of HONET over the Linear Lightwave Network (LLN) infrastructure. We identify a synchronization problem in LLN which complicates the implementation of slotted TDMA and T/WDMA networks. We propose a solution which permits us to implement HONET over LLN. The solution is based on using a new optical signal routing scheme called *rooted routing* instead of the originally proposed, shortest path routing. The impact of rooted routing on power loss and propagation delay is analyzed and an approach for minimizing power loss in LLNs with rooted routing is presented.

In order to increase scalability and capacity of HONET, we also propose multilevel and multifiber extensions of the basic HONET. The multilevel architecture is hierarchically organized. It exploits traffic locality and wavelength reuse. The multifiber HONET utilizes a multifiber infrastructure, such as multifiber tree or multifiber LLN.

CHAPTER 1

Introduction

The data rates required to meet users' communication needs in a local and metropolitan area environment have exploded over the last twenty years. Terminals are being replaced by workstations which are able to integrate sound, image and even full motion video with textual information. New multimedia applications will require bandwidths far exceeding that of today's 10 megabit per second (Mb/s) local area and 100 Mb/s metropolitan area networks. These applications will increase users' communication requirements to hundreds of megabytes per second, and possibly to rates in excess of 1 gigabit per second (Gb/s). The applications to be supported by future networks have very diverse characteristics. Some of them have very large bandwidth requirements, like supercomputer interconnections, supercomputer visualization [Vaz91] and high resolution uncompressed medical images [Gre93]. On the other hand, there are applications that require much smaller bandwidth, such as voice (64 kb/s) or video (1.5 Mb/s for compressed NTSC video [Gre93]). Some applications have very strict real-time requirements, (e.g., voice and video), which cannot tolerate large delay fluctuations. As a contrast to these, there are applications that can tolerate relatively large delays (e.g. file transfer, electronic mail). The applications can also differ in whether they require communication between a pair of stations (unicast) or they require one-to-many (multicast) or one-to-all (broadcast) communications. The examples of the last two are video distribution and voice and/or video conferencing.

Designing a network that can serve all the above mentioned applications poses a real challenge. The introduction of an optical fiber as a new transmission media makes building such a network a realistic possibility. What makes fiber optic networks so attractive is the very broad bandwidth of an optical fiber which cannot be matched by any other transmission media. It is estimated that the capacity of the low-attenuation bands at 1.3 μm and 1.5 μm is equal to 30 terahertz. Besides large capacity, the optical fiber is superior to other media due to very low attenuation losses (0.2 dB/km) and very low bit error rate (10^{-9} errors per bit compared to 10^{-5} errors per bit for copper wire). In addition to this, the technology evolved to the point where optical fiber become more cost-effective

than alternative copper links for high speed transmissions.

Networks that use optical fiber as a communication media such as FDDI [Ros86] and DQDB [DQD88] are already in use today. However, these are traditional network architectures where fiber merely replaces copper wire. The total capacity in these networks is time-shared among many users and each user operates at the aggregate bit rate at which the network operates. The total network capacity is limited by the speed of users' electrooptic interface and it is typically several hundred megabits per second. (it can be a few gigabits per second at most). Thus, these architectures cannot be extended to terabit per second capacities because of the electrooptic bottleneck. To overcome this bottleneck, we must look for different network architectures.

The most popular approach to realize terabit per second local and metropolitan area networks appears to be to use a passive optical network (PON) architecture where no electronic processing and electro-optical conversion is performed within the network. In PONs, all the users are connected to the common optical medium, which can be viewed as a "mass of glass". The optical spectrum is then divided into many different channels, each corresponding to different frequency (i.e., wavelength). This approach, called wavelength division multiplexing (WDM) [Bra90] makes possible the realization of a terabit per second network using several hundred to a thousand of channels at different wavelengths, each operating at moderate speeds of say, a few gigabits per second.

Next, we provide an overview of PON topologies, WDM architectures and access schemes which have been previously proposed for local and metropolitan area networks.

1.1 PON topologies

Passive optical networks can be implemented using star, bus, tree or mesh physical topology. We review next these options.

1.1.1 Star

The most common implementation of a passive optical network is based on the star topology as shown in Figure 1.1. The signals coming from all stations are combined at the star coupler and broadcast to all the outputs. Thus, each station receives all the signals and then selects the signal intended to it. The main limitation of this, as well as the other passive optical network topologies lies in optical power losses due to power splitting. Optical power transmitted by a sta-
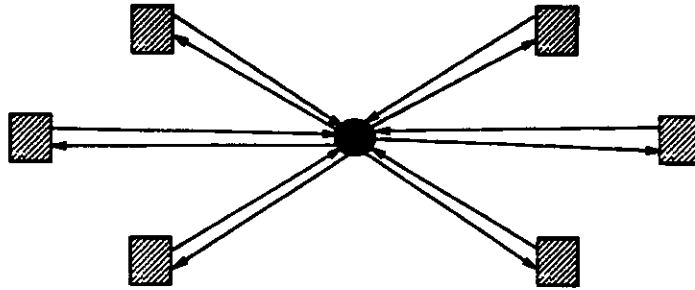
Figure 1.1: Broadcast star topology

tion is split evenly to all the stations in the network. The maximum ratio of the signal sent by a transmitter and the signal received by a receiver which ensures sufficient signal-to-noise ratio at the receiver for successful signal detection represents *power margin*. It is possible to support a data rate of 1 Gb/s with 10^{-9} errors per bit using a silicon avalanche photodiode with transmit power of 0 dBm and receiver sensitivity -40 dBm (which gives a power margin of 40 dB or 10000) [Hos90]. We need to ensure that the maximum power loss between any two stations, defined as the *power budget*, does not exceed the power margin. The power loss due to N-way power splitting can be expressed as $10\log_{10} N$ [dB] where N is the number of station in the network. If the fused biconical taper process is used for fiber coupling [Tek90], the excess losses are negligible. Assuming that maximum power margin is 40 dB, it would be possible to support up to 10,000 stations. It is difficult, however, to couple a large number of fibers using this technology. Dragone's Fourier optics coupler [Dra89] or modular star coupler would be more appropriate for a large network. The modular star coupler is illustrated in Figure 1.2. It is built by interconnecting several stages of 2×2 couplers. The number of stages in the modular star coupler is $\log_2 K$, where K is the number of ports.

1.1.2 Tree

In order to establish a star topology, it is necessary to install a link from each station to the central node (i.e. star coupler) in both directions. This can be, however, an expensive solution with respect to fiber layout cost if stations are distributed over a large geographical area. It is intuitive that a tree topology leads to lower-cost layouts than a star. This intuition is confirmed by quantitative results based on the comparison of a variety of topologies for a vast sample of station locations [BFG90]. Those results clearly show the cost advantage of trees over stars. Figure 1.3 illustrates the tree topology [GF88]. In a tree network

3
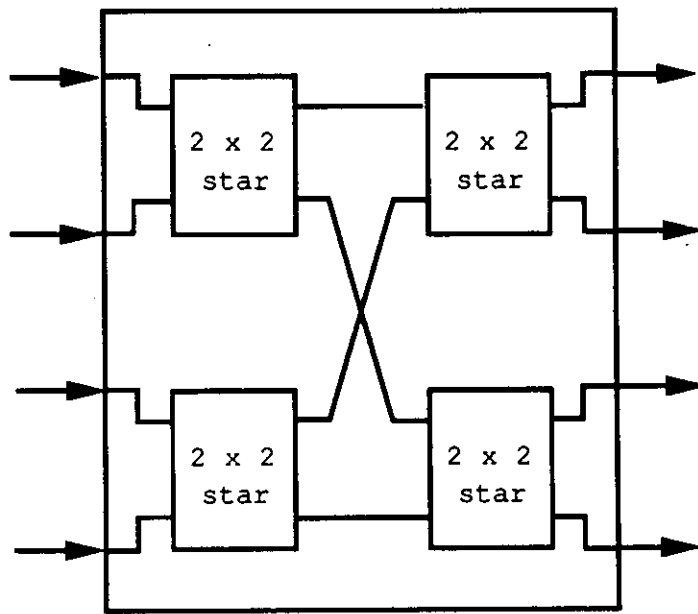
Figure 1.2: 4 × 4 modular star coupler
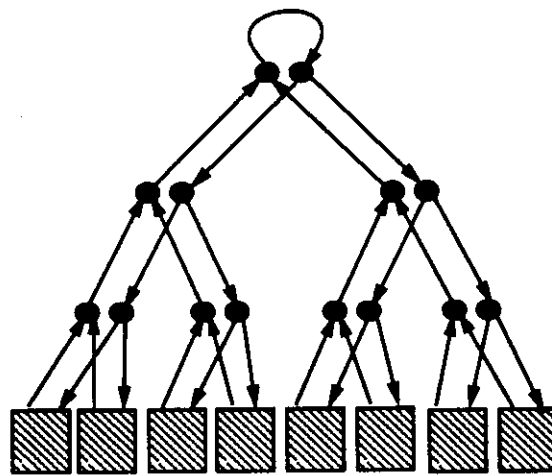
Figure 1.3: Tree-Net with 8 stations

based on a perfect binary tree with N stations, there are $2\log_2 N$ couplers on a path between a pair of stations. Assuming that each coupler incurs 3 dB losses (due to power splitting) the total power loss is $6\log_2 N$ [dB]. Thus, the power losses in the tree network grow logarithmically with the number of stations. This makes possible to support up to 64 stations with 40 dB power margin. Station connectivity can be significantly improved by introducing an all-optical amplifier at the root. Assuming an amplifier gain of 40 dB, the total power loss between between stations will be $3\log_2 N$ [dB] which makes possible to support up to 2048 stations.

Tree architecture offers a great potential for efficiently exploiting multiple fibers of commercial fiber trunks which can result in a significant increase of the network capacity [BGK93]. This solution is attractive because, although the multifiber cable is more expensive than single fiber cable, the marginal cost of an additional fiber is low, compared with installation and packaging costs. The payoff of the multifiber plant is very high, in that the number of available wavelengths is now amplified by a factor of the hundred.

1.1.3 Bus

The bus, illustrated in Figure 1.4 is another topology that can be used for PONs. Stations are connected to a linear, folded bus using couplers that first combine
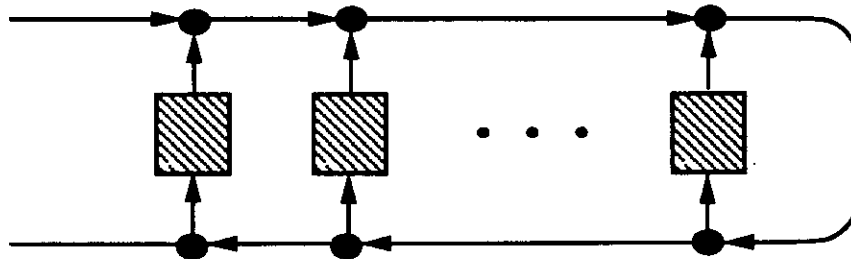
Figure 1.4: Bus topology

optical signals to the bus and then distribute the signals to the stations. If the couplers have uniform coupling ratio (i.e., 50% or 3 dB loss for tapping to the bus) the maximum power loss in the bus network (without link and coupler excess losses) is $(2N)\,3$ [dB]$= 6N$ [dB]. Therefore, the power loss grows linearly with the number of stations, which seriously limits the number of stations that can be supported in a network with this topology. In this case the maximum number of stations that can be supported is 6. It is possible to reduce losses to $10\log_{10} N^2$ [dB] if power distribution of couplers is optimal [Hen89]. This

result, however, is very difficult to achieve in a practical setting, with off-the-shelf couplers. The power losses will still grow linearly with the number of stations if the coupler's excess losses are not negligible. The problem of power losses can be alleviated to some extent using broadband optical amplifiers. To support 1000 stations on a bus, from 10 to 166 amplifiers are thus required (with optimized and uniform coupling ratios, respectively), as opposed to a single amplifier at the root of a tree.

1.1.4 Mesh topology (Linear Lightwave Network)

Recently, a new type of physical fiber layout infrastructure was proposed, namely the Linear Lightwave Network (LLN) [Ste91]. LLN is based on a mesh topology which can be partitioned into one or more trees using Linear Dividers-Combiners (LDCs), electrically controllable generalized optical switches (e.g., Ti:LiNbO$_3$ variable attenuators [KA88]) that are statically configured to distribute optical signals to specified outputs. LLN is not a passive network in a real sense since it uses LDCs which are active devices. However, it can still fit into passive optical network category because no electro-optic conversion or electronic processing is done on optical signals.

Each LLN tree provides full broadcast to all stations connected to it. Therefore, in LLN we can implement several subnets, each on a separate tree. By doing this, we can reduce power losses since power is being distributed only among the stations connected to the subnet. This also allows reuse of the same wavelength in separated subnets. Figure 1.5 illustrates a LLN with two tree subnets. Power losses within a subtree can be made as small as $10 \log_{10}(N-1)^2$ (where N is the number of stations connected to the subtree) by optimizing power distribution of LDCs as we will show in Chapter 4.

Although the subnets in LLN have a tree topology, there is a fundamental difference between the tree topology shown in Figure 1.3 and the one established in LLN. In a LLN tree as proposed by Stern [Ste91], optical signals travel on the shortest paths, and the signals are broadcast to all the stations except to the station that sent the signal. Thus, a transmitting station cannot receive its own transmission. Such signal distribution complicates coordination between stations in some multiaccess schemes as we will show later.

LLNs provide reconfigurability and fault tolerance. By changing power distribution at LDCs it is possible to change topology of the subnets and to overcome link or node failures. Also, power distribution can be tuned to minimize power budget.

6
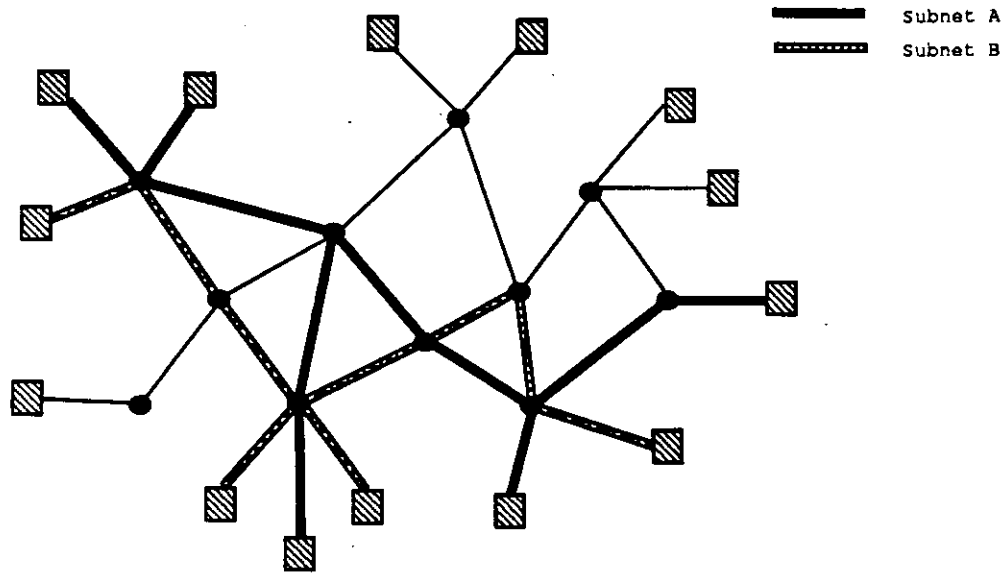
Figure 1.5: LLN with two subnets

The drawback of LLN is its complexity. It requires active devices (LDCs) and a separate network to control these devices. We will describe the LLN in more detail in Chapter 4.

1.2 WDM architectures and access schemes

Most proposed fiber optic WDM architectures in a local or metropolitan area environment can be divided into one of the following two categories:

- single-hop broadcast-and-select networks and

- multihop networks.

1.2.1 Single-hop broadcast-and-select networks

Single-hop networks can be defined as the networks where direct transmission can be achieved from a source to a destination. All input optical signals are combined and broadcast to all outputs thus providing for direct, "single-hop" communications from source to destination. The destination then selects the signals intended for it.

To permit multiple, simultaneous transmissions, wavelength division multiplexing (WDM) is used, and is often combined with time division multiplexing

(TDM). Figure 1.6 illustrates a scheme which uses both time and wavelength division multiaccess (T/WDMA). In this scheme time is divided into slots, i.e.,
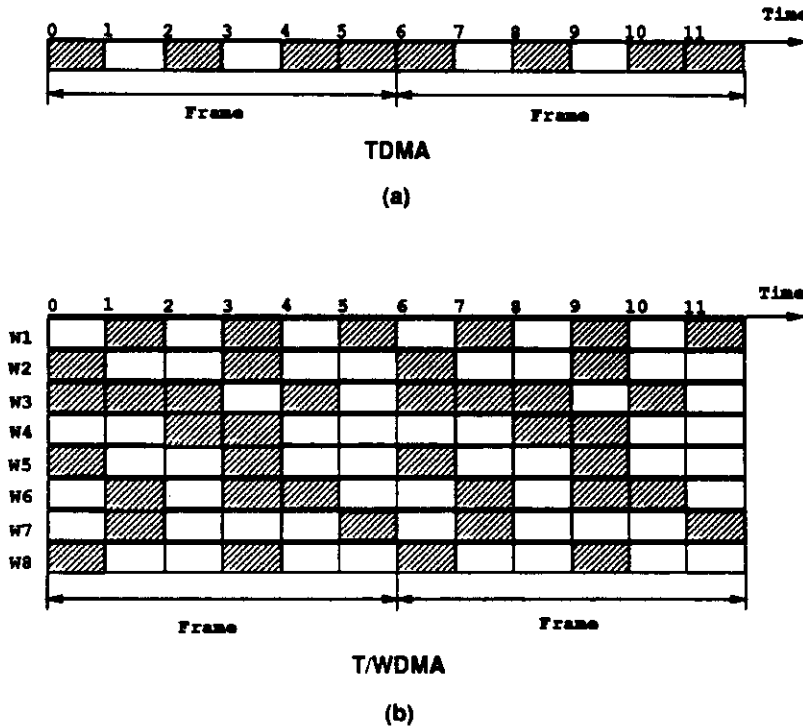
Figure 1.6: TDMA and T/WDMA

time intervals of fixed length in which a single station is allowed to transmit a packet. Time division multiaccess (TDMA) can be viewed as a special case of T/WDMA where only one wavelength is used. The user must thus "select" the wavelength and time slot at each transmission.

Several different possibilities exist, depending on whether transmitters, receivers, or both are tunable. In tunable transmitter fixed receiver configuration each receiver is assigned a fixed wavelength channel. A transmitter has to tune to the wavelength of the selected receiver. The problem with this configuration is that it cannot support multicast/ broadcast transmission. The configuration where transmitters are fixed and receivers tunable is similar to the previous one. The wavelength channels are now assigned to transmitters and a receiver has to tune to the wavelength of the appropriate transmitter. Multicast and broadcast transmissions can be supported by having several or all receivers tuned to the same source wavelength at the same time. The most flexible configuration is the one where both transmitters and receivers are tunable. It is shown in [LK92a]

that using both transmitters and receivers tunable provides better performance when the number of stations is larger than the number of wavelengths. However, the performance improvement comes at the expense of additional complexity of this configuration.

The tunable devices can be in principle replaced by arrays of fixed ones. However, the cost of such a network grows linearly with the increase in the number of optical channels (W). Thus, networks with large W would be prohibitively expensive.

In order to support connection oriented (real-time) traffic, a number of consecutive slots can be grouped into a frame. One or more slots in each frame can be reserved for a particular connection depending on the application data rate.

The T/WDMA scheme requires station synchronization in order to assure that transmissions occur in the assigned slots. This problem is very similar to the synchronization problem in TDMA satellite systems, and can be solved by borrowing techniques used in the satellite context. The problem can be easily solved in the case of star, tree and bus topologies. All stations have to be synchronized to a single reference point which can be the center of the star, the root of the tree or a point where the bus is folded. Each station has to know its propagation delay to the reference point in order to determine its transmission time. A station can easily determine its propagation delay by measuring the time it takes to receive its own transmission. In case of the LLN tree [Ste91] the synchronization is much more complicated. There is no single reference point that can assure proper synchronization. Also, since a station cannot receive its own transmission it is difficult to accurately determine propagation delays.

1.2.1.1 Technological and physical limitations

Efficient packet transmission in single-hop WDM networks require that a station be able to switch from one wavelength to another very fast. For example, in the T/WDMA scheme shown in Figure 1.6 a station may need to transmit or receive on different wavelengths in subsequent slots. Typically, this requires very fast tunable devices. However, current technology imposes severe limits to implementation of such a scheme due to limited tuning range and tuning speed of optical devices. External cavity, mechanically tuned devices have a broad tuning range. The tuning range of 57 nm has been reported for a mechanically tuned laser [WD83]. The number of WDM channels can be 1000 if highly selective mechanically tuned two-stage Fabry-Perot filters are used [Bra90]. The main disadvantage of mechanically tuned devices is very low tuning speed - on the order

9

of milliseconds. Due to the low tuning speed, these devices are not very suitable for T/WDMA. Acoustooptic lasers [CCC88] and filters [CLS89], on the other hand, have a medium tuning speed of around 10 μs and a broad tuning range of 87 nm and 400 nm, respectively. Due to low selectivity of these devices (channel spacing $\sim$1 nm), the number of WDM channels that can be covered is $\sim$100. An interesting property of acoustooptic filters is their multiwavelength filtering capability, which allows them to select more than one wavelength simultaneously. Semiconductor lasers and filters, such as distributed Bragg reflector tunable lasers [MMK87] and electrooptic tunable filters [WHA88] have tuning times in the order of nanoseconds but their physical characteristics limit the tuning range to no more than a few nanometers. Thus, these devices can operate only over a small number (typically less than 10) of wavelength channels. Generally, the faster the tuning time the smaller the tuning range. Recently, semiconductor lasers that can tune rapidly over a larger range have been demonstrated. In [JCC92] and [AKB92] semiconductor lasers that can tune over 30 nm and 57 nm were reported. These devices, however, are still being developed in research laboratories.

In addition or as an alternative to wavelength division multiplexing, subcarrier frequency division multiplexing has also been proposed [Dar87]. A number of microwave subcarriers can be multiplexed on the same wavelength. The advantage of this approach is that it uses mature RF/microwave technology. Retuning between subcarriers can be done much faster than between wavelengths. Also, spacing between subcarrier channels can be much closer than between wavelengths. However, a problem occurs in subcarrier frequency division multiaccess (SFDMA) systems if more than one station transmits on the same wavelength. The problem is that the wavelengths on which different stations transmit cannot be made exactly the same and that their relative difference causes destructive interference within the signal bandwidth [SEL91]. As a result, transmission failures may occur.

The switching speed in T/WDMA is also limited due to the dispersion phenomenon. Signals on different wavelengths travel at different velocities which results in different propagation delays. It is shown in [SH93] that for a distance of 125 km, the difference in propagation delays for wavelengths that differ by 200 nm is 400 ns, which is equivalent to 400 bits with 1 Gb/s rate. This causes a problem in synchronizing transmissions in a slotted T/WDMA system. The problem is illustrated in Figure 1.7. Figure 1.7.b shows timing of packet arrivals for the network in Figure 1.7.a where stations A and B transmit on wavelengths $W0$ and $W1$, respectively. In this example the transmitters are fixed and receivers tunable. We see that due to different propagation delays on wavelengths $W0$ and $W1$ receivers of stations C and D are not able to receive first slot on
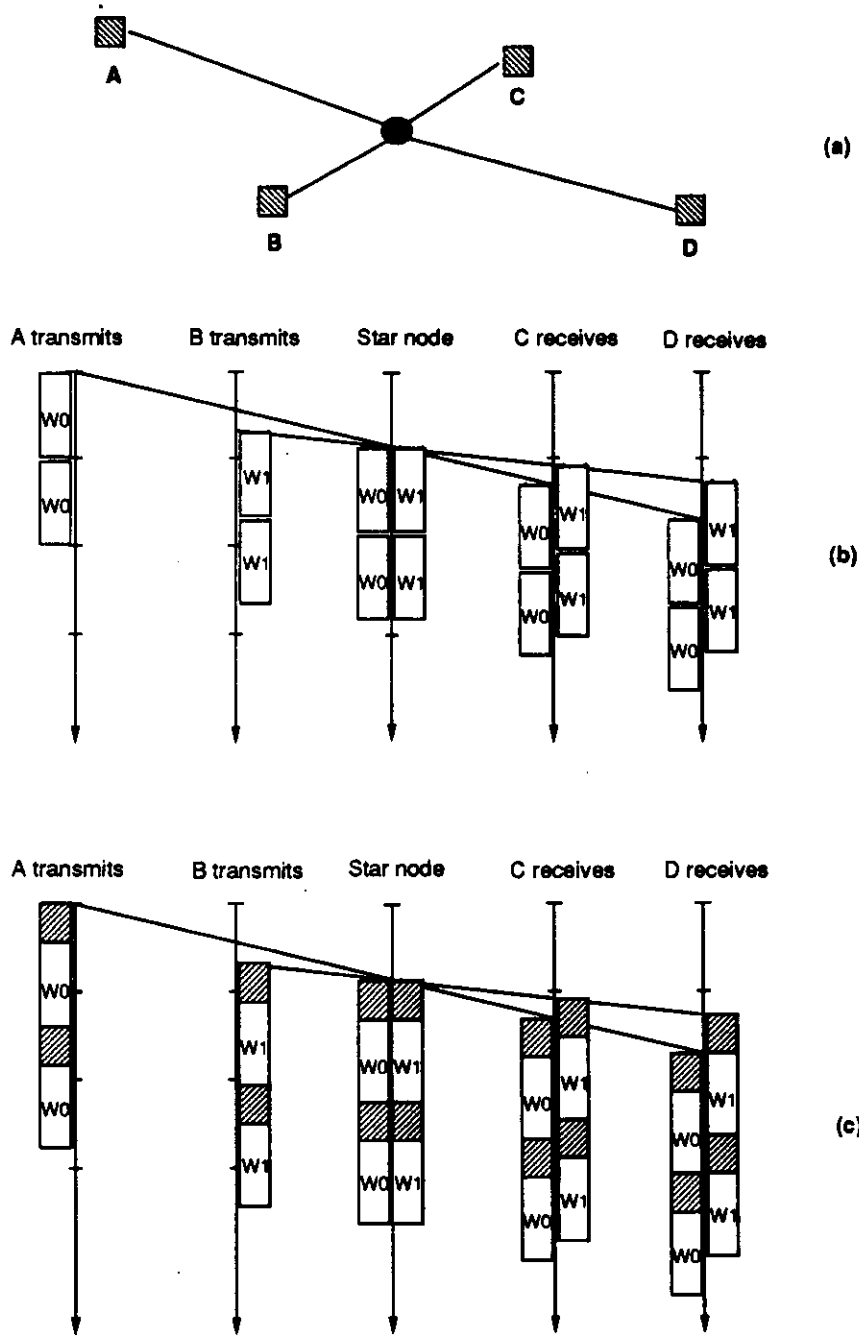
Figure 1.7: Wavelength dispersion and T/WDMA

11

$W0$ and then to switch to $W1$ for the second slot. Basically, a gap between slots has to be introduced in order to compensate the difference in propagation delays as shown in Figure 1.7.c. For a network with diameter of 125 km this gap is typically around 200 ns [SH93]. For larger network sizes, the difference in propagation delays and consequently the gap size increases. Thus, even if very fast tunable devices are available, we cannot perform tuning faster than allowed by the dispersion.

As we mentioned before, power splitting losses impose serious limits on the number of stations that can be supported in a PON. Optical amplifiers can to some extent overcome this problem. However, present optical amplifiers seriously reduce usable optical spectrum. The erbium-doped fiber amplifiers can provide a relatively flat gain only over a 30 nm bandwidth in the 1.5 μm band [Min91]. Semiconductor amplifiers that provide gain over 200 nm bandwidth are being developed [BSC90]. However, crosstalk can be a significant problem when a semiconductor amplifier is used to amplify several WDM channels simultaneously [RH90].

1.2.1.2 Experimental networks

Extensive research has been done in the past few years on wavelength division architectures and protocols. Some experimental networks have already been built.

The Bellcore's LAMBDANET system [GKV90] uses a combination of time and wavelength division multiplexing. Each node has a fixed transmitter and an array of 18 receivers. A grating demultiplexer is used to separate different optical channels. Each transmitter time-division multiplexes the traffic destined to all other nodes in a high-speed single wavelength data stream. Each receiving node simultaneously receives all the traffic, buffers it, and selects—using electronic circuits—the traffic destined for it. Two sets of experiments were performed, with 18 and 16 wavelengths, running at 1.5 Gb/s and 2 Gb/s, respectively.

IBM's Rainbow [DGL90] is a circuit-switched research prototype network consisting of 32 IBM PS/2 stations communicating with each other at 200 Mb/s data rates. It has fixed transmitters and slowly tunable Fabry-Perot filters for receivers. Because of the slow tunability of the receivers (submillisecond tuning time), an entire optical channel must be allocated to a source destination pair for the duration of their connection.

FOX (Fast Optical Cross Connect) [ACG86] is an experimental architecture for interconnection of processors with shared memories in parallel processing computers. It consists of two star networks, one for transmissions from processors to

memories and another for transmissions in the opposite direction. Each processor and memory module is equipped with a fast tunable transmitter and a fixed receiver.

HYPASS (Hybrid Packet Switching System) [AGK88] is a lightwave network architecture that also consists of two broadcast stars. One star is used for data transmissions and another for sending status information of output ports. Each sender has a tunable transmitter and a tunable receiver. Each receiving node has a fixed transmitter and receiver. HYPASS represent a hybrid electronic/optical switch, as it uses electronics for signal processing and optics for signal transport.

1.2.1.3 Protocols

Most single-hop WDM access protocols are proposed for networks that use tunable devices. The main problem in these networks is to coordinate transmission in such a way that both transmitter and receiver are tuned to the assigned wavelength.

The simplest wavelength division multiaccess protocol supports only connection oriented traffic. A pair of stations uses the assigned wavelength for the duration of its connection. Such a protocol is implemented in Rainbow network [DGL90] where each station has a fixed transmitter and a slowly tunable receiver. Each transmitter is assigned its own wavelength. A transmitter, when it has a packet to transmit, repeatedly sends requests to transmit to a particular destination, until it receives an acknowledgment. Each receiving station, when idle, polls all transmitters (by tuning on their wavelengths) to see if there is one that requests transmission, and returns an ack to it. Clearly, this protocol is not suitable for datagram traffic due to slow tuning and long connection establishment time (much longer than the duration of packet transmission). Also, the protocol does not allow a station to maintain more than one connection at a time.

In order to support packet-switched traffic and multiple connections, it is necessary to use a combination of time division and wavelength division multiplexing. Many such protocols have been recently reported in the literature.

The simplest packet-switched protocol that does not require any synchronization between stations can be implemented with one tunable transmitter and one fixed receiver at each station. Such a protocol is proposed for FOX architecture [ACG86]. Each receiver is assigned its own wavelength (i.e. the receiver is "wavelength addressable"). The transmitter tunes to the wavelength of the receiver and transmits a packet. A collision may occur if two or more stations simulateneously transmit to the same destination in which case the source has

to retransmit the packet. The source is informed about success of its transmission by an acknowledgment from the destination. However, the acknowledgment can also be lost due to collision. Alternatively, a station may use an additional tunable receiver that tunes to the wavelength of the intended destination and checks whether its transmission suffered a collision. Thus, this is a random access, Aloha type protocol which can be used only for low traffic loads. Random access protocols are inefficient and suffer from instability when the traffic load is high [Kle76]. Also, there is no guaranteed packet delay, since a packet may need to be retransmitted an arbitrary number of times due to packet collisions.

The other possibility is to use slotted protocols which require synchronization between stations. Many slotted protocols have been proposed. In [Dow91] a slotted variant of the previous protocol which uses a tunable transmitter and a fixed receiver was proposed which has improved performance but still suffers from the same limitations. In [CG88] a T/WDMA protocol is proposed where each station has a transmitter and receiver that are tunable over a number of frequencies. Each time-wavelength slot of the frame is permanently assigned to a source destination pair. The advantage of this protocol is that no pretransmission coordination is necessary since transmissions are predetermined. However, the problem is that the protocol is not scalable. The number of slots required per frame grows quadratically with the number of stations. In addition to this, the protocol is not sensitive to dynamic bandwidth requirements. In [CG88] is also proposed a partial fixed assignment where a number of transmitters or receivers share the same slot. However, such assignment can lead to collisions if two or more stations sharing the same slot are engaged in transmissions.

In order to overcome limitations of previous schemes, a number of protocols have been proposed that perform dynamic slot assignments. These protocols require a separate control channel for pretransmission coordination. The control channel is usually implemented using an additional wavelength. These schemes can be implemented using a single tunable transmitter-receiver pair where stations hops between control and data channels [HKS87, Meh90, JM92]. More efficient protocols use an additional, fixed transmitter-receiver pair for control channel [CDR90, SGK91, LK92b, CY91].

Slot reservation in the dynamic assignment schemes can be done using a random access technique [HKS87, Meh90, SGK91, LK92b] or using minislotted TDMA [CDR90, CY91, HRS92]. Generally, any protocol that has a random access component suffers from potential instability and unbounded delays which does not make it suitable for real-time applications. On the other hand, protocols that use TDMA reservation, where each minislot is assigned to one station, are

not scalable, since the number of minislots required grows with the number of stations. In Dynamic Time and Wavelength Division Multiaccess Scheme (DT-WDMA) [CDR90] the size of each data slot is equal to the size of N minislots where N is the number of stations. Having a very large number of stations would result in too large data slots which would adversely affect network efficiency.

In the dynamic slot assignment schemes the control channel becomes a bottleneck due to large processing requirements. Transmission of each packet requires processing of control packets. Each station has to process all the information transmitted over the control channel. One solution to this problem is to use more than one wavelength channel for control. In [HRS92] an extension of DT-WDMA is proposed which uses a separate control channel for each station. Thus, N wavelengths are used for control where N is the number of stations. There is also need to use tunable devices to implement those control channels. In [HRS92] a tunable transmitter and fixed receiver are used for the control channel.

In [GK91] protocols have been proposed that take into account limited tunability of fast tunable devices. Each station has a tunable transmitter and a number of fixed receivers. Each transmitter can tune over a limited number of contiguous wavelengths. The wavelengths are assigned to transmitters and receivers in such a way that each transmitter and each receiver can communicate in one hop. The main limitation is, however, that the number of receivers required per each station grows with the increase in the number of wavelengths.

Most of the protocols described above are intended for packet-switched traffic. However, none of the protocols proposed so far can integrate all traffic types efficiently. The protocol in [HRS92] makes an attempt to integrate both the packet switched and the circuit-switched traffic. However, this protocol is designed only for point-to-point traffic and therefore, broadcast/multicast cannot be supported. Also, the protocol requires very fast tuning, and a very large number of wavelengths. Practically, N wavelengths are wasted for the control.

1.2.1.4 Advantages and disadvantages of the single-hop approach

The advantage of the single-hop architectural concept is that real-time connections can be easily supported (for example, a particular slot in each time frame can be dedicated for a real-time connection). Also, multicast and/or broadcast can be easily achieved because of inherent broadcast nature of this type of networks. This, of course requires tunable receivers (or an array of fixed receivers), while transmitters can be fixed or tunable.

The main limitation of single-hop networks is the fast wavelength switching

requirement. While this requirement is not so critical for circuit-switched traffic, it is essential for efficient handling of packet-switched traffic. As we already pointed out, the present devices can achieve fast tuning only over a limited tuning range. This imposes limit on the aggregate throughput that can be achieved in the network.

Also, most efficient packet-switched protocols require additional bandwidth and (one or more wavelengths for control channel) and additional hardware (extra transmitter and receiver per station for control channel) and have large processing overhead for pretransmission coordination. The processing overhead for circuit-switched traffic is much smaller since it is done only for circuit set-up and circuit release. Since a connection can last minutes and perhaps hours, the processing overhead can be considered to be negligible.

1.2.2 Multihop networks

Instead of using a direct path from source to destination, multihop networks require some packets to travel across several hops. Basically, the multihop network is a store-and-forward network embedded in the passive optic network. The switches of the multihop network are represented by the user stations (thus, they are located at the periphery of the passive broadcast medium); the links consist of dedicated wavelength channels established between pairs of stations [AKH87]. Thus, over the physical broadcast topology, there is a virtual topology that determines the actual connectivity between the stations in the network.

As in conventional store-and-forward networks, packets may need to be queued at intermediate nodes. As an alternative for packet queueing, deflection or hot-potato routing has been proposed [Max87, AS91, ASZ92, BFG91]. A packet can be intentionally misrouted, but still reach its destination using a slightly longer path. The congestion problem in deflection routing networks can be elegantly solved because new packets are prevented from entering the network when congestion occurs.

Clearly, each hop incurs the penalty (in terms of additional packet delay and processing overhead) of an electro-optical conversion. Also, multihopping reduces network throughput. The effective network capacity is inversely proportional to the number of hops for packet transmissions. Thus, the virtual topology should be designed to minimize the number of hops.

16

1.2.2.1 Virtual topologies for the multihop networks

The regular virtual topologies that have been proposed for the multihop optical network are the ShuffleNet [AKH87], the de Bruijn graph [SR91], the torus [Max87, Aya89, GL92], the hypercube [LG92], the ring [KP92], and the dual bus [BMS92].

The most prominent example of virtual multihop network is ShuffleNet, proposed in [AKH87]. ShuffleNet exploits WDM to embed a perfect shuffle interconnection within a fully broadcast physical topology. A (p, k) ShuffleNet can be constructed with $N = kp^k$ nodes where p and k are positive integers. The nodes are arranged in k columns of p^k nodes each.

Figure 1.8 shows an example of an 8 station ShuffleNet virtual topology implemented on a physical tree. In this example each station has two pairs of "fixed"
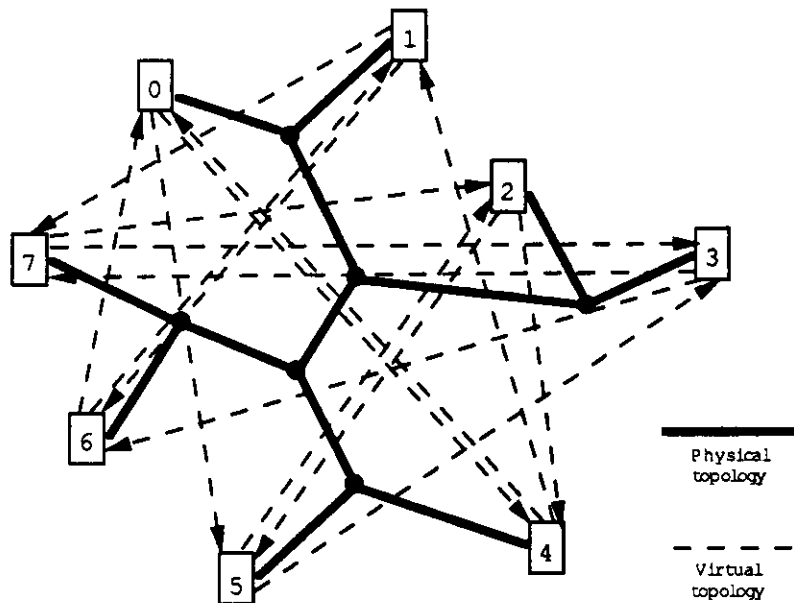
Figure 1.8: ShuffleNet virtual topology implemented on a physical tree

transmitters and receivers (i.e. $p = 2$). (In fact, these transmitters and receivers are not really fixed, but slowly tunable in order to make it possible to change virtual topology.) The number of wavelengths required in this case is twice the number of stations. Figure 1.9 shows a ShuffleNet virtual topology with 8 stations. The maximum hop distance between two nodes in the ShuffleNet is $2k - 1$. Since k grows logarithmically with N, the maximum as well as the average hop distance grows logarithmically as well.

De Bruijn graph is a perfect shuffle topology where all stations are organized
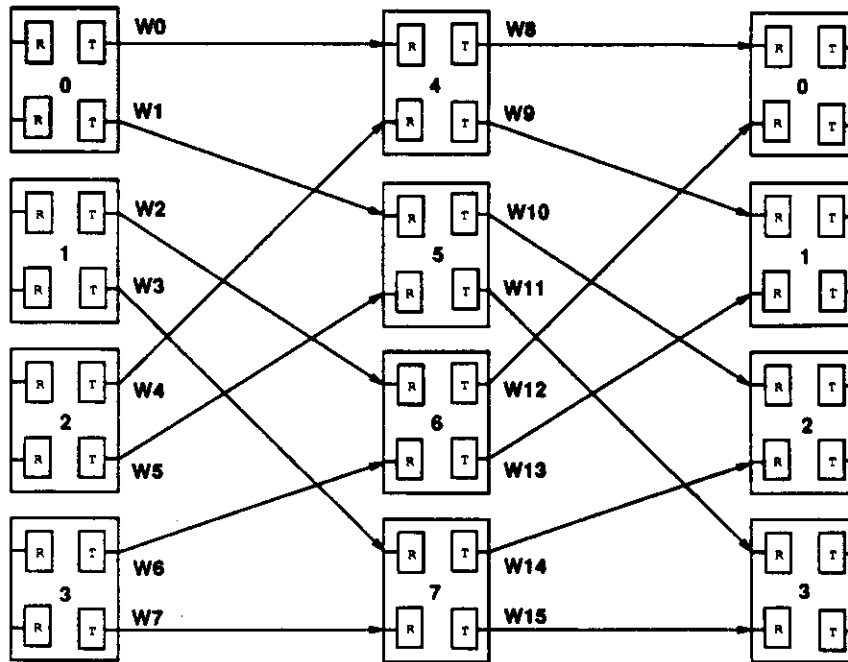
Figure 1.9: Virtual Topology of the Eight Station ShuffleNet

into a single column. Thus, the number of stations in de Bruijn graph is $N = p^k$. Figure 1.10.a shows the de Bruijn graph with $p = 2$ and $k = 3$. De Bruijn graph virtual topology corresponds to the state transition diagram of the shift register. Each state is represented by a p-ary number with k digits. The transition from one state to another corresponds to a left shift of the number followed by the concatenation of a single p-ary digit onto its right. In [SR91] the ShuffleNet and the de Bruijn graph have been compared and it is shown that de Bruijn graph has smaller average number of hops. The maximum number of hops for de Bruijn graph is k (more than k shifts and concatenations will cause the original number to repeat). However, the de Bruijn graph topology is not fully symmetric, which results in unevenly balanced link loads in a network based on this topology. As a result of this asymmetry, the maximum throughput supportable by this topology is smaller than in a Shufflenet with the same number of nodes and the same nodal degree.

In the toroid topology, the nodes are organized into a two-dimensional grid where the nodes of the first and the last row and column of the grid are connected thus forming a torus. Each nodes has four neighbors to which they can communicate directly. In [Max87] a Manhattan Street Network (MSN) is proposed where each station can transmit to two of its neighbors, and receive from other two as shown in Figure 1.10.b. Thus, such a network requires two transmit-

18
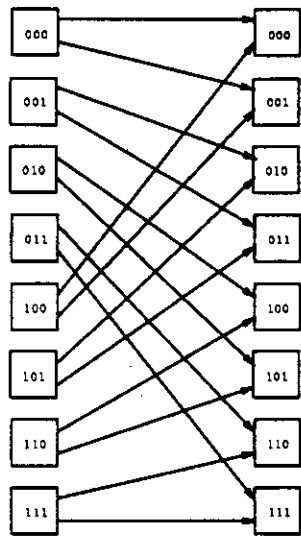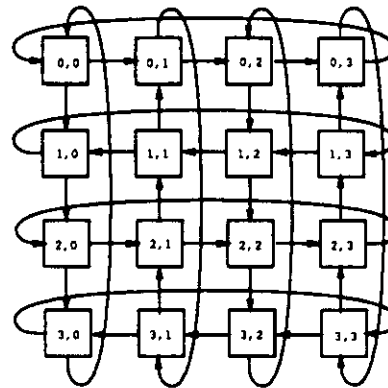
a) DeBruijn graph

b) Manhattan network
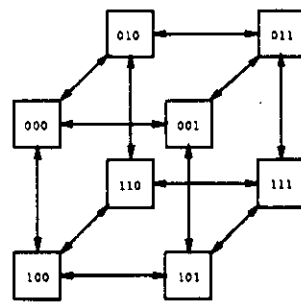
c) 3-dimensional binary hypercube
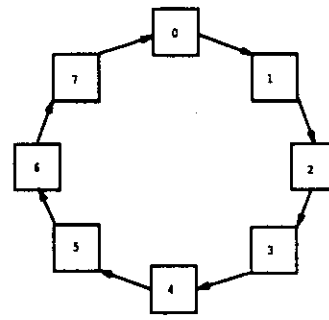
d) Ring

e) Dual bus

Figure 1.10: De Bruijn graph, Manhattan grid, hypercube, ring and dual bus virtual topologies

ters and two receivers. The connections are organized in the same manner as are the one-way streets of Manhattan. A bidirectional Manhattan grid, where each station can transmit to and receive from all four neighbors is also proposed. In this case each station needs four transmitters and four receivers. The advantage of this topology is that the deflection routing can be very efficiently implemented since the number of extra hops in case of deflection is small. The maximum and the average number of hops for this topology grows with the square root of the number of stations.

The simplest form of the hypercube topology is the binary hypercube which has $N = 2^p$ nodes, each of which has p neighbors. The binary hypercube with 8 nodes is shown in Figure 1.10.c. A node address consists of p binary digits and two nodes whose addresses differ in a single digit are neighbors. A node communicate with each neighbor bidirectionally. Thus, each node requires p transmitters and p receivers. This topology is evaluated in [LG92]. The advantage of this topology is that the number of hops grows logarithmically with the number of stations. The disadvantage is that the number of transceivers required per station also grows logarithmically with the number of stations. A generalized hypercube has been also proposed [Dow92, LG92]. In the generalized hypercube each address digit can have a different radix, which can be an arbitrary integer. The generalized hypercube provides more flexibility in accommodating different number of nodes but it has similar characteristics as the binary hypercube.

The ring virtual topology, shown in Figure 1.10.d, is used in the STARNET optical network [KP92]. The advantage of this topology is its simplicity. Each station transmits to the next station in the ring. The addition or removal of a station is relatively simple and it does not destroy the regularity of the topology as in the case of previously described virtual topologies. However, the problem with this topology is that the average number of hops required grows linearly with the number of nodes, which seriously limits the network throughput.

The dual bus topology [BMS92], shown in Figure 1.10.e, is another linear topology that is used for the multihop network. In this topology all nodes are organized in a row where each node is connected in both directions with the previous and the next node in the row. Networks based on the dual bus has same advantages and disadvantages as those based on the ring topology.

If traffic distribution is not uniform, regular virtual topologies are not optimal. It is possible to achieve better performance using some other non-regular virtual topologies. In [BG90] virtual topology optimization problem is addressed which minimizes mean packet delay. In order to construct optimal virtual topology a simulation annealing algorithm is used. In [LA91] the construction of optimal

topologies is studied based on minimization of the maximum link flow. A heuristic algorithm is developed that gives a suboptimal solution.

1.2.2.2 Shared-channel multihop networks

In order to utilize better wavelengths and to reduce the number of required wavelengths more than one transmitter and receiver can be allowed to access the same wavelength channel. Channel sharing also makes it possible to implement the same virtual topology using fewer transceivers per station or to implement more connected topologies using the same number of transceivers. It can provide additional flexibility in routing. The channels are shared using time division multiple access technique [HK88, BFG90, GL92]. Hluchyj and Karol [HK88] studied channel sharing for ShuffleNet and presented a routing algorithm for the shared-channel ShuffleNet that routes traffic along shortest paths in such a way that the traffic load on all channels is perfectly balanced when traffic is uniform. Bannister et al. [BFG90] studied optimal shared channel virtual topologies for nonuniform traffic. A perfect shuffle virtual topology is modified by doing branch exchange using a genetic algorithm. In [GL92] channel sharing in the Manhattan Street Network (MSN) is studied. In this network each row and each column represents a shared channel. It is shown that for the uniform traffic case the shared channel MSN can support higher aggregate network throughput than the original MSN. Ramaswami and Sivarajan [RS93] studied channel sharing using subcarrier multiplexing. They also developed a methodology for building a shared channel topology and showed that ShuffleNet and de Bruijn graph belongs to the class of topologies that permits channel sharing.

1.2.2.3 Advantages and disadvantages of the multihop approach

The main advantage of the multihop architecture is that it does not require tunable devices. The efficiency is not that good as in the single-hop network since a packet has to travel several hops. By choosing an appropriate virtual topology it is possible to minimize number of hops and to achieve very high throughputs. The network capacity grows with $O(N/\log N)$ if ShuffleNet or de Bruijn graph virtual topology is used. The multihop network can be implemented with a small number of fixed transmitters and receivers per station. Practically, the minimum number required is one transmitter and one receiver per station.

The advantage of the multihop architecture is that it supports datagram traffic efficiently and it is scalable. Aggregate throughput increases with the increase in the number of stations. In multihop networks where all stations are connected

to the same PON and use wavelength division multiplexing the throughput is, however, limited since the number of available wavelengths is limited. This problem can be easily overcome if the implementation of the multihop network is not restricted to a single "mass of glass" medium.

The main problem with the multihop architecture is that it does not support efficiently real-time traffic because of possible large delay fluctuations. A scheme has been proposed to overcome this problem [BFT91] but it results in a substantial increase in network overhead. Also, broadcast and multicast are not easy to implement.

1.3 Contributions of this research

In this dissertation we propose a fiber optic network for a local or metropolitan area environment that can support a wide range of applications using available optical technology. In Chapter 2 we describe HONET, a hybrid architecture which combines both the single-hop and the multihop approach in order to exploit the advantages of each. We study virtual topologies for HONET's multihop network that can be realized with TDMA channel sharing. We first consider the least expensive implementation of the multihop network based on single fixed transmitter-receiver pair per station. An analytic upper bound on the maximum throughput that can be achieved with any virtual topology is developed. It is shown that the ShuffleNet virtual topology has throughput which is very close to the upper bound. We also study virtual topologies where each station has more than one transmitter-receiver pair. It is shown that in this case the best throughput is obtained when no channel sharing is performed. The HONET architecture is then compared with other related architectures.

In Chapter 3 we propose a time and wavelength division multiaccess scheme for HONET's single-hop component. The same scheme is applicable to a general broadcast and select network that supports only circuit-switched traffic. The advantage of the proposed scheme is that it is able to support efficiently even a very large number of connections of very small data rates. It also allows each station to maintain many simultaneous connections. The unique feature of this scheme is that it can be efficiently implemented using tunable devices which have tuning (i.e., wavelength switching) time much longer than the TDM slot duration. We present first the performance analysis of this scheme for the uniform, single-rate, point-to-point traffic case. The main performance criteria are blocking probability and network throughput. An approximate analytical model based on Markov chains is developed and is shown to match very closely the simulation

results. The basic scheme is then extended by allowing receivers to exploit the multiwavelength filtering capability of acoustooptic tunable filters. It is shown that multiwavelength filtering capability can be used to improve performance. Finally, we perform a simulation study of the basic scheme for nonuniform traffic characterized by strong traffic locality. The results show that traffic locality can be exploited in order to improve the performance of this scheme.

In Chapter 4 we propose a solution for the synchronization problem in Linear Lightwave Networks. The lack of global synchronization in the basic LLN complicates the implementation of slotted TDMA and T/WDMA schemes. Our proposed solution is based on using a new optical signal routing scheme called *rooted routing* instead of the originally proposed, shortest path routing. This solution permits us to implement the HONET over the LLN. An impact of rooted routing on power losses and propagation delays is analyzed and an approach for optimizing power losses in LLNs with rooted routing is presented.

In Chapter 5 we propose multilevel and multifiber extensions of the basic HONET architecture in order to increase scalability and capacity. The multilevel architecture is hierarchically organized. It exploits traffic locality and wavelength reuse in order to increase the network capacity. A multifiber infrastructure, such as multifiber tree network or multifiber LLN, can also be used to increase capacity and scalability of the HONET architecture.

Conclusions and some directions for further research are given in Chapter 6.

CHAPTER 2

HONET Architecture and Packet Service

A promising approach to overcome limitations of single-hop and multihop networks is to implement a hybrid scheme in which both single-hop and multihop are included, and are used to support the services for which they are best suited. This approach is prompted by the observation that these schemes have complementary features: for example, single-hop networks can easily support real-time and broadcast/multicast (but are not appropriate for packet-switching); whereas multihop networks can easily handle data traffic (but do have difficulties in providing broadcast/multicast and real-time traffic support). By combining features of both architectures we could enjoy the benefits of both. Hybrid optical architectures that combine single-hop and multihop approaches have already been reported in the literature [KP92, GKB92].

STARNET [KP92] is a research prototype built at Stanford University. It is an optical network implemented over a broadcast star topology. STARNET handles packet-switched traffic using a multihop subnetwork and circuit switched traffic by a single-hop subnetwork. STARNET uses one fixed transmitter and two receivers, one fixed and another slowly tunable. Each station transmits on a unique wavelength. The transmitter is shared by both the multihop and the single-hop subnetwork. The fixed receiver is used by the multihop subnetwork and the tunable receiver by the single-hop subnetwork. The multihop subnetwork has ring virtual topology and it uses the FDDI protocol. The single-hop subnetwork operates like Rainbow [DGL90] where slowly tunable receiver tunes to the wavelength of the transmitter and stay on this wavelength for the duration of the connection. Connection establishment coordination for the single-hop subnetwork is performed over the multihop subnetwork.

MONET (Multilevel Optical Network) [GKB92] is another optical network proposal based on the idea of combining the single-hop and the multihop approaches. MONET consists of single-hop subnetworks (clusters) interconnected by a multihop backbone. Although MONET has advantages over existing single-hop and multihop networks in terms of supporting various traffic types, it still suffers of some limitations. Namely, within each single-hop cluster, datagram traffic is not handled efficiently. At the same time, the multihop backbone is

25

not suitable for real-time intercluster traffic. In MONET it is possible to use a pool of on-demand wavelengths for real-time intercluster connections, at the cost, however, of an additional tunable transmitter-receiver pair per station.

We investigate the hybrid concept more systematically than in previously reported studies and propose a new hybrid architecture called Hybrid Optical Network (HONET). HONET is based on the same concept as STARNET but it is more general. It is a synergy of two networks, a single-hop network that supports circuit-switched and multicast/broadcast traffic and a multihop network supporting packet-switched traffic. Both networks can use the same broadcast medium, but separate sets of wavelengths. (It is also possible to implement these two networks on separate fibers.) The single-hop and the multihop network, besides providing services for different traffic types, can also support each other in performing network control and management functions. The multihop network can be used for transmitting control data for access control to the single-hop network. For example, reservation of a particular wavelength channel or slot for a connection over the single-hop network can be done over the multihop network. Likewise, the single-hop network can provide broadcast of control data necessary for management of the multihop network. For example, information about virtual topology changes (like addition or removal of a station or changes in virtual links) can be sent to all the stations in a single broadcast message.

Generally, HONET can be viewed as a network which consists of the multihop network of an arbitrary virtual topology and the single-hop network based on a dynamically assigned T/WDMA scheme. STARNET can be viewed as a special case of HONET, in which the virtual topology of the multihop network is a ring, and WDMA (i.e., a special case of T/WDMA where the number of slots per frame is 1) is used over the single-hop network (i.e., the connection over the single-hop network uses an entire wavelength channel). HONET provides more flexibility in satisfying user's demands by allowing many different configurations. For example, a station can access the single-hop network using TDMA (which is a special case of T/WDMA where the number of wavelengths is 1) that requires only a fixed transmitter-receiver pair. Such a configuration would be the most suitable for small users that do not require high data rates. Stations' performance can be improved by using fast tunable devices and the T/WDMA scheme instead of TDMA.

We can implement T/WDMA even if some stations have only TDMA capability (i.e., fixed transmitters and receivers). A connection between a pair of stations that use fixed devices tuned on different wavelengths must then be established through gateways. Namely, a source station sets up a connection to a

gateway, a station that has fast tuning capability, by reserving slots within the TDMA frame. The gateway then sets up a connection to the destination station. The gateway basically operates as a circuit switch. Broadcast to stations that do not have tuning capability can be done in a similar manner. For each wavelength one gateway station is selected which, upon receiving broadcast data, retransmit· them on the assigned wavelength.

Some stations may have an additional, slowly tunable transmitter-receiver pair that can be used for applications that require very high data rates (e.g., 1 Gb/s). A pair of stations equipped with the additional devices can access a pool of on-demand wavelengths and reserve an available wavelength channel to use it exclusively for its connection.

The HONET's multihop network can be built using an arbitrary number of transmitters and receivers per station. The simplest implementation requires a single transmitter-receiver pair per station. Using additional devices increases the station's throughput and the total throughput as well.

2.1   Virtual topology for HONET's multihop network

As we mentioned before, the HONET multihop network can have an arbitrary virtual topology. Generally, the choice of virtual topology is restricted by the number of devices each station has if a wavelength channel is dedicated to each virtual link. If we allow channel sharing using TDMA, which we propose for HONET, the number of possible virtual topologies increases significantly. In this section we try to evaluate possible virtual topologies and to select those that would be the most suitable for HONET.

In the HONET architecture, real-time and other connection oriented applications are supported by the single-hop network. HONET's multihop network supports datagram traffic which can tolerate relatively large delays. The delay critical applications are handled by HONET's single-hop network. Therefore, the delay is not such an important criterion in choosing the best virtual topology for HONET. Our primary goal is to choose the virtual topology which maximizes throughput.

We consider first virtual topologies for HONET's multihop network with minimal hardware requirements: one fixed transmitter and receiver per station. If no channel sharing is permitted, each station can receive from or transmit to only one other station in a single hop. In such a case, the only virtual topology that can provide full connectivity among stations is the ring. Such a multihop network is used in STARNET [KP92]. Let us define the station's connectivity factor p as

27

the maximum number of stations the station can receive from or transmit to in a single hop (i.e., the maximum indegree or outdegree of each station). Clearly, without channel sharing we have $p = 1$.

We can increase p by allowing TDMA channel sharing. Figure 2.1 shows the ShuffleNet virtual topology with $p = 2$ implemented with a single transmitter/receiver pair. In this case four virtual links share a wavelength channel. We
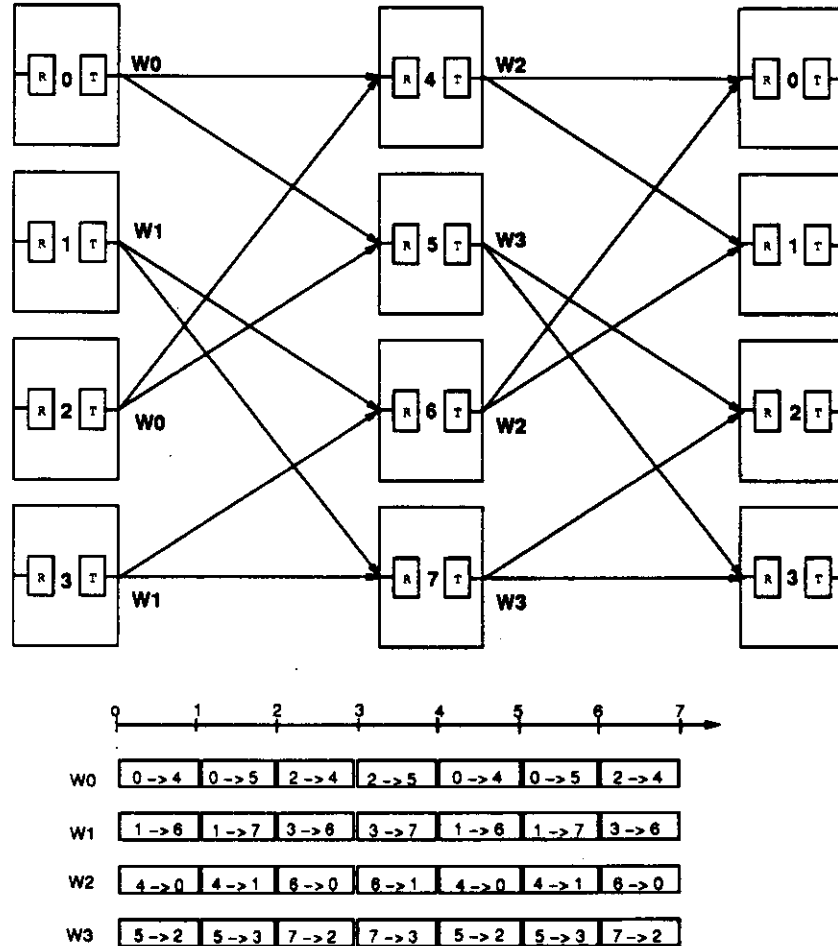
Figure 2.1: ShuffleNet virtual topology implemented using one transmitter and one receiver per station

know that the increase in p reduces the average number of hops which increases utilization of virtual links. On the other hand, the increase in p reduces the capacity of virtual links.

Let us try to find the optimal value for p which maximizes throughput per station. Let L be the arrival rate of traffic entering (leaving) the network (i.e., the

network throughput). We assume that each station generates an equal amount of traffic destined to all other stations (i.e., a uniform traffic distribution).

Let E be the average number of hops on a path from a source to a destination station. If the traffic distribution is uniform, each packet entering the network will require E transmissions on the average. Thus, the total arrival rate to the stations in the network which includes new and forwarded packets is $\lambda = LE$.

If each station has p incoming and outgoing links, and the capacity of a wavelength channel is C, the capacity of each virtual link is

$$C_{vl} = \frac{C}{p^2}$$

since p^2 virtual links share the same channel. The number of virtual links in the network is

$$n = Np$$

and the total network capacity

$$C_{tot} = nC_{vl} = \frac{NC}{p}$$

The total traffic in the network cannot exceed the total capacity. Thus,

$$\lambda < C_{tot}$$

which gives

$$L < \frac{NC}{pE} \tag{2.1}$$

Let us now try to determine the value E. From [HK88] we have that the lower bound for the expected number of hops in a network where each station is connected with p other stations is

$$E_{lb}(p) = \frac{p - p^{H+1} + NH(p-1)^2 + H(p-1)}{(N-1)(p-1)^2} \quad \text{if } p > 1 \tag{2.2}$$

$$E_{lb} = \frac{N}{2} \quad \text{if } p = 1 \tag{2.3}$$

where H represents the maximum number of hops from source to any of the destinations.

$$H = \begin{cases} \lceil \log_p(1 + N(p-1)) \rceil - 1 & \text{if } p > 1 \\ N - 1 & \text{if } p = 1 \end{cases} \tag{2.4}$$

Figure 2.2 shows how the lower bound on the average number of hops E_{lb} changes with p. Assuming that propagation delays dominate over transmission
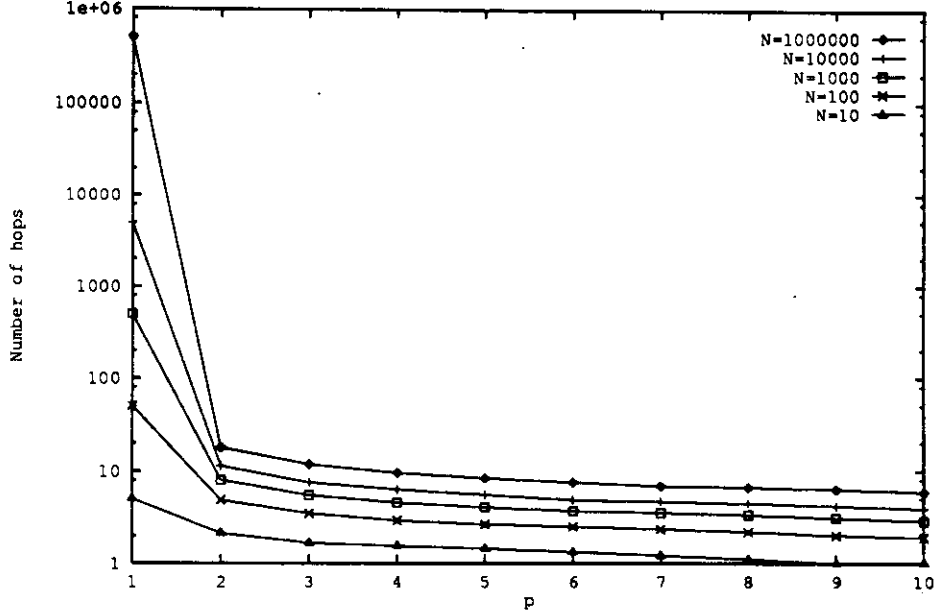
Figure 2.2: Lower bound for the average number of hops versus connectivity factor p when the number of stations N is 10, 100, 1000, 10000 an 1000000

and processing delays at intermediate stations, the average network delay is proportional to the average number of hops.

The upper bound for the aggregate throughput is thus

$$L_{ub} = \frac{NC}{pE_{lb}(p)} \tag{2.5}$$

and the upper bound for the mean throughput per station is

$$\sigma_{ub} = \frac{L_{ub}}{N} = \frac{C}{pE_{lb}(p)} \tag{2.6}$$

Figure 2.3 and Figure 2.4 show how L_{ub} and σ_{ub} depend on p when C is normalized to 1. **Parameter** N represents the number of stations in the network. We see a significant improvement in throughput with the increase in p from one to two for all cases of N. In fact, the improvement is higher when N is larger. This is an expected result since for $p = 1$ (i.e., ring) the number of hops grows linearly with N, while for $p > 1$ it grows logarithmically. We also see that when the number of stations is 10, 100 and 1000, σ_{ub} decreases for $p > 2$. The decrease is slower as the number of stations is larger. When the number of stations is 10000 and 1000000, σ_{ub} is almost the same for $p = 2$ and $p = 3$. In fact, σ_{ub} is slightly higher

30
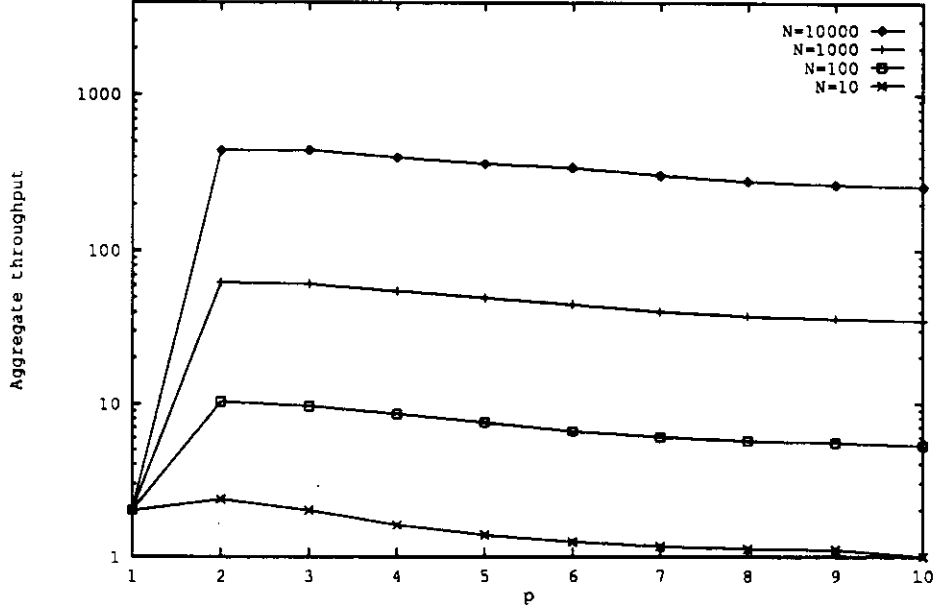
Figure 2.3: Upper bound for the aggregate network throughput L_{ub} versus connectivity factor p when the number of stations N is 10, 100, 1000 and 10000

for $p = 3$ in those cases. For $p > 3$, σ_{ub} slowly decreases. Thus, we see that the choice of $p = 2$ or $p = 3$ gives the optimal value for σ_{ub} in most cases of practical interest.

Let us consider the ShuffleNet topology (which requires that the number of station is $N = kp^k$ where k is an integer). It is shown in [HK88] that the average number of hops in ShuffleNet is

$$E(p, k) = \frac{kp^k(p-1)(3k-1) - 2k(p^k - 1)}{2(p-1)(kp^k - 1)} \qquad (2.7)$$

The maximum throughput that can be achieved in ShuffleNet is thus

$$L_{max} = \frac{NC}{pE(p, k)} \qquad (2.8)$$

and the maximum average throughput per station

$$\sigma_{max} = \frac{C}{pE(p, k)} \qquad (2.9)$$

Figure 2.5 shows the maximum average throughput per station in ShuffleNet for station connectivity $p = 2$ and $p = 3$. We see that the maximum throughput

31
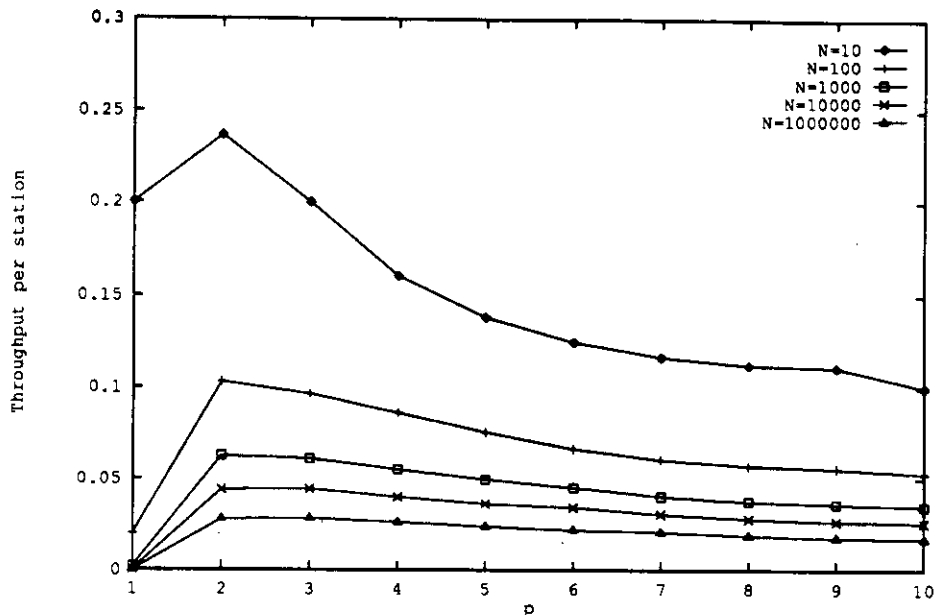
Figure 2.4: Upper bound for the average throughput per station σ_{ub} versus connectivity factor p when the number of stations N is 10, 100, 1000, 10000 and 1000000

that can be achieved using the ShuffleNet virtual topology is very close to the upper bound.

The above results imply that for the uniform traffic the best virtual topology (with respect to throughput) that can be built using a single fixed transmitter-receiver pair per station and time division multiple access for channel sharing will not have performance that differ significantly from ShuffleNet with $p = 2$ or $p = 3$. ShuffleNet with $p = 2$ is shown in Figure 2.1.

As we already pointed out, the throughput decreases very slowly with the increase in p when the number of stations is large. In such a case, it may be more beneficial to use higher p, since higher p reduces the number of hops, and therefore the delay. In addition to this, the increase in p reduces the number of wavelength channels since the number of channels is equal to N/p. Thus, by increasing connectivity we can allow more stations to be connected to the network if the number of wavelengths is limited.

In order to optimize both throughput and delay we introduce the power function P defined as a ratio of average throughput per station and average number
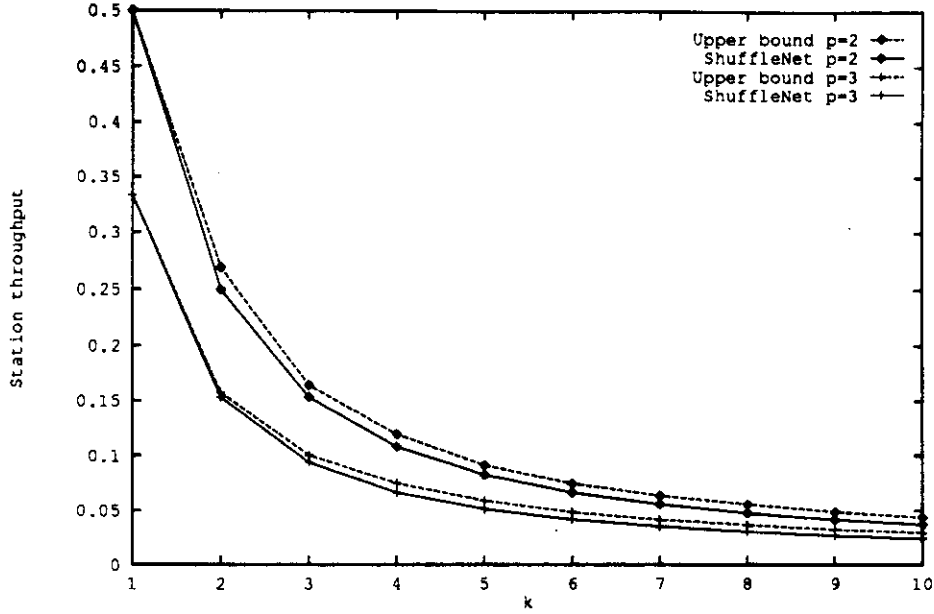
Figure 2.5: The maximum average throughput per station versus parameter k for ShuffleNet with stations' connectivity $p = 2$ and $p = 3$

of hops. Thus

$$P = \frac{\sigma}{E} \tag{2.10}$$

We have that

$$P \leq P_{ub} = \frac{\sigma_{ub}}{E_{lb}(p)} = \frac{C}{p(E_{lb}(p))^2} \tag{2.11}$$

Figure 2.6 shows the upper bound for the power function versus p for various numbers of stations N. We see that the optimal value p is higher than the one when the criteria was to maximize the throughput only.

Let us consider now how channel sharing affects the multihop network when each station has d transmitters and d receivers. Let s be the channel sharing factor which represents the number of virtual channels each transmitter and receiver is connected to. Thus, the station's connectivity factor is $p = ds$. The upper bound for aggregate throughput in this case is

$$L_{ub} = \frac{NCd}{pE_{lb}(p)} = \frac{NC}{sE_{lb}(p)} \tag{2.12}$$

and the upper bound for the mean throughput per station is

$$\sigma_{ub} = \frac{C}{sE_{lb}(p)} \tag{2.13}$$

33
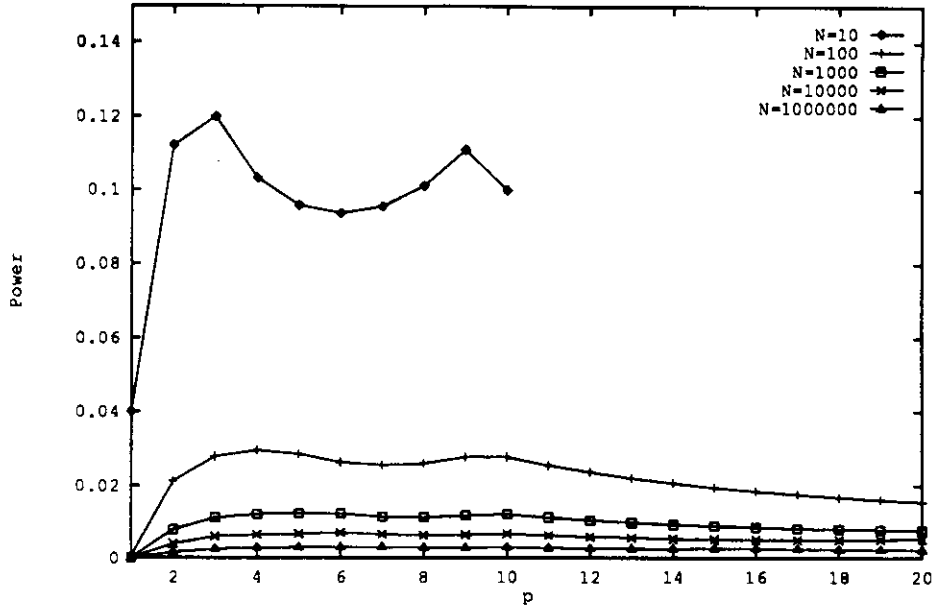
Figure 2.6: The upper bound for the power function versus p

Figure 2.7 shows how σ_{ub} changes with the increase in channel connectivity for different values of parameter d. We see that when $d > 1$ the best choice is to have $s = 1$, Thus, when the number of transmitters and receivers is greater than one, the highest throughput is achieved when no channel sharing is performed.

Figure 2.8 shows how P_{ub} changes with the increase in channel connectivity when $N = 1000$. We see from the figure that for $d = 1$, 2 and 3, channel sharing can improve the power.

2.2 Comparison of HONET with other networks

The advantage of the HONET architecture is that the bandwidth available for all traffic types is increased compared to the existing single-hop network architectures as illustrated in Figure 2.9. Efficient single-hop network protocols typically require two transmitter-receiver pairs per station, one for data and one for control traffic. A fixed transmitter and receiver are most often used for a control channel. In a single-hop network that supports packet-switching, its control channel carries a large amount of control traffic. In HONET we practically implement the control channel using a multihop network. As it is already shown, HONET's multihop network can be implemented with one fixed transmitter and receiver
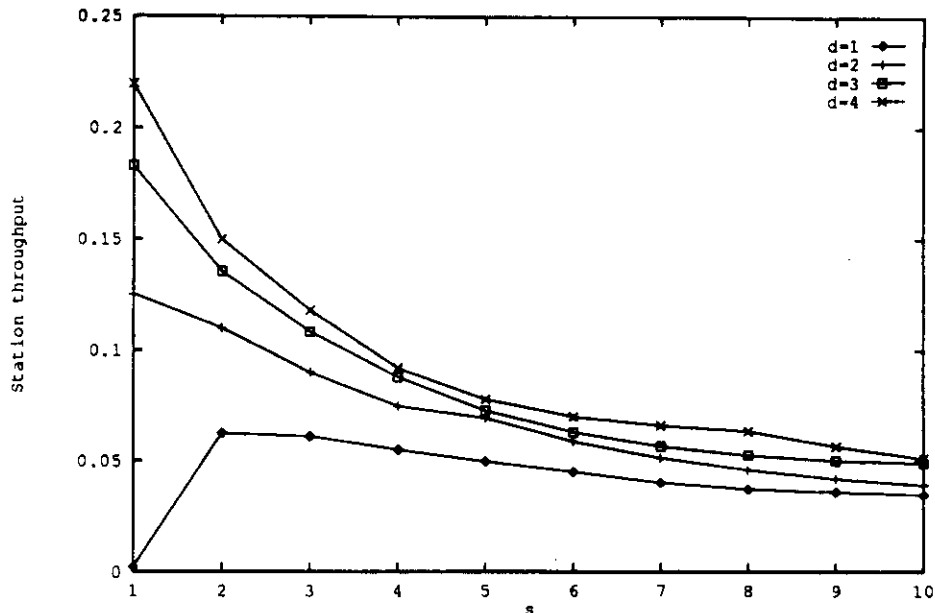
Figure 2.7: Upper bound for the average throughput per station versus channel sharing factor s when the number of stations is 1000 and the number of transmitter-receiver pairs per station d is 1, 2, 3, and 4.

per station. Thus, the cost of HONET is comparable to the cost of the classic single-hop network since they can be implemented with the same number of optical devices. As a difference from the control channel of the classic single-hop network which carries only control traffic, HONET's multihop network, in addition to the control traffic, carries also packet-switched traffic, thus freeing more capacity on the single-hop network. The amount of control traffic for HONET's single-hop network is significantly reduced since its data traffic is mostly connection oriented, and pretransmission coordination is done on a per connection basis. In order to estimate the amount of control traffic in HONET let us assume that the duration of a connection is exponentially distributed with mean duration of 100 seconds, and that the arrival rate of connection requests has Poisson distribution with mean of 100 connection requests per second. This results in a very high offered load of 10000 Erlangs (which practically represents the average number of connections in the system if there is no blocking). If each connection request requires exchange of 1 kb of control data, the aggregate control traffic is just 100 kb/s. Considering that throughput of the multihop network can be several Gb/s, we see that the control traffic uses a very small portion of HONET's multihop network resources. We can assume that practically all the capacity of

35
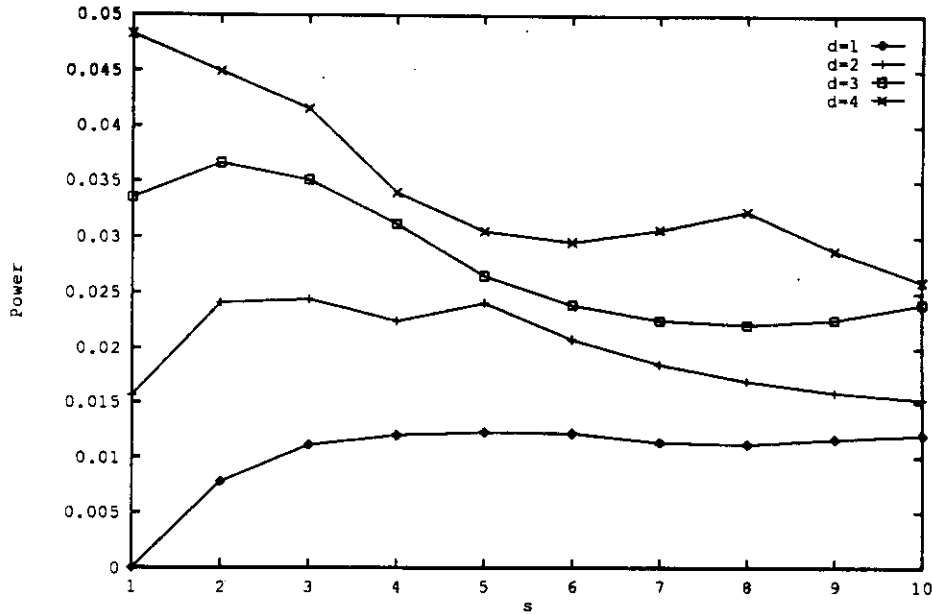
Figure 2.8: Upper bound for the power function versus channel sharing factor s when the number of stations is 1000 and the number of transmitter-receiver pairs per station d is 1, 2, 3, and 4.

HONET's multihop network is available for data traffic. Let us compare now the capacity of HONET and the single-hop network. Let us assume for the moment that the number of wavelengths that can be used by the single-hop and the multihop network is not limited. The maximum capacity of the single-hop network is NC where N is the number of stations in the network and C the capacity of a wavelength channel. In HONET, the maximum capacity is equal to the maximum capacity of the single-hop network which is also NC plus the maximum capacity of the multihop network which is in the case of the ShuffleNet topology (that uses $N/2$ wavelengths) approximately $\frac{NC}{2\log_2 N}$. Thus, we see that the maximum capacity of HONET is higher than the maximum capacity of the single-hop network. In reality, the capacity of the single-hop network is restricted by several factors, such as limited tuning range of fast tunable optical devices and limited amplification range of optical amplifiers if amplification is needed (in case of a large number of stations in the network). While the latter also holds for the multihop network, we should point out the implementation of the multihop network is not restricted to be on the single broadcast medium. In fact, the capacity of a multihop network can be much larger than the one of a single-hop network when the number of stations is large. As a result of this, the capacity of HONET can

36
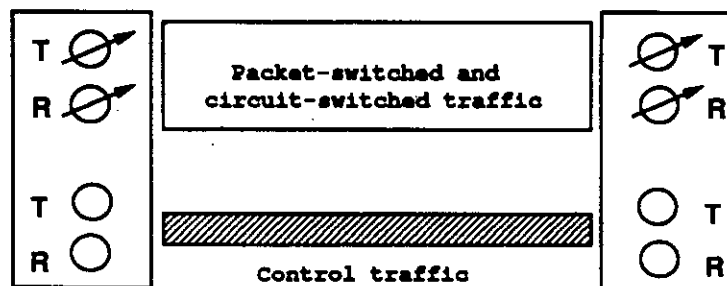
(a) Single-hop network
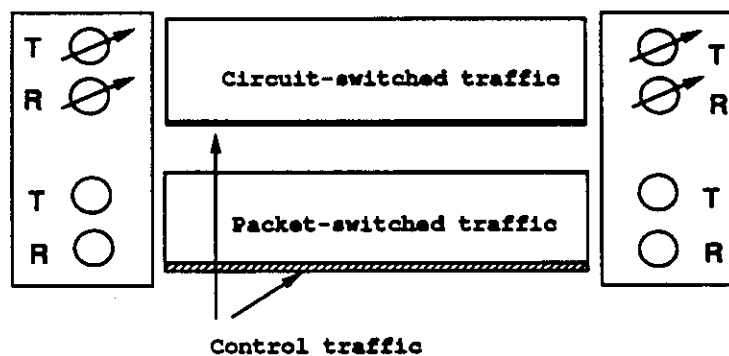
(b) HONET

Figure 2.9: Bandwidth of the classic single-hop network and HONET

be much higher than the capacity of the single-hop network.

Let us compare HONET with STARNET, the architecture that also combine a single-hop and a multihop network. In order to make a fair comparison to STARNET, let us consider the simplest HONET implementation which requires two fixed transmitter-receiver pairs, one for the single-hop network and one for the multihop network. Thus, the simplest HONET requires one more fixed transmitter than STARNET. Clearly, using a fixed transmitter receiver pair over the single-hop network reduces T/WDMA scheme to a pure TDMA. This seriously limits the capacity of the single-hop network. However, even though we can exploit much less optical bandwidth in such a network than in the STARNET single-hop network, we may, in fact, achieve better performance due to increased connectivity. In STARNET each station can communicate with only one other station at a time. The station is assigned an entire channel's bandwidth of, say,

37

1 Gb/s, and it uses the channel for a connection which may require a very small fraction of this bandwidth. We recall that most of real-time, connection oriented applications in use today require relatively low bandwidth. For example, the data rate for compressed video is only 1.5Mb/s and for voice 64kb/s. We also recall that a station can be a local area network gateway, or a mainframe, and thus, it may need to establish many simultaneous connections with other stations. In a STARNET with 500 stations up to 499 video connections can be supported while in a HONET with the same number of stations, the maximum number of video connections is more than 600. If the connections are multicast or broadcast, the number of connections that can be supported by STARNET is further reduced, while for HONET it remains the same. If we compare the multihop networks of HONET and STARNET we notice that we can have at least the same performance with HONET as with STARNET when the virtual topology is a ring. In HONET we are able to establish many different and more optimal virtual topologies by TDMA channel sharing. We should point out that implementing TDMA over the multihop network does not pose any additional cost, since the TDMA synchronization is provided by the single-hop network. As shown in Section 2.1, in case of uniform traffic the best candidate for the optimal virtual topology is ShuffleNet with degree 2 or 3 which provides significantly higher throughput than the ring when the number of stations is large.

CHAPTER 3

HONET: Circuit-switched Service with T/WDMA

3.1 Introduction

Ideally, the best T/WDMA performance is obtained with transmitters and receivers which can be retuned from slot to slot in negligible time. In reality, tuning time is finite and has direct impact on performance. One finds that system efficiency depends on the ratio of slot transmission time to tuning time and is thus dependent on slot size. The larger the slot the better the efficiency. Most applications, however, use only modest bandwidth and thus require small slots. For example, a voice connection at 64 kb/s requires 8 bits in a 125 μs frame. With 1 Gb/s channel speed the 8-bit transmission takes only 8 ns. Such a small slot size would require very fast tuning, on the order of a nanosecond. However, present optical devices of such speed have very narrow tuning range that can cover only a few wavelength channels. On the other hand, even if the tuning can be done infinitely fast, wavelength dispersion [SH93] requires to have idle periods between slots which can be as long as a couple hundreds of nanoseconds. If the required gap between slots is 100 ns, and the slot size 10 ns, at most only 10% of the network capacity can be utilized even with infinite tuning speed!

From the above discussion it is apparent that the major challenge in T/WDMA protocol design is the following: how to achieve efficient multiplexing of multirate circuits and/or packets with current available transmitter/receiver technology.

A possible solution to this problem is to use a large TDMA frame and/or large slots. Note that an increase in frame size will allow us to increase slot size as well, yet keeping the circuit rate constant. The use of large frame and large slots clearly simplifies the tuning problem. Since efficiency depends on slot time to tuning time ratio, the increase in slot size seems to be a step in the right direction. Indeed, some schemes have explored this approach. For example, in [HRS92] a design case is reported, which is based on a 1 ms frame subdivided into ten 100 μs data (packet or circuit) slots, for a 1 Gb/s channel speed. This means 100 kb packets and 100 Mb/s circuits. There are, however, several considerations

39

which restrict our choice of frame and slot size.

Let us first consider the choice of frame size. We recall that SONET frame size is 125 μs [BC89]. SONET is the physical layer infrastructure supporting ATM [Min89]. We also point out that the interconnection of the optical, T/WDMA network to SONET is highly desirable, in order to extend the circuit switched service to wide geographical areas. To this regard, we recall that both FDDI-II and DQDB [Sta93] use 125 μs frames to support the isochronous service in order to be SONET compatible. Likewise, interworking with SONET has been explored for HIPPI (High Performance Parallel Interface) systems operating at 800 and 1600 Mb/s [Ber92]. Given the need to interconnect to SONET, then the adoption of the 125 μs frame is certainly advantageous, in that it avoids fragmentation and reassembly of data units at the gateways interconnecting the optical star and SONET. Another reason against large frame sizes is packetization latency, which increases with frame size, and which may adversely affect some real-time applications. For voice connections, for instance, echo considerations discourage the use of frames larger than 1 ms.

Given that frame size is fixed to 125 μs, a proper slot size must be chosen. Here, we note that in a multirate traffic environment it would be difficult to efficiently utilize slots much larger than, say, 125 ns (i.e., 1 Mb/s circuit connection rate at 1 Gb/s channel rate). In fact, as previously discussed, even a 1 Mb/s rate is too high for voice. Acceptable efficiency could be restored, in this case, by using subrate multiplexing (i.e. transmission of one voice sample every 10 frames, say), or by concentration (i.e., multiplexing of several voice connections between the same source/destination pair).

It should be pointed out that the choice of relative small frame and slot sizes does not preclude efficient support of high speed circuits. The latter can be implemented by allocating multiple slots in the frame. Likewise, small slot size does not necessarily imply small packet size. A large packet can be mapped into several (not necessarily contiguous) slots, as is done for example in FDDI-II [Sta93]. This enables the implementation of integrated approaches where both packet and circuit service share the same T/WDMA frame. A major concern in those systems is the fact the very small packet size may imply very high header and processing overhead [HRS92].

In this chapter we propose a circuit-switched, T/WDMA solution which uses as a reference a 125 μs frame size (although, other frame sizes will also be explored for performance sensitivity purposes), and which is based on fixed transmitters and on tunable receivers with moderate tuning speed, such as those obtained with acoustooptic tunable filters. The choice of fixed transmitters and tunable

receivers was dictated by the need to support broadcast/ multicast applications efficiently. In addition to this, broadcasting is necessary in order to maintain proper synchronization among stations connected to the network. The rest of the chapter is organized as follows. In the second section we describe the T/WDMA scheme with subframe tuning. The slot assignment algorithm is described in the third section. In the fourth section we present the performance analysis of this scheme for uniform, single rate, point-to-point traffic. An approximate analytical model is developed and the results based on this model and simulations are presented. In the fifth section we modify the original scheme in order to exploit the multiwavelength filtering capability of acoustooptic tunable filters and show how this modification can improve performance. In the sixth section we study performance for one specific class of nonuniform traffic, such as traffic with localities.

3.2 T/WDMA scheme with subframe tuning

The main challenge in the design of a T/WDMA circuit-switching system is to achieve reasonable efficiency with small frame size and relatively large tuning times (i.e., the TDM slot transmission time is smaller than the tuning time). One way to compensate for the low transmission time to tuning time ratio is to use a pipelining technique, like the one proposed in [CY91]. The pipelining technique requires at least two (instead of one) tunable receivers at each station. Namely, while one receiver is receiving data, the other is being tuned to the wavelength for the next slot. This scheme allows us to make the slot size as small as the tuning interval. Thus, the slot size in a network with rapidly tunable devices can be as small as 10 ns (the minimum tuning time achieved in a laboratory today). Note, however, that pipelining alone cannot eliminate gaps between slots which are required due to wavelength dispersion.

The T/WDMA scheme proposed here is based on much slower acoustooptic devices (tuning speed of 10 μs). The reasons of this choice are: lower cost, simpler synchronization, and an order of magnitude larger tuning range (than achieved with nanosecond devices), which provides a potential for a significant increase in network capacity. Given this choice, pipelining alone is clearly not an adequate solution, since it does not allow fine grain multiplexing in the 125 μs TDM frame.

One basic strategy consists in subdividing a frame into subframes. We group a number of slots into a subframe and perform pipelining at the subframe level (i.e., from subframe to subframe, rather than from slot to slot). That is, stations perform retuning between subframes instead of between slots. With this

41

restriction, channel performance is not significantly affected by the relatively long tuning delays, as we shall see. In fact, the tuning time determines the minimum size of a subframe, but not the size of a slot within the subframe. We use the rule that the station retunes to another wavelength only after the current subframe ends. Therefore, the station cannot retune between two slots that belong to the same subframe. Figure 3.1 illustrates the concept. The figure shows a frame of

		Subframe 0				Subframe 1				Subframe 2				Subframe 3		
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
W0 (T0,T4)		D2		D2	D7	D4	D1		D1			D5		D6			
W1 (T1,T5)		D3	D5	D4	D6				D6		D3					D4	
W2 (T2,T6)		D1	D0	D0 D1	D1		D7	D7		D1	D1	D4		D3			
W3 (T3,T7)						D2	D3		D2	D2	D0	D0	D0	D1 D5 D7			D1

Figure 3.1: T/WDMA with subframes

16 slots organized into 4 subframes, each consisting of 4 slots. The system uses 4 wavelengths (W0, W1, W2 and W3) and has 8 stations. The transmitters are fixed, and the receivers tunable. Each transmitter is assigned a fixed wavelength on which it can transmit. If the number of stations is larger than the number of available wavelengths, some wavelengths must be shared by two or more transmitters. If we denote by Tk the transmitter of station k, then in our case the transmitters of stations 0 and 4 (T0 and T4) are tuned to W0, T1 and T5 to W1, T2 and T6 to W2 and, T3 and T7 to W3. Likewise, Dk denotes the receiver at destination k. A slot scheduled for transmission to destination k is marked with Dk. From the figure we see that in subframe 0 the receivers of stations 2 and 7 (D2 and D7) are tuned to W0; D3, D4, D5 and D6 to W1; and D0 and D1 to W2. Slot 2 on W2 is used for a multicast connection to D0 and D1. A destination tuned to a wavelength can receive from any station transmitting on this wavelength. For example, in subframe 0, D0 can receive on W2 from T2 or T6.

Let us now consider the arrival of a circuit-switched connection request between source 3 and destination 6. Since T3 transmits on W3, we need to allocate an empty slot on W3. We see that T3 and D6 are both idle in slots 0, 1 and 2 of subframe 0. However, since D6 is receiving in slot 3 on W1 and is not allowed to change wavelength during the subframe, we cannot allocate a slot for this connection during the first subframe. Such a case, when no slot can be allocated to a

connection because the destination is engaged on a different wavelength than the source, we call **wavelength conflict**. For the same reason we cannot allocate a slot for this connection in subframes 1, and 3 either (D6 is receiving on W1 in subframe 1 and on W0 in subframe 3). In subframe 2, D6 is available but there is no empty slot on W3. Thus, no slot can be allocated in any subframe, and the connection is blocked. As another example, let us consider a request for a connection between source 1 and destination 2. In this case the first available slot that can be allocated is slot 12 on W1.

In order to implement this scheme, each station needs a fixed transmitter and two tunable receivers. The receivers operate using a pipelining technique, as shown in Figure 3.2. The figure illustrates pipelining of receivers at station 1 for
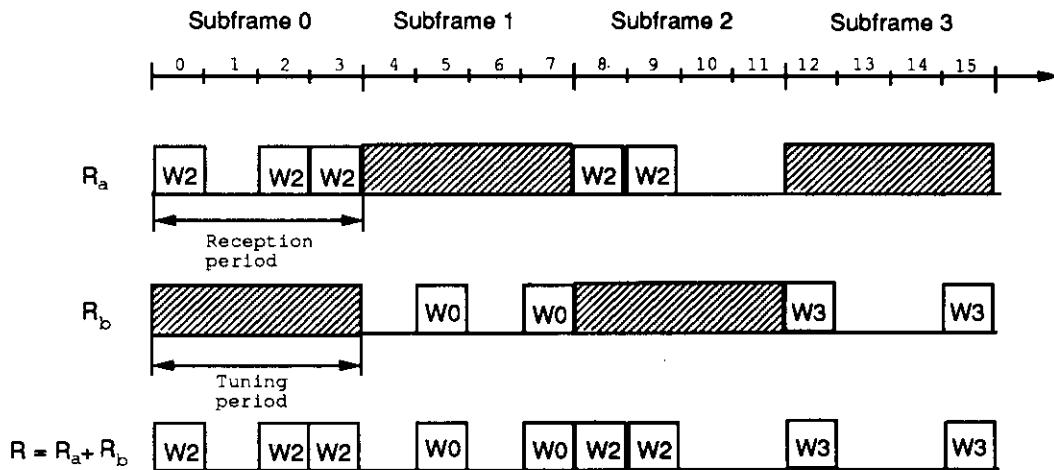
Figure 3.2: Pipelining of receivers Ra and Rb

the example from Figure 3.1. During a subframe one receiver is receiving while the other is retuning for the next subframe. Thus, only one receiver per station can be active in any subframe.

Since retuning is performed only between subframes, wavelength dispersion gaps between slots that belong to the same subframe can be eliminated. The gaps has to be placed only between subframes. If the number of slots is much larger than the number of subframes, the improvement in network efficiency can be significant. For example, if the wavelength dispersion gap is 100 ns, the number of subframes 10 and frame size 125 μs, up to 99% of network capacity can be used.

In order to support connections with different data rates, each subframe can be partitioned into slots of different sizes. For example, each subframe can con-

43

sist of a number of video slots, and some of them can be further partitioned
into a number of smaller slots to be used for voice connections as illustrated in
Figure 3.3.
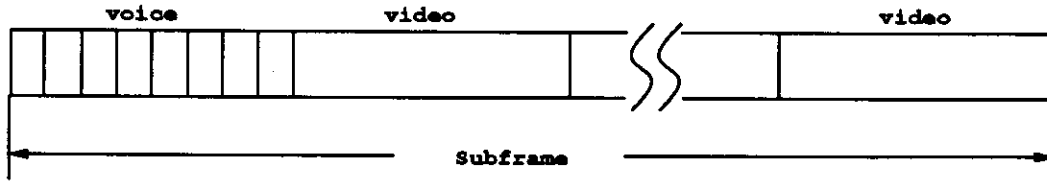
Figure 3.3: Voice and video slots in a subframe

The scheme proposed here assumes that pretransmission coordination is per-
formed over HONET's multihop network. It should be noted that "coordination"
does not have critical impact on performance in our case since it generates a very
small amount of control traffic if circuit-switched connections are assumed to last
several minutes and perhaps hours. Control traffic can be given higher priority
over data traffic in order to reduce the coordination delays. Although the coordi-
nation delays may not be negligible they are irrelevant in case of circuit-switching
since they are associated only with call set-up and release (e.g., establishing a
long distance phone call can take several seconds). In contrast, coordination de-
lays play a major role in T/WDMA networks used for packet switching or for
integrated packet and circuit-switching [HRS92].

3.3   Slot assignment algorithm

The objective of this assignment algorithm is to minimize the blocking probability
of future calls. The basic strategy is to pack as many connections as possible in
as few subframes as possible. This way, more subframes will be available for
unrestricted allocation of new connections. Specifically, for a connection request
to destination d that uses wavelength w the following slot assignment algorithm
is used:

Step 1:   If d is active on w in one or more subframes, select among these the
subframe that has an empty slot and with the largest number of slots already
allocated to d. If such a subframe is found, allocate one of its empty slots; else
go to Step 2.

44

Step 2: Select randomly a subframe in which d is idle and there is an empty slot. If such a subframe is found, activate d on w in this subframe and allocate one of its empty slots; else the connection is blocked.

This assignment algorithm tends to minimize the number of subframes used by a destination on each wavelength. Consequently, blocking probability is reduced, as compared to arbitrary assignment.

Upon termination of a connection to destination d on wavelength w, the slot used by the connection is released. If this was the only connection to d on w in the subframe, d is deactivated on w in this subframe.

The presented slot assignment algorithm can be implemented using centralized or distributed approach.

The centralized algorithm requires a single control station. Stations send the connection request via the packet-switched network to the control station which performs the algorithm and informs the stations about the slot assignment for the new connection by broadcasting via the single-hop network or by sending a packet (to the stations that are involved in the new connection). The central station keeps track of all the information necessary for executing the slot assignment algorithm.

In the distributed version of the algorithm, there are more than one control station. We can partition the existing circuit-switched traffic into point-to-point and broadcast/multicast. We can then assign separate sets of subframes to each of these traffic classes. The broadcast/multicast traffic can be handled by a single control station in the same manner as it is done for the point-to-point traffic using the centralized algorithm. For the point-to-point traffic the distributed algorithm requires W control stations where W is the number of wavelengths used in the network. The control station for wavelength w is one of the stations that transmit on w. Basically, each control station keeps track of which slots are busy on its wavelength. On the other hand, each destination keeps track of wavelengths it is using in each subframe. If a station s transmitting on w wants to establish a connection to station d, the following protocol is executed:

1. station s sends a connection request to its control station via the packet network.

2. the control station sends a packet which includes a list of idle slots on w to station d and waits for response from station d.

3. station d selects the slot to be used for the connection, according to the

algorithm specified above, sends the packet to the control station informing it about its choice and tunes to the selected subframe and slot. If there is no subframe that can be used for the connection, d notifies the control station that the connection cannot be established.

4. after receiving the packet from station d, the control station informs station s to start its connection.

Each connection request requires transmission of four packets (two packets in the case the source station is also the control station). This can be considered as a small control overhead if we take into account that each connection last a relatively long time.

3.4   Performance analysis

Since the proposed T/WDMA with subframe tuning is designed to support circuit-switched communications, our analysis will be focused on the typical performance parameters of a circuit-switched system, namely, blocking probability and network load. Note that the performance of the control network (i.e., HONET's multihop network carrying the control traffic for this scheme) has no impact on these performance parameters provided that the capacity of the control network is large enough to support control traffic (i.e., connection establishment requests). As we mentioned above, control network delay may affect the delay in establishing connections but it will not significantly impact the above mentioned performance measures, since the duration of a connection is much longer than the connection setup time.

3.4.1   Assumptions

In this performance study of a network of N stations we use the following assumptions:

- Point-to-point connections. We will limit our analysis to point-to-point traffic. It should be clear however that this scheme can support multicast and broadcast traffic as well.

- Uniform traffic. Source and destination stations are chosen randomly using the uniform distribution.

- Poisson arrivals, exponential service times. Each station generates call requests according to a Poisson process. The average duration of a connection

46

is exponentially distributed. The mean time between connection setups and the average duration of a connection are expected to be much longer than the frame length.

- No queueing of connection requests. If a connection is blocked it is immediately discarded.

- Uniform circuit connection data rates. For simplicity, we assume that a connection requires only one slot per frame (i.e., no multirate applications). However, the analysis could be extended to handle also multirate traffic.

- T/WDMA carries only user traffic. For the sake of simplicity, we assume that only user traffic (i.e., no control traffic) is transmitted over the T/WDMA network.

- Balanced wavelength allocation. The assignment of wavelengths to stations' transmitters is determined using the following formula:

$$t_s = s \bmod W$$

where s is the station number (stations are numbered from 0 to $N-1$).

3.4.2 Analytical model

The exact analytical model for the scheme described above is very complex. We develop an approximate analytical model which is, as it will be shown later, accurate enough when the number of stations is large (100 or more). Since the scheme is used for a metropolitan area network, the accuracy of the approximate model should be sufficient in most cases of practical interest. We use an iterative procedure to solve this analytical model.

Let us first introduce some notation:

N Number of stations.

W Number of wavelengths.

S Number of subframes in a frame.

K Number of slots in a subframe.

λ Aggregate arrival rate.

T Average duration of a connection.

L Average offered load $(L = \lambda T)$.

We begin our analysis with calculating the average duration of a period in which a destination is receiving on wavelength w in the TDM frame. In order to find it we model the number of slots used by a destination on wavelength w as a continuous time Markov process shown in Figure 3.4. The arrival rate to the
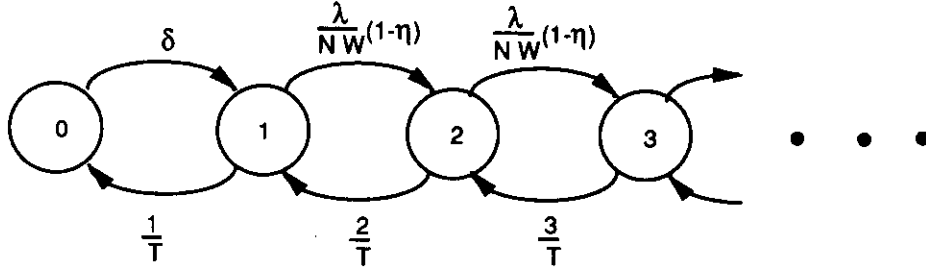
Figure 3.4: Number of slots used by a destination on wavelength w

destination from a station whose transmitter uses wavelength w is $\frac{\lambda}{NW}$. We use here an approximation that the number of stations assigned to each wavelength is $\frac{N}{W}$ which is true only if this number is an integer. η represents the probability that there is no available slot where the connection can be placed. Note that η can be assumed to be independent of the number of slots in use, due to our large population assumption. We will determine this probability later. We make an approximation using an infinite number of states model, since number of slots used by a destination cannot exceed the total number of slots at each wavelength. We denote the arrival rate in state 0 as δ. The actual value δ is not relevant in this analysis, as we will show soon.

Let p_i be the probability that a destination is using i slots on w. Solving for p_0 we have

$$p_0 = \frac{1}{1 + \delta T \frac{NW}{\lambda T(1-\eta)}(e^{\frac{\lambda T}{NW}(1-\eta)} - 1)} \tag{3.1}$$

Using this result we can find the average busy period B (i.e. the average duration of a period during which a destination is receiving on w). We have that

$$p_0 = \frac{I}{I + B} \tag{3.2}$$

where I is the average idle period. We also have

$$I = \int_0^\infty e^{-\delta} dt = \frac{1}{\delta} \tag{3.3}$$

48

Using Equation (3.1), (3.2) and (3.3) we get the following expression for the average busy period:

$$B = \frac{1 - p_0}{p_0} I = T(e^{\frac{L(1-\eta)}{NW}} - 1)\frac{NW}{L(1 - \eta)} \tag{3.4}$$

We now calculate the probability that a receiver is using i wavelengths simultaneously — π_i. In order to determine this probability we simplify the model by assuming that the destination uses at most one subframe per each wavelength, which is a reasonable approximation if the number of stations is large. In such a case, each destination has a small number of connections on each wavelength, and the probability that it requires more than one subframe to accommodate these connections is small. We also approximate the period during which a destination uses a wavelength (i.e., busy period B) as exponentially distributed. This allows us to model the number of wavelengths (i.e., subframes) simultaneously used by the destination as a continuous time Markov process shown in Figure 3.5. γ_i
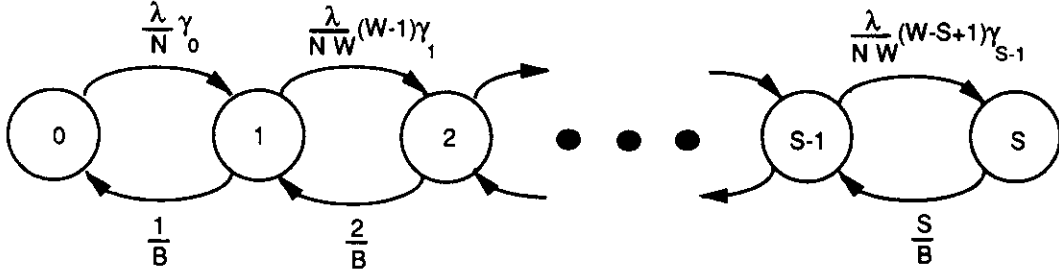
Figure 3.5: Number of wavelengths used by a destination

represents arrival rate in state i. If destination d is in state i, this means that d is using i subframes. Thus, a connection on a new wavelength w can be allocated if there is an empty slot in any of $S - i$ subframes not used by d. Let α_i be the probability that all slots in i subframes on this wavelength are full. We have then

$$\gamma_i = \begin{cases} \frac{\lambda}{NW}(W - i + 1)(1 - \alpha_{S-i}) & i \leq W \\ 0 & i > W \end{cases} \tag{3.5}$$

Let us now calculate the probability α_i. Since the arrivals to the system have Poisson distribution and a subframe is chosen randomly with uniform distribution, in order to calculate α_i we can use an assumption that arrivals to each subframe have Poisson distribution with the same arrival rate. The arrival rate to i subframes is thus equal to the sum of arrival rates to each of these i subframes.

We now proceed with determining the arrival rate to a subframe on wavelength w. If the number of wavelengths W is greater than number of subframes S, and

49

the number of wavelengths (subframes) used by a destination is S, any connection request arrival requiring one of $W - S$ wavelengths not used by the destination will be immediately discarded. Thus, these arrivals do not contribute to the arrival rate to any of the subframes. The effective arrival rate to a subframe can be calculated as

$$\sigma = \frac{\lambda}{SW}(1 - \pi_S(\frac{W - S}{W}))$$ (3.6)

The blocking probability α_i is

$$\alpha_i = P_{loss}(\sigma i T, K i)$$ (3.7)

where P_{loss} is Erlang's loss formula defined as

$$P_{loss}(x, n) = \frac{\frac{x^n}{n!}}{\sum_{i=0}^{n} \frac{x^i}{i!}}$$ (3.8)

γ_i is the probability that there is an empty slot in $S - i$ subframes not used by the destination. Thus,

$$\gamma_i = 1 - \alpha_{S-i}$$ (3.9)

Solving for π_i we have

$$\pi_i = \frac{\frac{B^i}{i!} \prod_{k=0}^{i-1} \gamma_k}{1 + \sum_{j=1}^{S} \frac{B^j}{j!} \prod_{k=0}^{j-1} \gamma_k}$$ (3.10)

The blocking probability can be expressed as

$$P_B = \sum_{i=0}^{S} q_i \pi_i$$ (3.11)

where q_i is the conditional probability that a connection is blocked given that the destination is already using i wavelengths (subframes). Given that i subframes are used, a new connection on w will be blocked if: a) w is not used by the destination and all slots in other $S - i$ subframes are full or b) if the destination is receiving on w in one of these subframes but all slots in that subframe and remaining $S - i$ subframes are full. The blocking probability in the former case is equal to the probability α_{S-i}. In the latter case the blocking probability β_{S-i} has to be calculated using different approach since the probability of an arrival to a particular subframe is not independent of the number of destinations receiving in this subframe.

Thus, the probability q_i can be expressed as

$$q_i = \begin{cases} \alpha_S & i = 0 \\ \frac{W-i}{W}\alpha_{S-i} + \frac{i}{W}\beta_{S-i} & 0 < i < S \\ \frac{W-S}{W} + \frac{S}{W}\beta_0 & i = S \end{cases}$$ (3.12)

50

We now proceed with calculating the probability β_i. For a subframe on wavelength w we have that the conditional probability that a call arrives in the subframe given that j destinations are assigned to the subframe is

$$P[\text{call arrives}|j \text{ destinations}] = \frac{j}{N} \tag{3.13}$$

A destination is assigned to the subframe on w if it is already receiving or it is about to receive in this subframe. Let θ be the probability that a destination is assigned to the subframe. We assume that this probability is independent of the number of destinations using the subframe which is actually an approximation. In reality, if the number of destinations is larger, it is more likely that all slots in the subframe are full. In such a case, destinations will try to place their calls in a subframe that has an empty slot.

$$\theta = \frac{1}{S}[(1 - \pi_S) + \frac{S}{W}\pi_S] \tag{3.14}$$

The probability that j destinations are assigned to the subframe is given by the Bernoulli distribution

$$P[j \text{ destinations}] = \left(\begin{array}{c} N \\ j \end{array} \right) \theta^j(1 - \theta)^{N-j} \tag{3.15}$$

Using expressions (3.13) and (3.15) we have

$$P[\text{call arrives}, j \text{ destinations}] = \frac{j}{N} \left(\begin{array}{c} N \\ j \end{array} \right) \theta^j(1 - \theta)^{N-j} \tag{3.16}$$

and

$$P[\text{call arrives}] = \sum_{j=0}^{N} \frac{j}{N} \left(\begin{array}{c} N \\ j \end{array} \right) \theta^j(1 - \theta)^{N-j} = \theta \tag{3.17}$$

The a posteriori probability that there are j destinations assigned to the subframe given that a call for this subframe arrives is

$$P[j \text{ destinations}|\text{call arrives}] = \frac{1}{\theta}\frac{j}{N} \left(\begin{array}{c} N \\ j \end{array} \right) \theta^j(1 - \theta)^{N-j} \tag{3.18}$$

Given that j destinations are assigned to the tagged subframe, the effective load to one of the remaining $S - 1$ subframes on wavelength w is

$$\Omega_j = \frac{L}{(S - 1)W}\frac{N - j}{N}[(1 - \pi_S) + \frac{S - 1}{W - 1}\pi_S] \tag{3.19}$$

The conditional probability that a call finds all slots full in the assigned subframe and i other subframes given that j destinations are assigned to the subframe is

$$\beta_i^{(j)} = P_{loss}(\frac{L}{NW}j + \Omega_j i, K(i+1)) \tag{3.20}$$

The probability that a call finds all slots full in the assigned subframe and i other subframes is then

$$\beta_i = \sum_{j=0}^{N} \beta_i^{(j)} P[j \text{ destinations}|\text{call arrives}] \tag{3.21}$$

Using Equation (3.18) and (3.20) we get

$$\beta_i = \frac{1}{\theta N} \sum_{j=1}^{N} \binom{N}{j} j(\theta)^j (1-\theta)^{N-j} P_{loss}(\frac{L}{NW}j + \Omega_j i, K(i+1)) \tag{3.22}$$

Finally, we can calculate probability η using probabilities β. We have

$$\eta = \frac{1}{1-\pi_0} \sum_{i=1}^{S} \beta_{S-i} \pi_i \tag{3.23}$$

Using the expressions derived so far and applying the following iterative procedure we can compute the system blocking probability. We define $\pi_i(n)$, $\alpha_i(n)$, $\beta_i(n)$, $q_i(n)$, $\eta(n)$ and $P_B(n)$ as the values obtained for π_i, α_i, β_i, q_i, η and P_B at the end of the nth iteration. We start with some initial values $\pi_S(n)$, $\eta(0)$ and $P_B(n)$. One simple initial estimate is to set $\pi_S(0) = 0$ and $\eta(0) = P_B(0) = P_{loss}(L/W, SK)$ using formula (3.8). We then apply the following iterative procedure:

Step 1: Let $n = 1$.

Step 2: Calculate $\alpha_i(n)$ using expression (3.7). Find $\beta_i(n)$ using expressions (3.14), (3.19) and (3.22).

Step 3: Calculate $\pi_i(n)$ using expression (3.10). Find $q_i(n)$ using expression (3.12).

Step 4: Calculate $\eta(n)$ using expression (3.23). Find $P_B(n)$ using expression (3.11).

Step 5: If the difference between $P_B(n)$ and $P_B(n-1)$ is smaller than a threshold value, stop. Otherwise, set $n = n + 1$ and go to Step 2.

Let us know analyze the case where the number of slots per subframe K is infinite, which gives the lower bound for blocking probability for given load, number of subframes and number of stations. This is, in fact, the probability that a wavelength conflict occurs. In such a case, the probability that a subframe is full is zero, and consequently, the probability that a destination uses more than one subframe per wavelength is also zero. Thus, we have $\eta = 0$, α_i, $\beta_i = 0$ for $i = 1, \ldots, S$, $q_i = 0$ for $i < S$ and $q_S = (W - S)/W$. In such a case we can obtain a relatively simple closed form solution for blocking probability. The formula for blocking probability reduces to

$$P_B^{lb} = \lim_{K \to \infty} P_B = \begin{cases} \dfrac{W-S}{W} \dfrac{\dbinom{W}{S}(e^{\frac{L}{NW}} -1)^S}{\sum_{j=0}^{S} \dbinom{W}{j}(e^{\frac{L}{NW}} -1)^j}, & S \leq W \\ \\ 0, & S > W \end{cases} \tag{3.24}$$

As shown later, P_B^{lb} is a monotonically increasing function of W for $W \geq S$ where S is the number of subframes per frame. Thus, for $K = \infty$ the performance degrades with the increase in W which is rather unexpected. However, this can be explained by the fact that for $W > S$ a call is blocked because the destination is receiving on wavelengths different from that of the transmitter. An increase in W decreases the probability of both transmitter and receiver being on the same wavelength, and thus increases the blocking probability. We also see that blocking probability does not depend on L but on the ratio L/N, i.e. the average load per station. Thus, for fixed blocking probability, the total load grows linearly with the increase in the number of stations.

Equation (3.24) can be also used for the case when the number of wavelengths is large. The increase in the number of wavelengths results in the decrease in arrival rate per subframe on a wavelength and in the decrease of the probability that a subframe is full. For W large enough this probability becomes close to 0.

The blocking probability approaches asymptotically the constant P_B^* with the increase in number of wavelengths.

$$P_B^* = \lim_{W \to \infty} P_B = \frac{\frac{(\frac{L}{N})^S}{S!}}{\sum_{j=0}^{S} \frac{(\frac{L}{N})^j}{j!}} = P_{loss}(\frac{L}{N}, S) \tag{3.25}$$

We see that this result is in fact Erlang's loss formula where the offered load is $\frac{L}{N}$ and the number of servers is S. This can be explained by observing that for W close to infinity the probability that a destination is receiving more than one connection on the same wavelength is close to zero. Therefore, each destination

has only one connection per subframe, which implies that it can have at most S connections simultaneously. The offered load to each destination is $\frac{1}{N}$ of the total offered load if the traffic is uniform.

3.4.3 Results

We present here the results obtained using presented analytical model and by simulation. Assuming that frame size is 125 μs and tuning speed (i.e., minimum subframe size) is 10 μs, we can partition the frame into 10 subframes. Figure 3.6 shows the blocking probability versus the number of wavelengths W for 10 subframes, 120 stations, offered load $L = 5000$ Erlangs and various values of K (i.e., number of slots per subframe). In this and later experiments we use very high loads which are not typical of normal operation and which account for high blocking probabilities. We do this in order to increase the accuracy of the results for a given simulation run time. This approach is justified by the fact that our main purpose is the relative comparison with other schemes, and the study of sensitivity to parameter changes. The blocking probability will be much lower under normal operating loads. The results obtained using the analytical model and simulations are represented with curves and points, respectively. It can be seen that the analytical model matches closely the simulation results. All the curves in Figure 3.6 except the curve for $K = 1$ (which represents the classic T/WDMA where slot size is equal to subframe size) show an optimal value for W where the blocking probability is minimal. For $W > S$ all the curves converge to the envelope which is in fact the lower bound P_B^{lb} from Equation (3.24). We see that at the optimal operating point (i.e., optimal value of W) the performance improves with the increase in K, i.e., the number of slots per subframe. However, with the increase in W, the large K advantage diminishes. For example, we see from the figure that for $W \geq 60$ the blocking probability is the same for any $K \geq 10$. Also, the optimal operating value W decreases with the increase in K. The blocking probability reduction caused by an increase in K (i.e., decrease in slot size) can be explained by the fact that with a large number of small slots in a subframe it is more likely to find a free slot and therefore "piggyback" a new call on a subframe on which the receiver is already tuned. This effect, coupled with the low value of W tends to reduce blocking. It should be noted, however, that there is a practical lower bound on slot size, and therefore upperbound on K, due to the need of guardbands between slots. In our example, the largest value $K = 100$ implies a slot size of 125 ns, which can be still comfortably implemented.

In Figure 3.7 the number of stations is increased ten times to $N = 1200$ and the load is doubled to $L = 10000$. Thus, the average load per station is decreased

Figure 3.6: Blocking probability for $L = 5000$ and $N = 120$. The number S of subframes is 10, and the number K of slots per subframe ranges from 1 to 100

five times. We see a significant reduction in the blocking probability because of a reduction in the wavelength conflicts. This is explained by the fact that the average load per station decreases with an increase in the number of stations. We also see an increase in the optimal W. Overall, we note that an increase in the number of stations improves performance. If the number of stations were infinite, P_B^{lb} would be zero (i.e., no wavelength conflicts) for all values of W. In such a case, all the curves would be monotonically decreasing with W.

Figure 3.8 shows the blocking probability versus number of wavelengths for fixed requested bandwidth E (i.e., the average number of connections L times the data rate of a connection) for different connection rates (which are expressed in slots per connection per frame, C). We see that when W is small, using low rate applications gives better performance as a result of better statistical multiplexing. However, when W is large, using smaller number of high rate applications gives much better performance. This is because for large W the limiting factor is not the network capacity but the probability of wavelength conflicts, which increases with the increase in the number of connections.

In previous examples we have studied how blocking probability changes while keeping the offered load constant. It seems more interesting for practical purposes
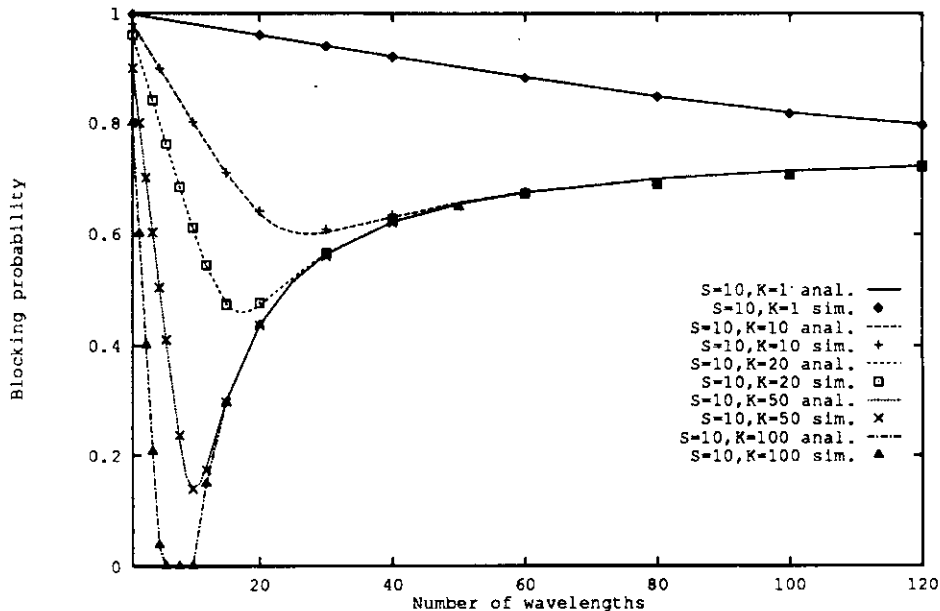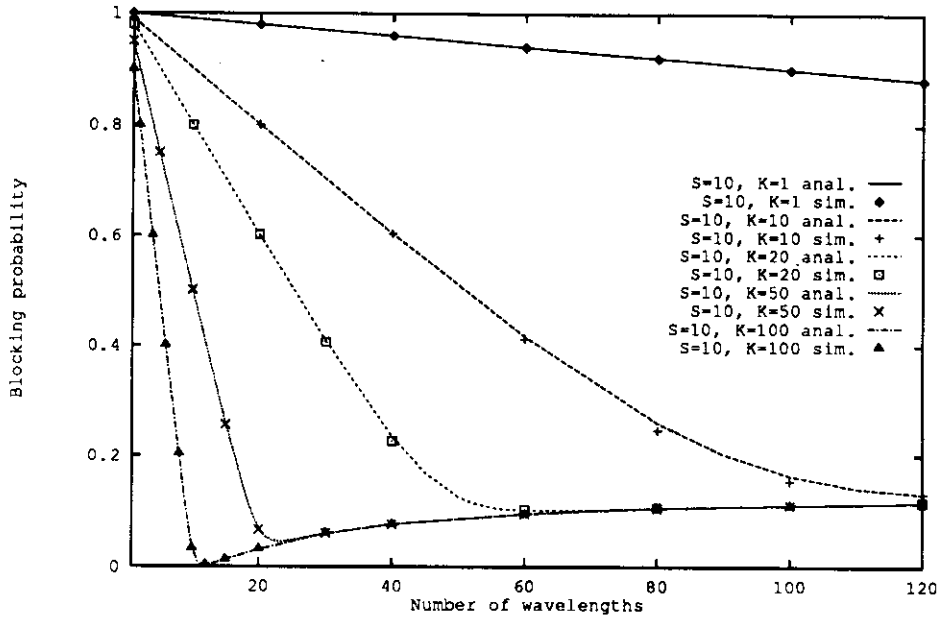
Figure 3.7: Blocking probability for $L = 10000$ and $N = 1200$. The number S of subframes is 10 and the number K of slots per subframe ranges from 1 to 100

to study offered load while keeping blocking probability (i.e., quality of service) fixed. It should be noted that obtaining data for such a study using simulation is a difficult task. The input to the simulation program is offered load, not blocking probability, so we have to guess the load which will result in required blocking probability. Since blocking probability is a monotonically increasing function of load, we can develop an iterative method that finds required blocking probability with specified accuracy in a finite number of steps (i.e., simulations). We apply this iterative procedure to our analytical model, which requires significantly less computation time than the simulation. Figure 3.9 shows the offered load versus the number of wavelengths for a fixed number of slots (i.e., 1000 slots) in the frame and different subframe sizes. Thus, the product $S \times K$ is fixed. The blocking probability is fixed to 0.01 and the number of stations is 1200. It can be seen that performance improves with the increase in the number of subframes. The increase in the number of subframes can be achieved by reducing the subframe size. For the fixed frame size the subframe size can be reduced either by increasing the tuning speed of optical devices or using additional interleaved receivers in the pipelining technique. Since the tuning speed is determined by the receiver technology (the tuning speed of acoustooptic filters is limited by the time necessary to set up the acoustic wave within the filter), this approach can make only minor increase in the

56
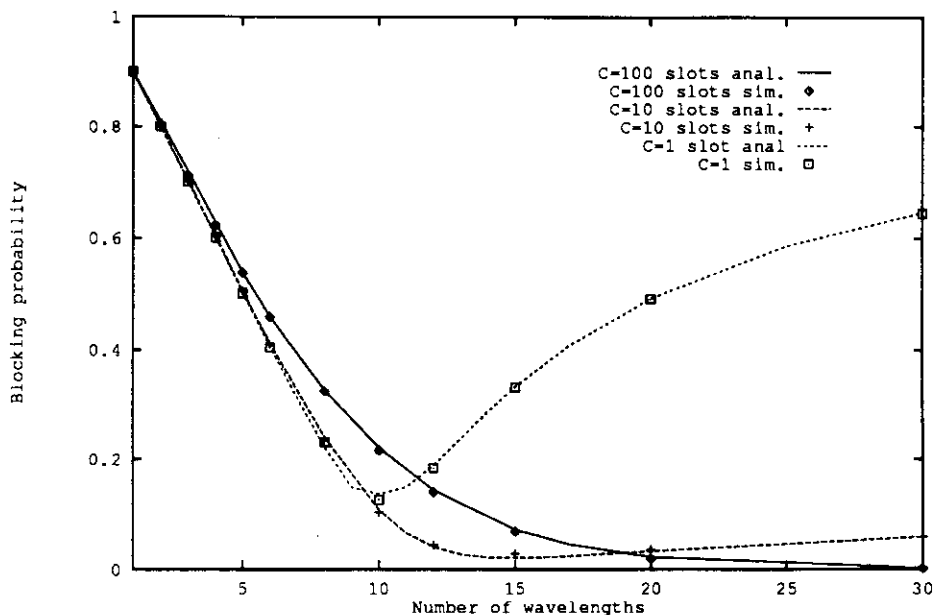
Figure 3.8: Blocking probability for $E = 10000$ (10 Gb/s), $S = 10$, $K = 100$ and $N = 120$. The number C of slots per connection ranges from 1 to 100.

number of subframes unless a different receiver technology is used. Figure 3.10 illustrates the pipelining technique with three tunable receivers per station. Note that by using r receivers per station (where $r > 1$) we effectively make the subframe length $r - 1$ times smaller than the tuning period, which in turn allow us to increase the number of subframes per frame by the same factor. The number of subframes could also be increased if we were allowed to increase the frame size. For example, we could achieve $S = 50$ with 10 μs acoustooptic devices by increasing frame size from 125 μs to 625 μs. The optimal value for W increases with the increase in S. This figure can illustrate the potential of this T/WDMA scheme. With 50 subframes and 20 slots per subframe the offered load at the blocking probability is around 70000 when the number of wavelengths is 80. The case when $S = 1000$ and $K = 1$ corresponds to the classic T/WDMA scheme, where wavelengths can be changed after each slot. Note that in this case the classic T/WDMA cannot operate for W greater than 10, since it requires fast tunable semiconductor devices (i.e. less then 100 ns tuning time) whose tuning range at present is limited to a few nanometers. We see that in most of other cases we can consistently achieve loads much higher than 10000 by choosing the optimal number of wavelengths. Thus, we can improve performance using our proposed T/WDMA scheme with slower acoustooptic devices that have much
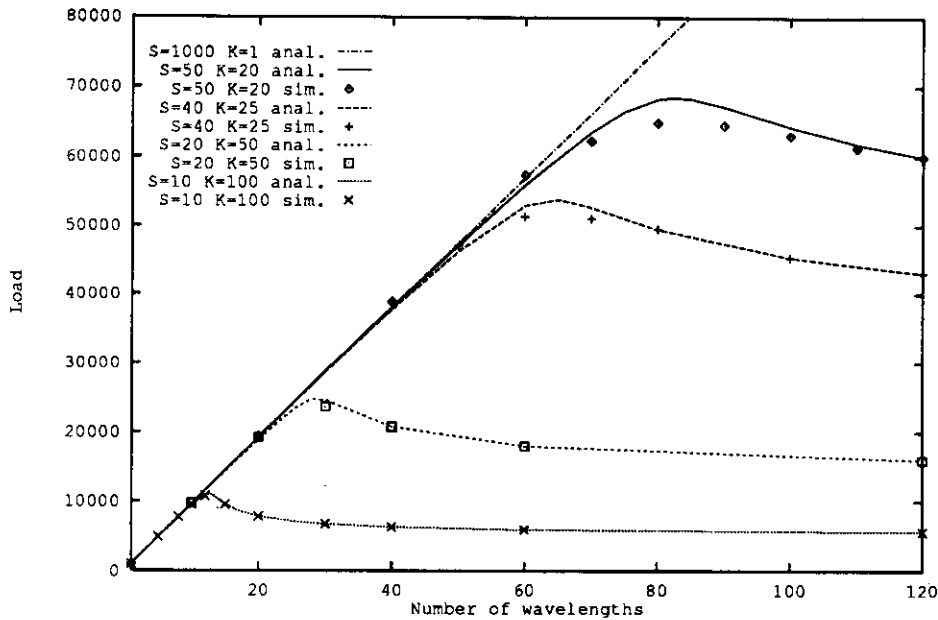
Figure 3.9: Offered load for $Pb = 0.01$ and $N = 1200$. The number S of subframes ranges from 10 to 1000, and the number K of slots per subframe from 1 to 100. The number of slots per frame $S \times K$ is held constant at 1000.

larger tuning range. In other words, the larger tuning range more than offsets the lower tuning speed.

Let us now analyze performance for the case when only one receiver (instead of two) is used per station. In such a case, each destination is able to receive in every other subframe, i.e., 50% of time (the other 50% of the time its receiver is retuning). In our analysis we assume that half the destinations are able to receive in even numbered subframes and the other half in odd numbered subframes. In such a case the system is practically partitioned into two separate subsystems, one consisting of even numbered subframes and the other one of odd numbered subframes. In each subsystem only half of stations are able to receive. The load to each subsystem is half of the total load. Thus, we can analyze the one receiver case using our analytical model if parameters L, S and N are reduced by two. Figure 3.11 shows a performance comparison for the cases when one and two receivers per station are used. We notice that using two receivers instead of one improves performance. One may verify that the effect of doubling the number of receivers (from one to two) is similar to doubling the number of subframes (while keeping the total number of slots in the frame constant). This similarity can be explained by observing that, when each destination is using on average only a

58
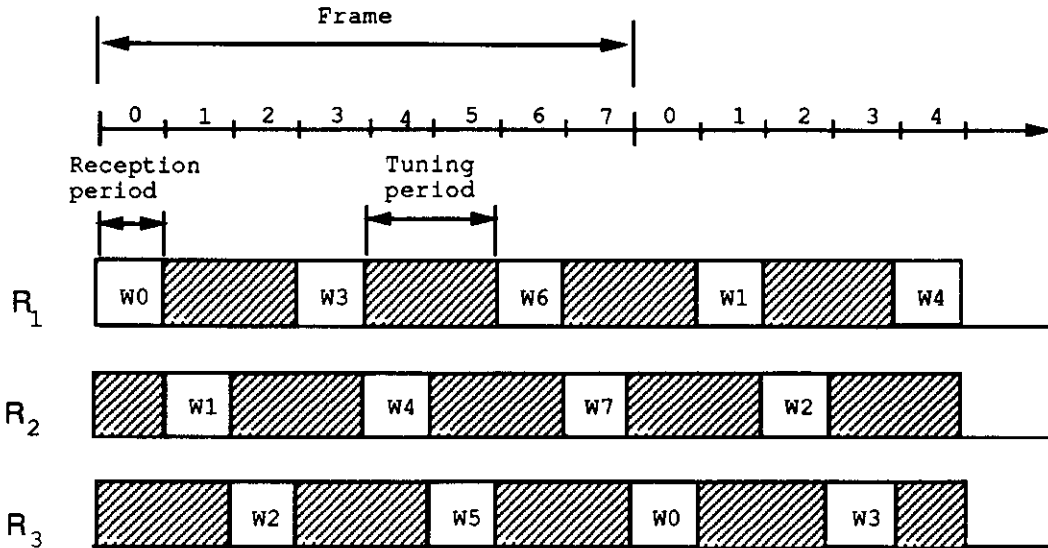
Figure 3.10: Pipelining with three receivers per station

small portion of the optical channel bandwidth, it makes little difference whether its capacity is equal to the capacity of the optical channel (if two receivers are used) or just half that capacity (if one receiver is used). From these results it appears that it would be more reasonable to use one receiver per station instead of two if we were able to double the number of subframes. If this scheme is used only for point to point traffic, this is indeed true. However, using a single receiver per station would make transmission of multicast/broadcast traffic difficult or even impossible since for such traffic all destinations must be able to receive in the same subframe. Given the importance of multicast and broadcast support, we opt for the solution which uses two receivers per station. In the rest of this chapter we will consider only the two receivers case.

3.5 T/WDMA scheme with multiwavelength selective receivers

As can be seen from the previous results, performance may degrade significantly when the number of wavelengths is larger than the number of subframes. This performance degradation is a result of blocking caused by the wavelength conflict problem (the situation when a call cannot be allocated because the destination is already tuned to another wavelength). For example, if we have 120 stations and allocate one wavelength to each of them (i.e., W=120), and partition the frame
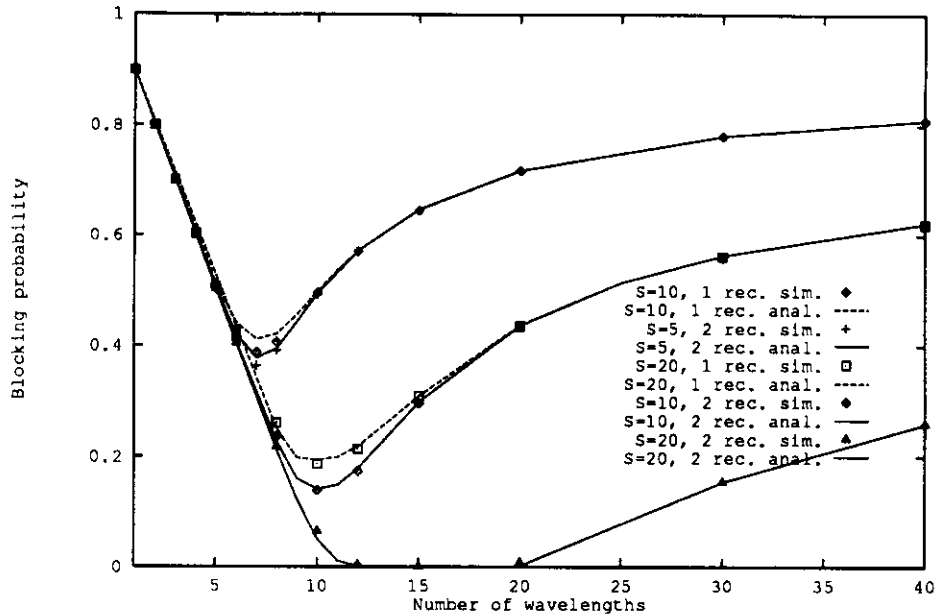
Figure 3.11: Blocking probability for $L = 5000$, $S \times K = 500$, $N = 120$ and 1 or 2 receivers per station

into 10 subframes and 100 slots per subframe, the blocking probability will be 0.72 if the load is 5000 Erlangs (see Figure 3.6). Thus, only around 1400 slots (i.e., calls) can be allocated on the average even if the total number of slots is 120000 (which is the upper bound on the maximum possible number of simultaneous connections). Thus, only around 1% of the total number of slots is used. One solution to reduce the blocking probability caused by the wavelength conflict problem is to increase the number of subframes, as we mentioned earlier. Here we propose another approach, which exploits the multiwavelength filtering capability of acoustooptic tunable filters [Che90] and does not require any additional hardware.

Figure 3.12 shows a receiver that uses an acoustooptic tunable filter and direct detection. The acoustooptic tunable filter selects a number of wavelengths which are forwarded to a broadband photodetector. The selection of up to five arbitrary wavelengths has been reported [CLS89], and it is likely that this number will increase in the future. Since direct detection is used, the receiver receives a superposition of signals on the selected wavelengths. Clearly, if more than one of the selected wavelengths carries a transmission simultaneously, a collision will occur at the receiver.

60
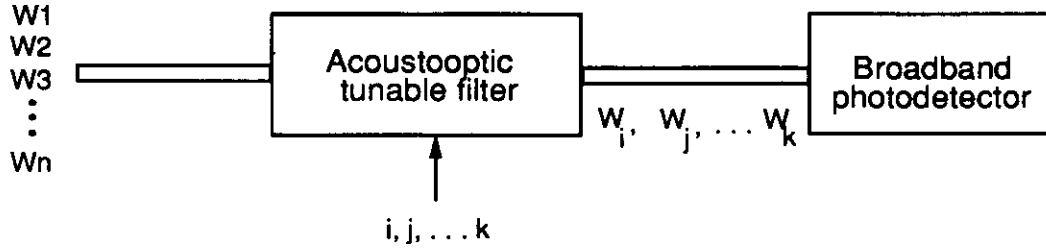
Figure 3.12: A receiver with multiwavelength selectivity

Using multiwavelength selectivity we can allow a destination to be tuned on more than one wavelength simultaneously during a subframe, thus reducing blocking caused by wavelength conflicts. However, if a destination is tuned on more than one wavelength during a subframe, we must make sure that, when the destination is receiving on one of the wavelengths, no station is transmitting on the remaining wavelengths during the same slot, in order to avoid collisions. Thus, slots on all the selected wavelengths must be reserved, or, more precisely blocked from access by other stations. A slot remains blocked for the entire duration of the connection. Therefore, the new scheme reduces blocking caused by wavelength conflicts at the expense of wasting extra slots. When wavelength conflict is the predominant cause of blocking, most of the slots are not used anyway. Thus, the waste of extra slots has only a minor adverse effect on performance while the reduction in wavelength conflicts has a major beneficial effect, as we shall show later.

In order to estimate the performance gain from using multiwavelength selectivity let us calculate the blocking probability for the case when number of slots per subframe is infinite, and maximum number of wavelengths selected is W_m. In this case in each subframe a destination can receive on up to W_m different wavelengths. Since the number of subframes is S, the maximum number of wavelengths a destination can receive on is SW_m. Replacing S with SW_m in expression 3.24 we get the following result for the blocking probability.

$$
P_B^{lb} = \lim_{K \to \infty} P_B = \begin{cases} \dfrac{W - SW_m}{W} \dfrac{\binom{W}{SW_m} (e^{\frac{k}{NW}} - 1)^{SW_m}}{\sum_{j=0}^{SW_m} \binom{W}{j} (e^{\frac{k}{NW}} - 1)^j}, & SW_m \leq W \\[4mm] 0, & SW_m > W \end{cases} \tag{3.26}
$$

This result is in fact a lower bound for blocking probability for the multiwavelength selectivity case. Thus, using wavelength selectivity of degree W_m has practically the same effect on the lower bound, as W_m times increase in the

61

number of subframes.

The freedom of retuning during a subframe, however, requires to reintroduce wavelength dispersion gaps between slots of the subframe. Thus, the multiwavelength selectivity approach will have an adverse effect on performance if the necessary wavelength dispersion gap is large compared to the slot size.

With multiwavelength selectivity, slot allocation becomes more complex. Beside allocating slots for transmission, additional slots must be blocked. For slot allocation we propose two schemes: without and with reuse of blocked slots. Both schemes are discussed below.

3.5.1 Reservation without reuse of blocked slots

In this scheme, the slots blocked by one destination are considered busy by all other destinations. This essentially means that if a destination is tuned to M wavelengths during a subframe, M slots (instead of one) will be used for each of its connections.

A slot i is available in a given subframe for a transmission to destination d on wavelength w if all the following conditions are met in the subframe:

1. slot i on w is idle and d is idle in slot i

2. every slot j in which d is receiving (on another wavelength) is idle on w

When slot i on wavelength w is allocated for a transmission, several slots need to be marked as "blocked" in the subframe under consideration. Specifically:

- mark slot i on all wavelengths selected in the station d filter, except for the slot which is now busy on wavelength w (vertical marking)

- if w is a newly selected wavelength, mark every slot j on w such that d is receiving in that slot (on another wavelength, say w') (horizontal marking)

When transmission in slot i on wavelength w is completed, slot i on all the wavelengths selected in the station d filter in the subframe is released. Furthermore, if slot i was the only slot used by d on wavelength w in the subframe, d does not need to be on w anymore and all slots on w blocked by d can be released.

Figure 3.13 illustrates slot allocation and marking in a subframe with four slots and four wavelengths.

Figure 3.13a shows a slot allocation for the case where destination 1 is receiving on W0 and W1, and destination 3 on W2. The slots used for transmissions

	0	1	2	3
W0	D1	B1		
W1	B1	D1		
W2		D3		
W3				

(a)   D1 selects W0,W1; D3 selects W2

	0	1	2	3
W0	D1	B1	D1	
W1	B1	D1	B1	
W2		D3		
W3				

(b)   D1 selects W0,W1; D3 selects W2

	0	1	2	3
W0	D1	B1	D1	B1
W1	B1	D1	B1	B1
W2		D3		
W3	B1	B1	B1	D1

(c)   D1 selects W0,W1W3; D3 selects W2

Figure 3.13: Slot allocation in a subframe when multiwavelength selectivity is used

63

to destination k are marked with Dk as before, and the slots blocked by destination k are marked with Bk. We see that for each connection to destination 1, an extra slot is blocked. Figure 3.13b shows the slot allocation after arrival of a new connection request to destination 1, in time slot 2 on wavelength 0. Since destination 1 is using also W1, a slot must be blocked on this wavelength in the same time slot (vertical marking). Figure 3.13c shows the slot allocation after arrival of another connection to 1 on W3. Since destination 1 is receiving on W0 and W1, in order to allocate the connection in this subframe, destination 1 has to add W3 to the set of its selected wavelengths. In order to be able to do this, in addition to allocating the connection to slot 3 on W3, slots 0, 1, and 2 on W3 must be blocked (horizontal marking). Also, slot 3 on W0 and W1 must be blocked (vertical marking). If any of these slots were busy or marked, the connection would be blocked in this subframe. We see that the allocation of this connection reduces the number of idle slots from 9 to 3.

For a connection request to destination d that requires wavelength w, the following slot assignment and marking algorithm is used:

Step 1: If d is active on w in one or more subframes, select from these subframes the one that has an available slot and has the maximum number of slots already assigned to d. If such a subframe is found, allocate randomly one of its available slots and mark the corresponding slots; else go to Step 2.

Step 2: From the set of subframes where d is receiving on less than the maximum number of wavelengths that can be selected by its filter (W_m), select randomly a subframe that has an available slot that can be allocated and where the number of wavelengths used by d is smallest. If such a subframe is found, add w to the set of wavelengths selected by destination d in this subframe, allocate randomly one of its available slots and mark the corresponding slots; else the connection is blocked.

Upon termination of a connection to destination d on wavelength w, the slot used by the connection and the corresponding blocked slots are released. If this was the only connection to d on w in the subframe, w is removed from the set of wavelengths selected by destination d.

This algorithm can be easily implemented using centralized approach. The distributed version of the algorithm would require much more complex connection establishment procedure than the one described in section 4.3.

Figure 3.14 shows the simulation results for the T/WDMA scheme that uses

multiwavelength selectivity. The four solid line curves show blocking probability
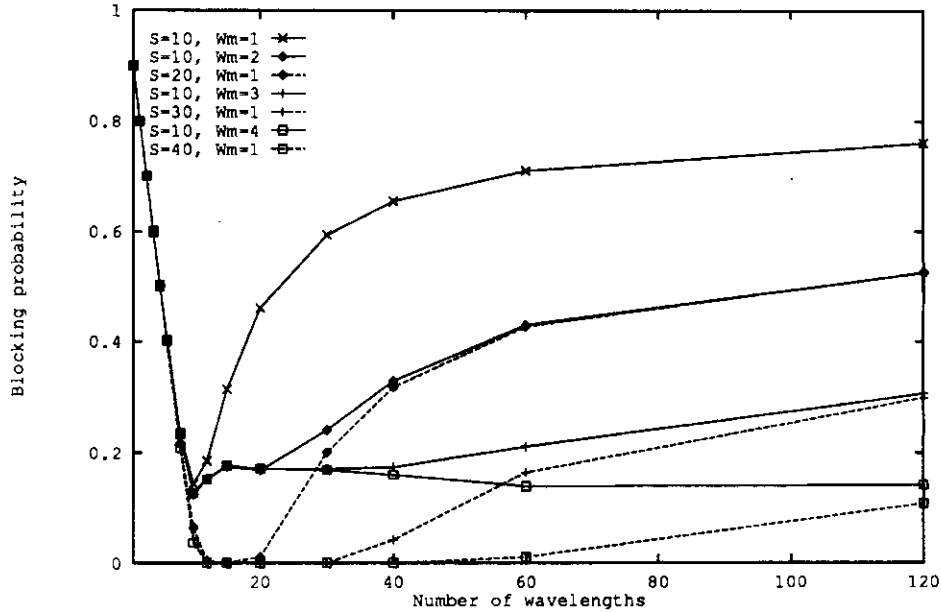
Figure 3.14: Blocking probability for $L = 6000$, $S \times K = 600$ and $N = 120$ when multiwavelength selectivity is used. The number of filter selectable wavelengths W_m ranges from 1 to 4.

versus number of wavelengths for W_m=1,2,3 and 4 where W_m is the degree of wavelength selectivity of the acoustooptic filter. We note a major improvement in blocking probability as W_m increases from 1 to 4, when the total number of wavelengths is larger than 10. The three dashed curves represent the blocking probability lower bounds for $W_m = 2$, 3 and 4, respectively, obtained using our analytical model in which S is replaced with SW_m and K with K/W_m. We see that wavelength selectivity has a similar effect on blocking probability as the increase in the number of subframes. For example, the curve ($S = 10$, $W_m = 2$) and ($S = 20$, $W_m = 1$) are almost the same for large W. The increase in wavelength selectivity, however, gives less improvement than the corresponding increase in number of subframes since some slots are wasted because of blocking. With an increase in the number of wavelengths, the degradation due to blocked slot waste diminishes; thus, the corresponding solid and dashed curves asymptotically converge.

65

3.5.2   Reservations with reuse of blocked slots

Since the purpose of blocking slots is to make sure that nobody transmits in them, we can allow more than one destination to block the same slots at the same time. By doing so, we reduce the total number of blocked slots and thus make more slots available for transmissions.

The previously stated slot availability conditions has to be slightly modified in the case of reusing blocked slots. In this case, a slot i is available in a given subframe for a transmission to destination d on wavelength w if all the following conditions are met in the subframe:

1. slot i on w is idle and d is idle in slot i

2. every slot j in which d is receiving (on another wavelength) is idle or blocked on w

Slot assignment and marking are carried out as before (both vertically and horizontally), except for the fact that now multiple blockings are allowed on the same slot and wavelength. This approach reduces the waste caused by slot blockings.

Figure 3.15 illustrates slot allocation and marking when the reuse of blocked slots is allowed. Figure 3.15a and 3.15b show the slot allocation before and after the arrival of a connection to destination 0 on wavelength 1. After the horizontal marking, we see that slots 0 and 2 on wavelength 1 are blocked both by destination 0 and destination 1. Note that we would not be able to allocate a slot for this connection if reuse were not allowed. A blocked slot can be reused by an arbitrary number of stations.

Figure 3.16 replicates the experiments reported in Figure 3.14, except that now reuse is allowed.

Comparing the curves of Figures 3.14 and 3.16, we see an overall improvement in performance if reuse is allowed. This result is expected since slot reuse reduces the overall number of blocked slots and thus makes more slots available for transmissions. The asymptotic convergence of the solid line curves to their corresponding dashed curves is now much faster.

3.6   Performance analysis for traffic with localities

We have studied so far the performance assuming that the traffic distribution is uniform. In reality, it is often the case that the stations form communities of interest where most communication takes place. Thus, the stations form a

Subframe

	0	1	2	3
W0	D1	B1	D1	
W1	B1	D1	B1	
W2	B0		D0	
W3	D0		B0	

(a)   D0 selects W2,W3; D1 selects W0,W1

	0	1	2	3
W0	D1	B1	D1	
W1	B0, B1	D1	B0, B1	D0
W2	B0		D0	B0
W3	D0		B0	B0

(b)   D0 selects W1,W2,W3; D1 selects W0,W1

Figure 3.15: Reuse of blocked slots

number of clusters and most traffic is local within those clusters. In this section we study the effect of traffic localities on the performance of the proposed T/WDMA scheme. The performance analysis is based only on simulation results. The analytical model for this case would be considerably more complex than the one developed for uniform traffic.

In order to model the effect of traffic localities on performance, we assume that stations can be grouped into a number of clusters. We introduce new performance parameters: the number of clusters C and the locality factor a. The locality factor represents a fraction of traffic that is "local", (i.e., remains within a cluster). The calls that are not local are called remote or intercluster calls. We also assume that the traffic distribution within each cluster and on the intercluster level is uniform. We study only a two level organization where we distinguish between local and remote traffic. Note that in reality the clusters can be hierarchically
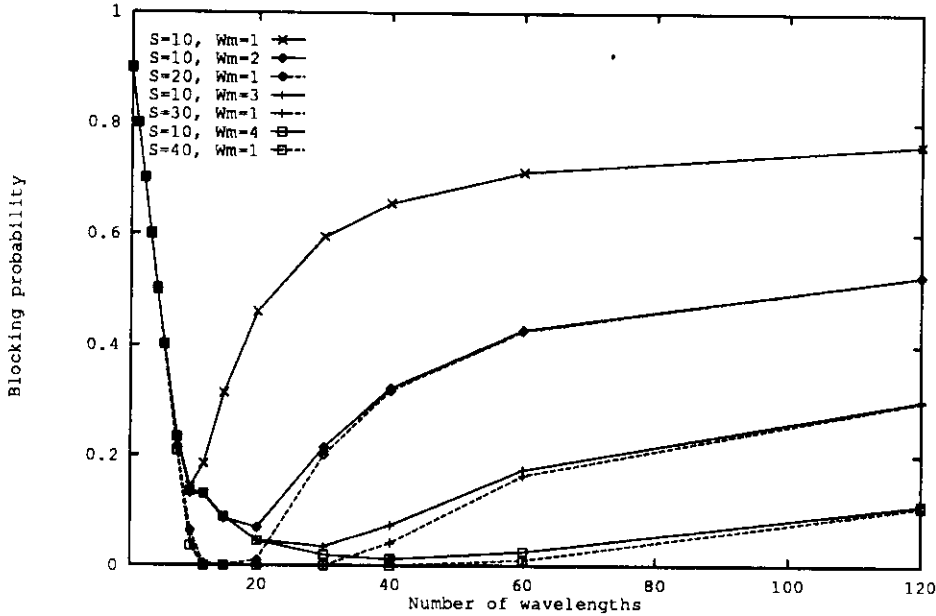
Figure 3.16: Blocking probability for $L = 6000$, $S \times K = 600$ and $N = 120$ with multiwavelength selectivity and reuse of reserved slots. The number of filter selectable wavelengths W_m ranges from 1 to 4.

organized into several levels where clusters can form a larger cluster on the higher hierarchical level.

Handling of traffic with localities by the proposed T/WDMA network opens an interesting problem: how to assign wavelengths to transmitting stations. We study two different wavelength assignment approaches.

The first approach called MINW minimizes the number of wavelengths assigned to each cluster. The total number of W wavelengths is partitioned into C sets of wavelengths in such a way that every wavelength belongs to at least one of those sets and the number of wavelengths used in each set is minimized. For example, if the total number of wavelengths is 20 and the number of clusters is 10, we would assign 2 wavelengths to each cluster. It seems reasonable to minimize the number of wavelengths used by stations belonging to the same cluster in order to minimize the number of subframes needed by local traffic which could lead to overall performance improvement.

The second approach called MAXW maximizes the number of wavelengths assigned to each cluster. In other words, we try to assign as many different wavelengths as possible to stations within a cluster. The maximum number of

wavelengths that can be assigned to each cluster is $\min(N_c, W)$ where N_c is the number of stations within the cluster.

Figure 3.17, 3.18 and 3.19 compare these two wavelength assignment approaches for a network of 500 stations that can be partitioned into 10 clusters. Figure 3.17 shows the blocking probability versus number of wavelengths for local (i.e., intracluster traffic). The locality factor a varies from 0.5 (i.e., 50% of traffic is intracluster) to 0.9. Figure 3.18 shows the same performance measures for re-

Figure 3.17: Blocking probability for intracluster traffic when $L = 5000$, $S = 10$, $K = 50$, $N = 500$ and the number of clusters is 10

mote (i.e., intercluster) traffic. Figure 3.19 shows the average blocking probability for both local and remote traffic.

Let us study first the approach that minimizes the number of wavelengths per cluster. We see that the total blocking probability and the blocking probability for local traffic always improve with the increase in traffic localities. This can be explained by observing that for local traffic a small number of wavelengths is used and thus the probability of wavelength conflicts is reduced. The blocking probability for remote traffic increases with the increase in traffic localities for small W. In case of small W, the load on each wavelength is high and a destination needs more than one subframe to assign to a particular wavelength. The number of subframes that a destination uses for a "local" wavelength (i.e. wavelength
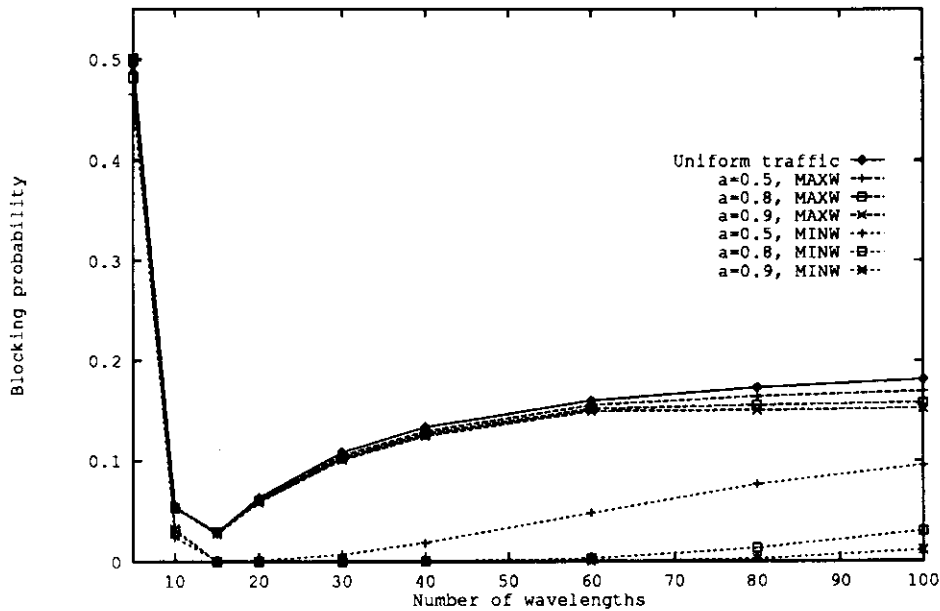
Figure 3.18: Blocking probability for intercluster traffic when $L = 5000$, $S = 10$, $K = 50$, $N = 500$ and the number of clusters is 10

used by stations within its cluster) increases with a which reduces the number of subframes that are available for a "remote" wavelength. With an increase in the number of wavelengths, each wavelength becomes less loaded and the number of subframes used by a "local" wavelength is reduced. The overall effect is that more subframes become available to "remote" wavelengths. Since the load offered to a destination is lower on "remote" wavelengths due to traffic locality, the blocking of remote calls caused by wavelength conflicts is reduced which results in overall reduction in blocking. In other words, using this wavelength assignment approach it is possible to pack more local calls into fewer number of subframes which allows remote calls to use more subframes. This, in fact, reduces wavelength conflicts for both local and remote calls.

When the number of wavelengths per cluster is maximized we can notice that the performance of the T/WDMA scheme for both local and remote traffic are almost the same as they are in the case when the traffic is uniformly distributed.

If we compare blocking probabilities for those two approaches, we can notice that the MAXW approach performs better when the number of wavelengths is small. For larger number of wavelengths, as well as at the optimal operating point the MINW approach is superior for both local and remote traffic and can
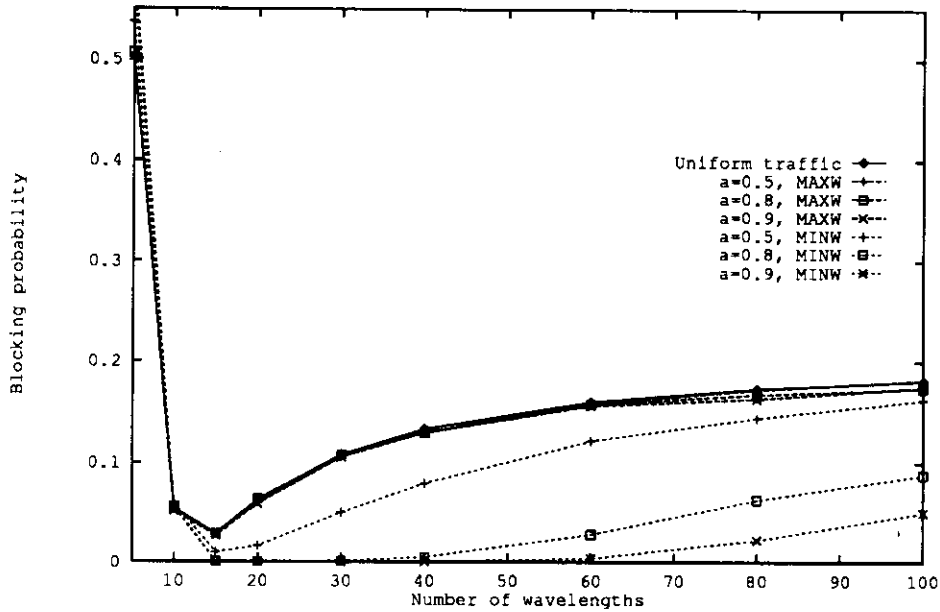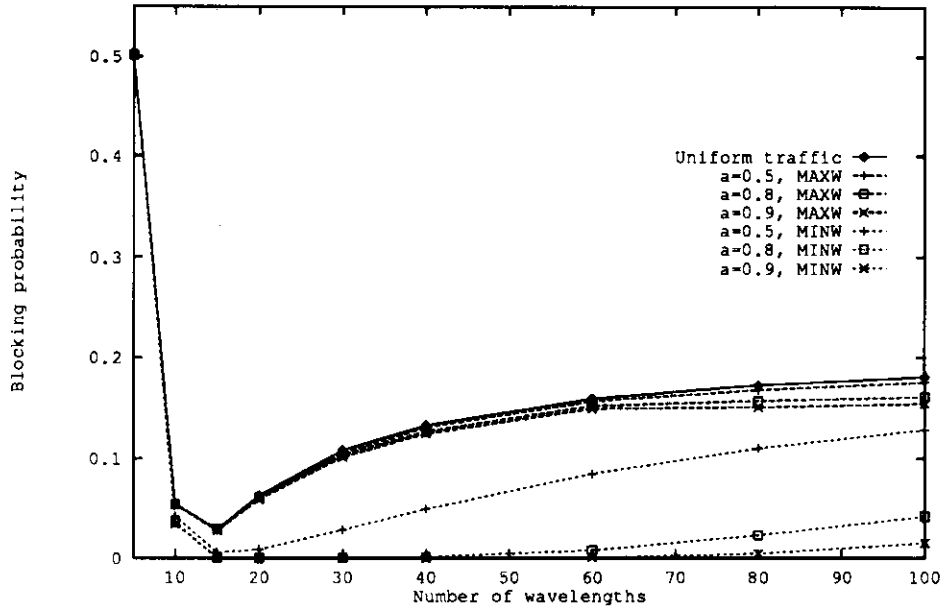
Figure 3.19: Blocking probability for both intra- and intercluster traffic when $L = 5000$, $S = 10$, $K = 50$, $N = 500$ and the number of clusters is 10

significantly improve performance if traffic distribution is characterized by high traffic localities.

3.7 Conclusions

We have proposed a novel T/WDMA scheme for circuit-switched traffic and large user population realizable with available optical technology. An approximate analytical model of this scheme for uniform point-to-point single-rate traffic was developed and performance analysis based on this model was presented and compared to simulation results. It was shown that analytical model matches very closely the results obtained by simulations. The results show that our scheme has the potential to outperform T/WDMA schemes which are based on very fast tunable devices (nanosecond tuning times), in spite of the fact that our scheme uses much slower acoustooptic devices (microsecond tuning time). Basically, low tuning speed is compensated by a much larger tuning range. It is also shown that the performance of this scheme is largely affected by the number of tuning intervals (subframes) used. The more subframes the better the performance. Finally, it is shown that the multiwavelength filtering capability of the acoustooptic filters

as well as traffic localities can be effectively exploited to improve performance.

In most of the analysis presented in this chapter we assumed that all stations use a pair of acoustooptic tunable receivers. This may be too costly for some stations. It is possible to implement this scheme even if stations have different capabilities. A station has to have at least one slowly tunable receiver. Note that such a station would be able to establish direct, single-hop connections only with those stations that transmit on the wavelength to which the destination is tuned to, since its receiver is not able to retune during the frame (the tuning time is larger than the frame size). Thus, receiving capability of such a station would be limited. This is equivalent to using one subframe per frame. Practically, each destination can view different partitioning of the frame into subframes. As we pointed out, the number of subframes S is crucial factor for the performance of this scheme. The more subframes, the better the performance. The gateway station would typically be required to have a much larger number of subframes than a small station. By increasing the amount of hardware (i.e., the number of receivers) we can effectively increase S and improve performance of a station. For example, using M acoustooptic filters would increase the number of subframes $M - 1$ times compared to the two receivers case. Alternatively, the multiwavelength filtering capability can be used to achieve a similar effect.

CHAPTER 4

HONET Implementation on a Linear Lightwave Network

The Linear Lightwave Network (LLN) is an attractive infrastructure for implementation of the HONET architecture. As we already mentioned in Chapter 1, one of the main advantages of LLN over other passive optical networks is its capability to change the network topology dynamically which makes possible to overcome link and node failures. In this chapter we address the synchronization problem of LLN. Lack of global synchronization in the original LLN complicates the implementation of time division multiaccess schemes such as those used in HONET. The LLN architecture is described and a new optical signal routing scheme is proposed. This scheme provides global synchronization and makes LLN suitable for implementing HONET. We then evaluate power budget for LLN that uses the new scheme.

4.1 Overview of LLN

The LLN [Ste90, Ste91, BSB91] is a recently proposed lightwave network architecture. As its name implies, LLN performs only linear operations on optical signals: power combining, splitting, and linear (nonregenerative) optical amplification. LLN consists of stations and nodes interconnected among themselves via bidirectional fiber optic links. The key component of LLN is a device called the Linear Divider-Combiner (LDC) installed at each network node. LDC is a multiport device that performs switching, splitting and multiplexing of optical signals. LDC is, in fact, an electrically controllable optical switch with additional features such as multicasting (an input signal can be distributed to several outputs) and multiplexing (several input signals can be combined on a single output link). Figure 4.1 shows a 4×4 LDC with four pairs of incoming and outgoing fiber links. Input lines represent incoming optical fibers and output lines outgoing fibers.

A Linear Divider-Combiner of arbitrary size can be built using 2×2 directional couplers with adjustable coupling ratio implemented in Ti:LiNbO$_3$ technology
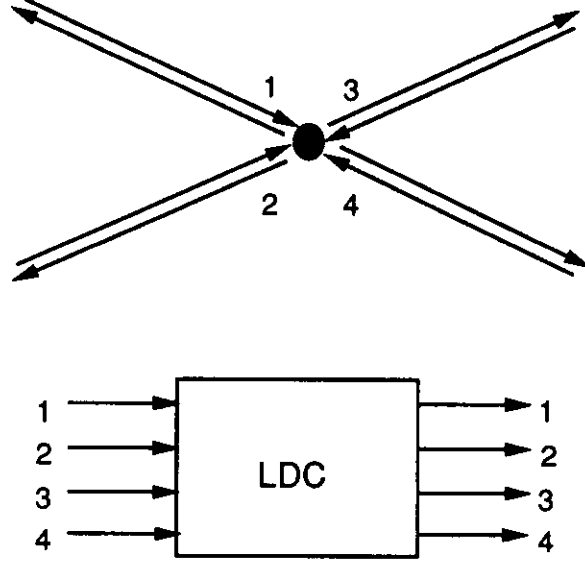
Figure 4.1: 4×4 Linear Divider-Combiner

[KA88]. Figure 4.2 illustrates implementation of a 4×4 LDC using these couplers. We note that the first two stages are "divider" stages, and the last two are "combiner" stages.

The LDC's switching function is defined by a **power transfer matrix**. A $K \times K$ LDC will have the following power transfer matrix:

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK} \end{bmatrix}$$

where matrix element a_{ij} represents the fraction of optical signal from input port i sent to output port j. The matrix elements can be expressed as

$$a_{ij} = \delta_{ij}\sigma_{ij}, \quad i,j \in \{1,2,\ldots,K\} \tag{4.1}$$

where δ_{ij} represents the portion of power from input port i split to output port j (through the divider stages), and σ_{ij} the portion of power from input port i combined into output port j (through the combiner stages). Note that δ_{ij} and σ_{ij} must satisfy the following sets of constraints:

$$\sum_{j=1}^{K} \delta_{ij} \le 1, \quad i \in \{1,2,\ldots,K\} \tag{4.2}$$

Figure 4.2: Implementation of a 4 × 4 LDC using 2 × 2 directional couplers

$$\sum_{i=1}^{K} \sigma_{ij} \le 1, \quad j \in \{1, 2, \ldots, K\} \tag{4.3}$$

Power splitting and combining in a LDC is illustrated in Figure 4.3.

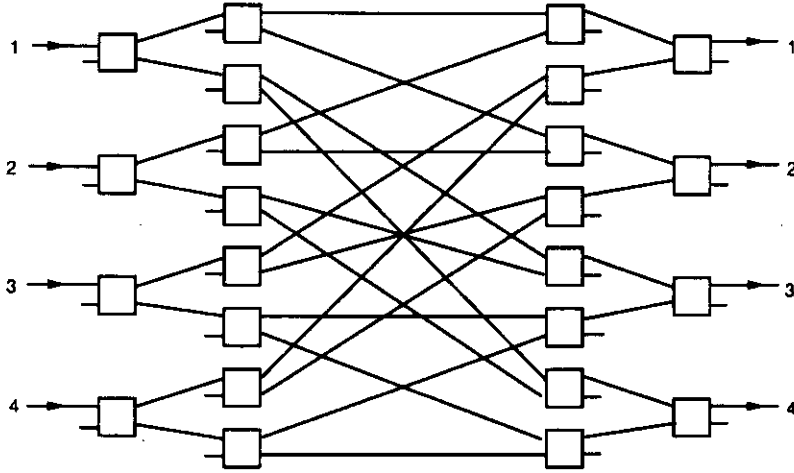LLN can have an arbitrary mesh physical topology. Figure 4.4 shows a LLN with six stations (A, B, C, D, E and F) and five nodes (g, h, i, j and k). Each edge in the graph represents a bidirectional link, i.e. it may be viewed as two unidirectional fiber links going in opposite directions. By adjusting the δ and σ parameters, arbitrary connectivity clusters can be created. Stations that communicate among themselves belong to the same cluster. Suppose that we have two clusters: cluster 1 with stations B, D, and F and cluster 2 with stations A, C and E. The stations that belong to the same cluster form a subnet that has a tree topology. The subnet topology must be a tree in order to avoid signal interference that would happen if the same signal reaches its destination traveling on multiple paths. Also, clusters are disjoint by definition. Therefore, different subnets cannot share links. In Figure 4.4 the dotted edges represent links that are used by subnet 1 and the bold edges links used by subnet 2. Suppose now that link gh failed. Subnet 1 can be reconfigured by changing the LDC power transfer matrices in order to add links hi and ij. Similarly, in the case of failure of link hk, subnet 2 can be reconfigured by adding link hi. This example shows the potential of LLNs in improving network fault tolerance.

In the original LLN proposal, the routing of optical signals in a subnet is performed using **shortest path routing**, as illustrated in Figure 4.5. If A transmits, the routing pattern in Figure 4.5.a is followed. If C transmits, the pattern shown in Figure 4.5.b is followed. In general, a separate broadcast tree
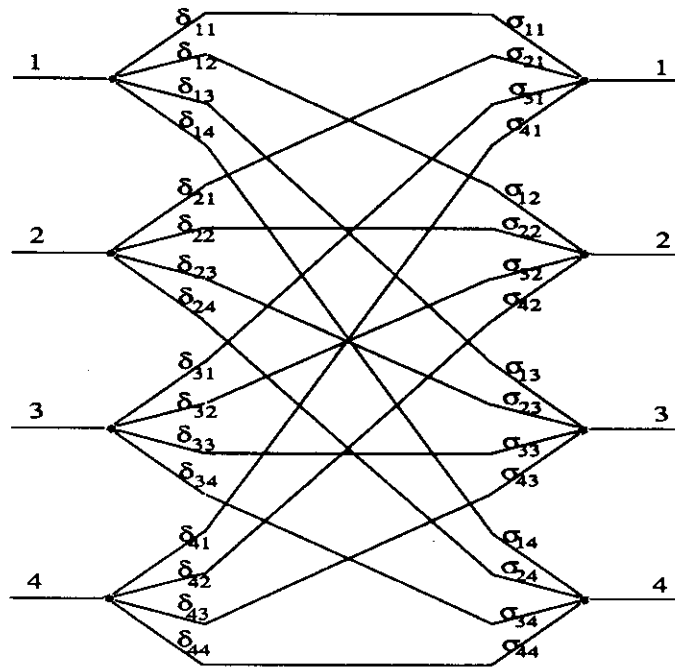
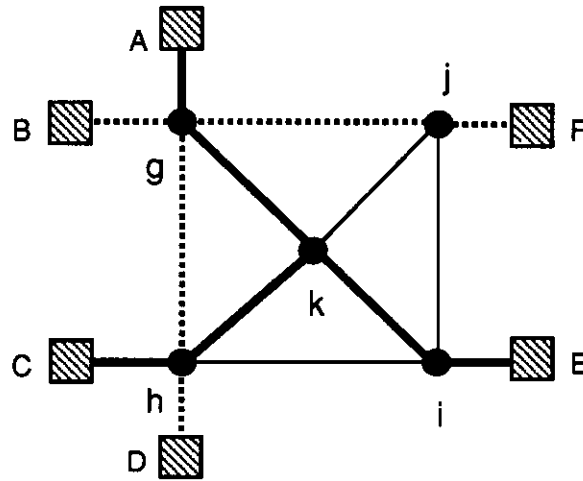Figure 4.3: Power splitting and combining in a LDC

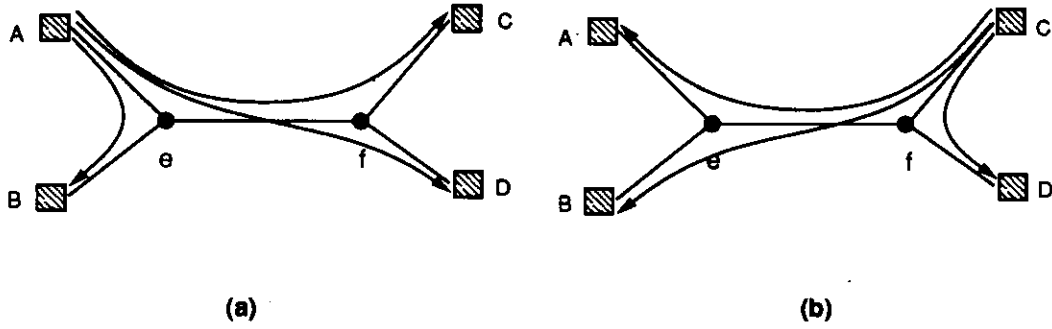
Figure 4.4: A LLN on a mesh topology

Figure 4.5: Shortest path routing

is associated with each source. LDCs are configured to support such distinct broadcast trees. For example, the power transfer matrices for the LDCs at nodes e and f in the network shown in Figure 4.5 have the following values if the power distribution at the LDCs is uniform (i.e. the power from each input port is split equally among all selected output ports, and each output port receives equal portion of the power from all selected input ports):

$$M_e = \begin{bmatrix} a_{AA} & a_{AB} & a_{Af} \\ a_{BA} & a_{BB} & a_{Bf} \\ a_{fA} & a_{fB} & a_{ff} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

$$M_f = \begin{bmatrix} a_{CC} & a_{CD} & a_{Ce} \\ a_{DC} & a_{DD} & a_{De} \\ a_{eC} & a_{eD} & a_{ee} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

4.2 TDMA synchronization problem

In star topology networks all transmissions are directed toward a central node, from which they are then broadcast to all nodes. Synchronization is required to coordinate transmissions so that no two transmissions overlap at the central node. This can be achieved by determining for each node the time instants when it can begin its transmission. By measuring the propagation delays from each node to the center it is possible to schedule transmissions so that they arrive at the central node at the desired time. Thus, it is possible to fully utilize each slot in the frame, achieving maximum channel capacity.

However, synchronization is much more difficult to achieve in a LLN network than in a star network. What makes the synchronization problem difficult is the fact that in LLN there is no single synchronization point, so that transmissions

must be synchronized with respect to not just a single (central) node but several nodes. Furthermore, these synchronization requirements may be conflicting, making it impossible to achieve the optimal utilization of a shared channel. Figure 4.6 illustrates the problem. The network on the Figure 4.6.a consists of the stations A, B, C and D interconnected among themselves via the network nodes e and f. These stations and nodes, along with the links incident on them, form a tree.

In this example we assign 1/2 of the total channel capacity to stations A and C to broadcast their traffic to all other stations. The requirement for broadcast connections is more restrictive than for point-to-point connections. It dictates that an optical signal must not collide with another signal anywhere in the network. We assume that slot size is one time unit, and that propagation delays are normalized to slot size. Figure 4.6.a shows the topology of a network where link labels represent normalized propagation delays. The TDMA frame carries two slots, one slot each for stations A and C. Let us assume that station A transmit at times $2k$ $(k = 0, 1, 2, \ldots)$. Station A's slots arrive at node e at times $2k + 1.3$ and at node f at times $2k + 1.6$. In order to have the transmission arrivals from A and C synchronized at e, station C should start its transmissions at times $(2k + 1) + 1.3 - 1.2 = 2k + 1.1$. On the other hand, in order to synchronize the transmission arrivals at node f, C should transmit at times $(2k + 1) + 1.6 - 0.9 = 2k + 1.7$. Obviously, we are not able to satisfy both requirements at the same time, and therefore cannot avoid transmission collisions. Figures 4.6.b and 4.6.c show the cases where the synchronization is performed with respect to nodes e and f, respectively. In both cases collisions occur due to misalignment of transmission arrivals. In the first case station D receives collided transmissions and in the second case station B. In order to avoid collisions it is necessary to introduce gaps between slots. In this example, we need to insert a gap of length 0.6, as shown in Figure 4.6.d which would reduce channel utilization to 77%.

The synchronization problem can be solved by reducing the frame to a size that is equal to a common divisor of propagation delays, as stated by Stern [Ste91]. However, the requirement for having the frame size dependent on the propagation delays is not practical. First of all, this requirement may make the frame size too small for any practical application. Besides, propagation delays may vary with temperature changes, thus requiring continuous adjustments in the frame size. Furthermore, the propagation delays will also change when topological reconfigurations occur.

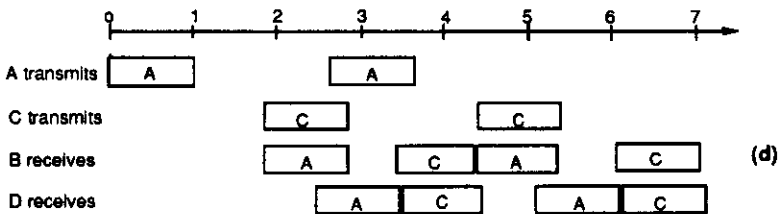There are other approaches that can be taken to solve the synchronization
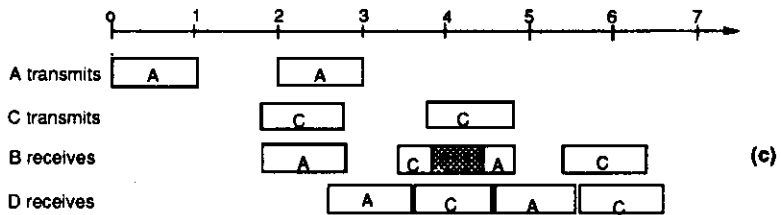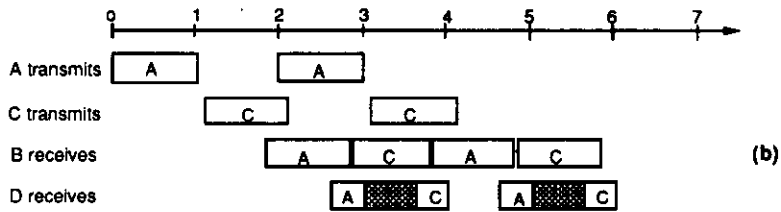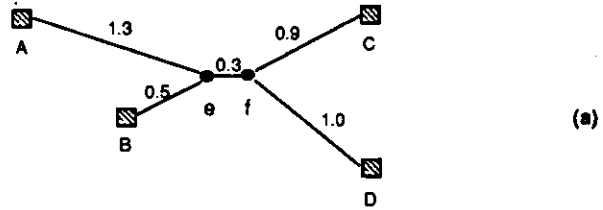
Figure 4.6: Synchronization problem in a tree topology network

problem. The slot size can be chosen to be equal to the sum of the transmission time and the maximum propagation delay between two nodes. This allows a transmission to drain out of the network before the next transmission starts. The disadvantage of this scheme is a low channel efficiency. The efficiency depends on the transmission time to maximum propagation delay ratio, and may become very poor in metropolitan area high-speed networks. Another approach that can be applied to networks that use both time and wavelength division multiplexing is to optimize network throughput by appropriate scheduling [Ste91]. This requires a complex scheduling algorithm which is expensive to implement. Stern [Ste91] also proposed a simple assignment procedure, called Pseudorandom scheduling. Each node is given the opportunity to transmit on a particular channel at times defined by a predefined pseudorandom sequence. The advantage of this approach is that the synchronization is simplified, requiring each source to be synchronized only with its destination. However, this approach has a relatively low efficiency, around 50%. Also, this method is not well suited to broadcast/multicast traffic.

In general, the previously mentioned approaches do not achieve optimal channel utilization. The more efficiency we want to obtain, the more complexity in synchronization and scheduling we must introduce. Yet, the result will not be optimal.

4.3   A solution for the TDMA synchronization problem

From the previous discussion on TDMA synchronization we observed that the difficulty with general LLN networks is that they require synchronizing the arrivals of transmissions at different nodes — a condition which cannot be easily satisfied. In the case of a star or a rooted tree (such as Tree-Net [GF88]), we only need to synchronize transmissions with respect to one node, i.e., the root. This is easy to achieve, and yields the maximum utilization of the shared channel. The same idea can be applied to LLN. A node is selected to be the root, and all optical signals, instead of traveling on the shortest paths, must go through that root node before reaching the destination(s). This routing scheme we call **rooted routing**. Figure 4.7 illustrates rooted routing. We may arbitrarily choose node f to be the root node. If A transmits, the routing pattern in Figure 4.7.a is followed. If C transmits, the pattern shown in Figure 4.7.b is followed. The signal is broadcast to all stations, including the station originating the signal. The stations can use this "echo" signal as an acknowledgment of a successful transmission (i.e., no collision occurred), or as a power reference to tune their transmit power level, or as a time reference to measure the round-trip propagation delay (for transmission

synchronization). Thus, rooted routing is a simple and straightforward solution

(a)   **(b)**
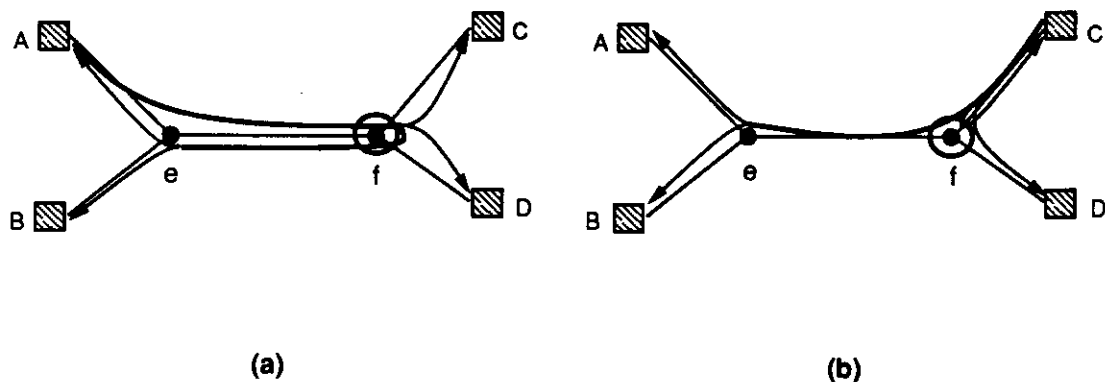
Figure 4.7: Rooted routing

of the TDMA synchronization problem. The stations need only to synchronize their transmissions to node f. By synchronizing their transmissions to f, stations A and B automatically achieve synchronization at node e, for example.

Figure 4.8 shows transmission scheduling for the example from Figure 4.6 when rooted routing is used with the root at node f. We see that all stations receive the transmissions without collisions and there is no need to introduce any gaps between slots. Thus, we can practically achieve 100% efficiency of the TDMA channel.

From this example we see that a transmission from A to B has to pass through links Ae, ef, fe and eB on its way to the destination. This is a longer path than in the case of shortest path routing where the signal goes only through node e before reaching station B. Therefore, the propagation delay increases. Also, the power loss may increase.

In order to implement this scheme, the Linear Dividers-Combiners must be properly configured. In the example, the power transfer matrices for nodes e and f, with the uniform power distribution at the LDCs, should have the following values:

$$M_e = \begin{bmatrix} a_{AA} & a_{AB} & a_{Af} \\ a_{BA} & a_{BB} & a_{Bf} \\ a_{fA} & a_{fB} & a_{ff} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

$$M_f = \begin{bmatrix} a_{CC} & a_{CD} & a_{Ce} \\ a_{DC} & a_{DD} & a_{De} \\ a_{eC} & a_{eD} & a_{ee} \end{bmatrix} = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix}$$

Note that all the elements of matrix M_f have value 1/9. If the LDC at node f were replaced by a 3×3 star coupler the matrix elements would have value 1/3.
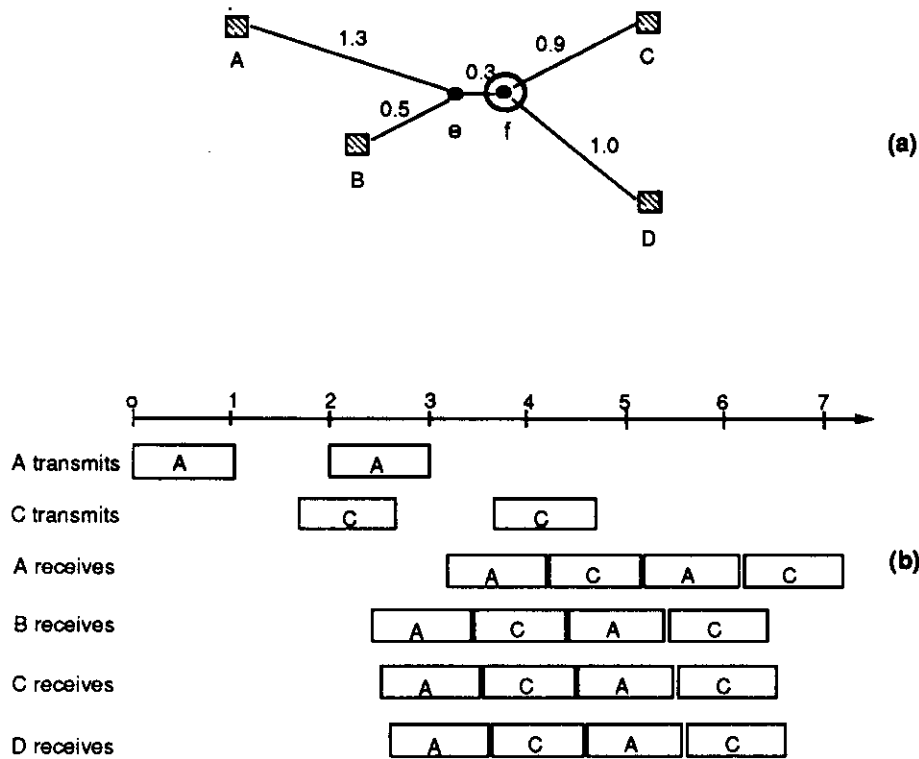
Figure 4.8: Synchronization with rooted routing

Thus, instead of attenuating each signal by a factor of 3 (due to power splitting), the LDC attenuates each signal 9 times. This is a consequence of the requirement that power distribution in LDCs has to satisfy constraints (4.1), (4.2) and (4.3). Basically, it is the penalty we pay in order to enjoy the reconfiguration flexibility of LLN. We further discuss this limitation of LDCs in the next section.

If all stations are at the same distance from the root, the maximum propagation delay in a LLN with rooted routing is the same as in a LLN with shortest path routing. If this is not the case, the delay increases when rooted routing is used. The actual increase depends on the network topology and the choice of the root node. It is shown in Appendix A that the maximum propagation delay of rooted routing is at most twice the maximum propagation delay of shortest path routing if the degree of the root is at least two (i.e., the root has at least two bidirectional links connected to it).

Any node in the network can be chosen as the central node. One of the criteria for choosing the center is to minimize the average or the maximum propagation delay. This can be done by choosing the node whose average or maximum distance from all stations is minimal. Another criterion for choosing the root node is to

minimize power losses. As it will be shown in the next section, the optimal power budget does not depend on the choice of the root node when link and excess losses are negligible. When link loss becomes the dominant factor (as is the case in large metropolitan networks), minimizing propagation delays will also minimize power losses, since both propagation delay and link loss are proportional to link lengths.

4.4   Power budget analysis

The main limitation in the number of stations that can be supported in an optical network is due to power losses. As we already mentioned in Chapter 1, the maximum ratio of the signal sent by a transmitter and the signal received by a receiver which ensures sufficient signal-to-noise ratio at the receiver for successful signal detection represents the *power margin*. We need to ensure that the maximum power loss between any two stations, defined as *power budget*, does not exceed the power margin. Therefore, the main concern in the power budget analysis is to evaluate the maximum power loss in a network, and to minimize it whenever possible. In addition to power budget, the dynamic range is also an important factor. Receiver dynamic range can be defined as the difference in power levels of the largest and the smallest signal that the receiver can handle. Thus, we must ensure that the difference in the maximum and the minimum power loss between transmitters and the receiver is smaller than the receiver's dynamic range. In rooted routing the dynamic range problem is easily resolved by requiring that each transmitter calibrate its power so that its own echo power is equal to the power received from a reference station. Since all the signals are routed through the same root, the calibration with the reference station ensures that every station receives the same power from all sources [GF88].

Next, we analyze the power budget of a LLN with rooted routing. There are four factors that contribute to the power losses in this network: (1) *splitting losses* due to the splitting of optical power at a node to several outgoing links, (2) *link losses* due to the attenuation of an optical signal on a fiber optic link, (3) *excess losses* that occur at nodes due to the imperfect coupling of incoming and outgoing links with couplers and (4) *combining losses*.

The combining losses are the result of the LDC's limitation in multiplexing optical signals. This limitation was formulated in Section 2 by the set of constraints (4.3) which state the fact that it is not possible to build a power coupler that combines two or more uncorrelated inputs and delivers them to its output losslessly [Hen89]. For example, if we have a $K \times K$ LDC and we want to combine the signals from all K incoming links just to a single outgoing link, say r, ideally

we would like to have:

$$a_{ir} = 1, \quad i \in \{1, \ldots, K\}$$
$$a_{ij} = 0, \quad i \in \{1, \ldots, K\}, \quad j \in \{1, \ldots, r-1, r+1, \ldots, K\},$$

that gives

$$\sum_{i=1}^{K} a_{ir} = \sum_{i=1}^{K} \sigma_{ir} = K$$

Since generally $K > 1$, this assignment of the matrix coefficients does not give a feasible solution. In order to have a feasible solution, the power combining from an input port to output port r must be reduced K times. Therefore $\frac{K-1}{K}$ of total incoming power will be lost in this LDC due to power combining.

In this power budget analysis we are mainly interested in the effect of splitting and combining losses on total power loss. When the number of stations is large, these losses are dominant. While the excess and link losses can be further reduced with technology advancements, the splitting and combining losses cannot be eliminated. Thus, we begin our analysis studying the ideal case where excess and link losses do not exist.

Let us consider the power budget in a LLN with rooted routing implemented on the topology shown in Figure 4.9. We choose node h to be the root node. We
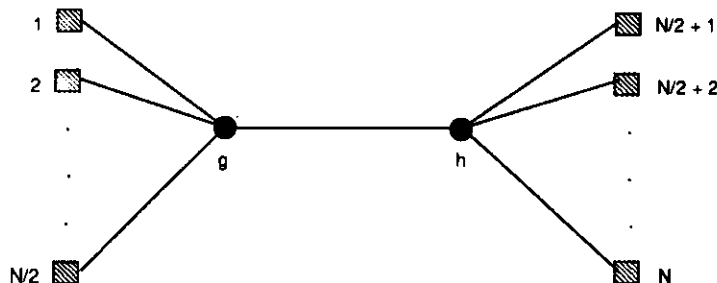
Figure 4.9: A tree topology with N stations and two nodes

consider first the ideal case when power distribution and multiplexing at LDCs is uniform. This means that that at node g an equal portion of power is combined from each link ig to link gh and an equal portion of power is distributed from link hg to each link gi where $i \in 1, \ldots, N/2$. Likewise, at node h an equal portion of power from each input port is distributed to each output port.

The maximum power loss in this network occurs between the stations i and j where $i, j \in \{1, \ldots, \frac{N}{2}\}$. A signal from station i after passing node g is attenuated

$\frac{N}{2}$ times due to the combining losses. After passing node h, the signal on link hg is attenuated $(\frac{N}{2} + 1)^2$ times due to the power splitting and combining at node h. Finally, the signal on link gj is again attenuated $\frac{N}{2}$ times due to power splitting at node g. The total power loss of the signal sent from station i to station j is

$$P[\text{dB}] = 10 \log_{10}((\frac{N}{2} + 1)^2(\frac{N}{2})^2) \approx 10 \log_{10}((\frac{N}{2})^4) \tag{4.4}$$

If shortest path routing is used, the maximum power loss for a LLN with uniform power distribution at LDCs occurs between the stations i and j, where $i \in \{1, \ldots, \frac{N}{2}\}$ and $j \in \{\frac{N}{2} + 1, \ldots, N\}$. The signal is attenuated at each LDC $(\frac{N}{2})^2$ times due to splitting and combining losses. Thus, the maximum power loss is

$$P[\text{dB}] = 10 \log_{10}((\frac{N}{2})^2(\frac{N}{2})^2) = 10 \log_{10}((\frac{N}{2})^4) \tag{4.5}$$

Power budget can be reduced in LLN with both types of routing if power distribution and multiplexing at each LDC is optimized rather than being chosen as uniform. We need to solve the following optimization problem: Determine the power transfer matrices for all nodes in the network, such that the maximum power loss between any pair of nodes is minimal. Let us first introduce the following notation:

N Total number of stations in the network.

R Root node.

$K^{(l)}$ Number of links incident to node l.

$r^{(l)}$ Link incident to non-root node l and on a path from that node to the root.

$m_i^{(l)}$ Node or station to which node l is connected via link i.

$n_i^{(l)}$ Number of stations that can be reached from node l via link i.

$S_i^{(l)}$ Set of stations that can be reached from node l via link i.

$Lopt_i^{in(l)}$ Optimal ratio between power transmitted by station $s \in S_i^{(l)}$ and power at input port i of node l $(i \neq r^{(l)})$.

$Lopt_i^{out(l)}$ Optimal ratio between power at output port i of node l $(i \neq r^{(l)})$ and power received by station $s \in S_i^{(l)}$.

$E^{(l)}$ Excess loss at node l expressed as ratio between total power entering the node and total power leaving the node.

D_{kl} Power loss on link (k, l).

$a_{ij}^{(l)}$ Power transfer matrix element for node l.

For example, the network shown in Figure 4.10 has the following parameters:
$N = 5$, $R = d$, $K^{(c)} = 3$, $K^{(d)} = 3$, $K^{(e)} = 3$, $r^{(c)} = 3$, $r^{(e)} = 1$, $m_1^{(c)} = A$,
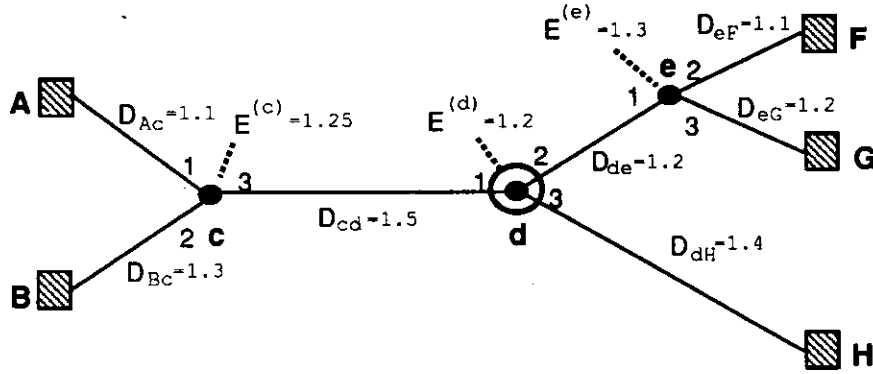
Figure 4.10: A network with five stations and three nodes

$m_2^{(c)} = B$, $m_3^{(c)} = d$, $m_1^{(d)} = c$, $m_2^{(d)} = e$, $m_3^{(d)} = H$, $m_1^{(e)} = d$, $m_2^{(e)} = F$,
$m_3^{(e)} = G$, $n_1^{(c)} = 1$, $n_2^{(c)} = 1$, $n_3^{(c)} = 3$, $n_1^{(d)} = 2$, $n_2^{(d)} = 2$, $n_3^{(d)} = 1$, $n_1^{(e)} = 3$,
$n_2^{(e)} = 1$, $n_3^{(e)} = 1$, $S_1^{(c)} = \{A\}$, $S_2^{(c)} = \{B\}$, $S_3^{(c)} = \{F, G, H\}$, $S_1^{(d)} = \{A, B\}$,
$S_2^{(d)} = \{F, G\}$, $S_3^{(d)} = \{H\}$, $S_1^{(e)} = \{A, B, H\}$, $S_2^{(e)} = \{F\}$, $S_3^{(e)} = \{G\}$.

The solution for the power budget optimization problem in a LLN with rooted routing which takes into account all types of losses is presented in Appendix B. If the links have the same losses in both directions (e.g., $D_{ij} = D_{ji}$), it is shown in the appendix that $Lopt_i^{in(l)} = Lopt_i^{out(l)} = Lopt_i^{(l)}$. In such a case we have the following result for the optimal power budget:

$$P[\text{dB}] = 10 \log_{10}(E^{(R)}(\sum_{k=1}^{K^{(R)}} Lopt_k^{(R)})^2) \tag{4.6}$$

where

$$p = m_i^{(l)}$$

and

$$Lopt_i^{(l)} = \begin{cases} D_{pl} & \text{if p is a station} \\ (\sum_{k=1, k \neq r^{(p)}}^{K^{(p)}} Lopt_k^{(p)}) E^{(p)} D_{pl} & \text{if p is a node} \end{cases}$$

It is shown in Appendix B that when power distribution is optimal, power loss between any pair of stations is the same. This essentially means that a receiver

will receive signals at the same power level from any transmitting station (assuming that all stations transmit the same power). Conversely, all stations receive the same power from the same transmitter. Thus, the dynamic range required by the receiver in a LLN with rooted routing and optimal power distribution is practically zero.

As an example, let us calculate the optimal power budget for the network shown in Figure 4.10. There we have: $Lopt_1^{(c)} = D_{Ac} = 1.1$, $Lopt_2^{(c)} = D_{Bc} = 1.3$, $Lopt_2^{(e)} = D_{eF} = 1.1$, $Lopt_3^{(e)} = D_{eG} = 1.2$, $Lopt_1^{(d)} = (Lopt_1^{(c)} + Lopt_2^{(c)})E^{(c)}D_{cd} = 4.5$, $Lopt_2^{(d)} = (Lopt_2^{(e)} + Lopt_3^{(e)})E^{(e)}D_{de} = 3.56$, $Lopt_3^{(d)} = D_{dH} = 1.4$, $P[dB] = 10\log_{10}(E^{(d)}(Lopt_1^{(d)} + Lopt_2^{(d)} + Lopt_3^{(d)}))^2 = 20.30$.

From this general solution we can get a solution for the ideal case. There we have

$$E^{(l)} = 1$$
$$D_{kl} = 1$$
$$Lopt_i^{(l)} = n_i^{(l)}$$
$$\sum_{k=1}^{K^{(R)}} Lopt_k^{(R)} = \sum_{k=1}^{K^{(R)}} n_k^{(R)} = N$$

and the optimal power budget is

$$P[dB] = 10\log_{10} N^2$$
$$= 10\log_{10} N + 10\log_{10} N \tag{4.7}$$

This is an interesting result that shows that in the ideal case the optimal power budget does not depend on the network topology nor on the choice of the root node. The optimal power budget is determined only by the number of stations in the network. Also, we can see that the combining losses contribute to the total power loss as much as the splitting losses do.

It is shown in Appendix C that the optimal power budget in the ideal case when shortest path routing is used is

$$P[dB] = 10\log_{10}(N-1)^2$$
$$= 10\log_{10}(N-1) + 10\log_{10}(N-1) \tag{4.8}$$

The results for the optimal power budget for both routing schemes are very similar. The difference is that when rooted routing is used, a signal is split among N stations, and when shortest path routing is used, it is split among $N-1$ stations.

87

Let us consider again the network shown in Figure 4.9. The optimal power budget for this LLN when rooted routing and shortest path routing are used is given in expressions (4.7) and (4.8), respectively. If we compare these results with the power budget when power distribution at LDCs is uniform, given in expressions (4.4) and (4.5), we can see that the power budget is reduced significantly by optimizing the power distribution.

If we take into account the excess and link losses, and assume that all nodes have the same loss E and all links have the same loss D, we can show from expression (4.6) that the maximum power loss in the optimized LLN with rooted routing is

$$\begin{aligned} P[\text{dB}] &= 10\log_{10}(E(\frac{N}{2}D^2E + \frac{N}{2}D)^2) \\ &\approx 10\log_{10}((\frac{N}{2})^2D^4E^3) \quad \text{if } DE \gg 1 \end{aligned}$$

Let us estimate now the maximum power loss in the LLN with shortest path routing and optimal power distribution. We can determine a simple upper bound for this loss if we use the same power distribution which was optimal for the ideal case (i.e., suboptimal power distribution). In such a case the maximum power loss is equal to the splitting and combining losses (which are the same as in the ideal case) increased by link and excess losses encountered on the path between stations i and j where $i \in \{1, \ldots, \frac{N}{2}\}$, $j \in \{\frac{N}{2} + 1, \ldots, N\}$. Thus, we have that the maximum power loss in LLN with shortest path routing and optimal power distribution is

$$P[\text{dB}] \leq 10\log_{10}((N - 1)^2D^3E^2)$$

From this example we can make the following observations. Shortest path routing is generally better than rooted routing when the link and/or excess losses are dominant. When the link and excess losses are small compared to the splitting and combining losses, the performances of both routing schemes are almost the same.

With a power margin of 40 dB we find from expression (4.7) that it is possible to support up to 100 stations in a LLN subnet with rooted routing in the ideal case. If it is necessary to support more stations or if the excess and/or link losses are significant, optical amplification should be used.

One possibility is to install an optical amplifier at each output port of each node. In order to compensate for excess and link losses, the amplifiers with a constant amplification gain can be used. Each amplifier should have an amplification gain equal to the excess losses within its node and to the losses on its outgoing link. These losses are constant because they depend only on the size of a LDC

(i.e. on the number of directional couplers within the LDC a signal has to pass going from an input to an output port) and the link length. If such amplifiers are used the power loss analysis for LLN with rooted routing in the ideal case gives the exact solutions. Due to the simplicity of these solutions, the complexity of the network optimization and management can be greatly reduced. If we want to use the amplification to overcome the splitting and combining losses, the amplifiers whose gain can be dynamically controlled would be preferred because these losses depend on the topology of a LLN subnet that can be changed dynamically.

However, the installation of an optical amplifier per each link would drastically increase the network cost. Another, more economical solution that can be applied to LLNs with rooted routing requires adding a single optical amplifier. In this scheme a single unidirectional fiber link is added to the root node making a self-loop. Thus, this scheme requires an additional input and output port at the root node. The optical amplification is performed only at that link using the Erbium-doped fiber amplifier [NNA91]. Figure 4.11 illustrates the scheme. Optical signals coming to the root node from all its input links (except from
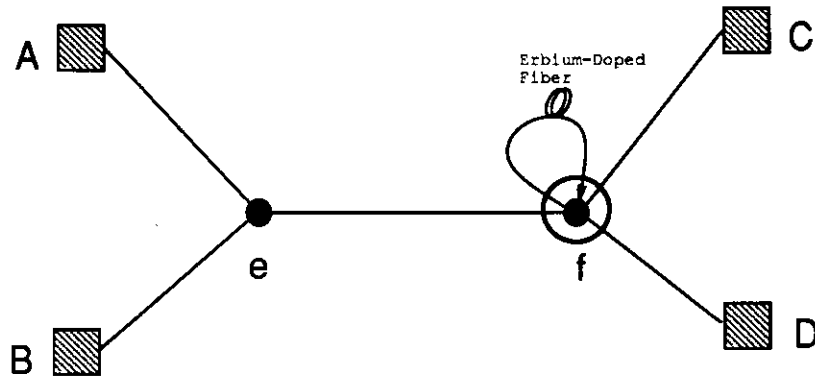
Figure 4.11: Loop link at the root node

the loop link) are multiplexed to the loop link, where optical amplification is performed. The amplified signal is then sent back to the root node where it is distributed to all its output links (except to the loop link). This routing scheme is defined by the following power transfer matrix for the node f when power distribution and multiplexing is uniform:

$$
M_f = \begin{bmatrix} a_{CC} & a_{CD} & a_{Ce} & a_{Cf} \\ a_{DC} & a_{DD} & a_{De} & a_{Df} \\ a_{eC} & a_{eD} & a_{ee} & a_{ef} \\ a_{fC} & a_{fD} & a_{fe} & a_{ff} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix}
$$

The power budget optimization for this scheme is considered in Appendix D. If the amplification gain at the loop link is $G[\text{dB}]$, the optimal power budget for such a network is

$$P[\text{dB}] = 10 \log_{10}(E^{(R)}(\sum_{k=1, k \neq r^{(R)}}^{K^{(R)}} Lopt_k^{(R)}))^2 - G[\text{dB}] \qquad (4.9)$$

Provided that G is high enough to compensate the power losses from stations to the loop link, we have the following expression for the optimal power budget

$$P[\text{dB}] = 10 \log_{10}(E^{(R)}(\sum_{k=1, k \neq r^{(R)}}^{K^{(R)}} Lopt_k^{(R)})) \qquad (4.10)$$

and for the ideal case

$$P[\text{dB}] = 10 \log_{10} N \qquad (4.11)$$

Thus, only splitting losses are incurred and power budget becomes equal to the power budget of the star network with N stations. The number of stations can be increased up to 10000 which should be sufficient in most cases.

4.5   Conclusion

In this chapter we proposed a new scheme for optical signal routing in Linear Lightwave Networks, called rooted routing. This new scheme overcomes the TDMA synchronization problem that exists in LLNs when shortest path routing is used. The rooted routing scheme, however, increases propagation delays and power losses. The power budget optimization problem for LLN with rooted routing is presented and solved, and it is shown that, when the excess and link losses are small compared to the splitting losses, and when power distribution is optimized, both routing schemes perform almost the same. Possibilities of using optical amplification in LLNs were also discussed and it was shown that the power budget of a LLN subnet that uses rooted routing can be significantly improved using a single optical amplifier.

CHAPTER 5

Multilevel and Multifiber HONET

One of problems with the HONET architecture discussed so far is scalability. The capacity of the single-hop network implemented using T/WDMA is limited due to the limited tuning range of optical transmitters and receivers, and it cannot be easily extended without major additions in station's complexity (an additional transmitter and receiver at each station are needed). Also, the number of wavelengths available for implementing the multihop network may not be sufficient if the number of stations is very large. This would require that more stations share the same wavelength which would reduce capacity of virtual links and consequently the overall network throughput. We showed in the previous section that the throughput decreases with the increase in channel sharing. Another major limitation is power budget. While power budget limitation of the multihop network can be overcome by eliminating the restriction that all stations have to be connected to the same passive optical medium (which would also allow wavelength reuse within separate PONs) this cannot be done for a broadcast-and-select single-hop network. Optical amplification can to some extent alleviate the power budget problem. However, the optical amplifiers existing today have a limited bandwidth of about 30 nm. With 1 nm spacing between channels, this gives us only 30 wavelengths [Min91]. Thus, we cannot expect to support more than on the order of, say, several hundreds of stations on the same broadcast medium.

In order to overcome these limitations we introduce multilevel and multifiber HONET, which we describe next.

5.1 Multilevel HONET

Multilevel HONET is an extension of the basic HONET architecture. In the multilevel HONET stations are grouped into communities of interest (i.e. clusters or subnets where most of communication takes place) which are then connected to the same PON. Clusters are then interconnected by a multilevel hierarchical structure. If most traffic is localized within clusters, this approach can significantly increase network capacity since wavelengths can be reused in separate

PONs.

The idea of hierarchical multilevel networks is not new. Stern [Ste90] proposed hierarchical organization for Linear Lightwave Network (LLN) with wavelength selective LDCs using single-hop overlaid networks, as illustrated in Figure 5.1. The single-hop networks use a set of contiguous wavelengths, called waveband. A unique waveband is assigned to each hierarchical level. At each hierarchical level
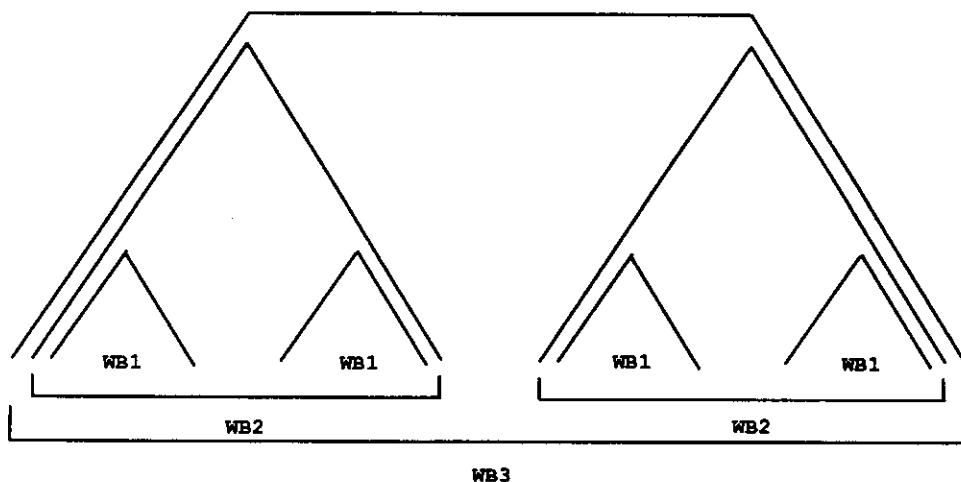

Figure 5.1: Single-hop overlaid networks

the stations communicate via a single-hop network. The problem with this organization is that it cannot overcome the power budget limitation since power is split among all the stations in the network at the highest hierarchical level. In addition to this, T/WDMA implemented over this network would require that each station have either a fast tunable transmitter-receiver pair with a tuning range that can cover all wavebands or a pair of devices for each waveband. MONET [GKB92] also has a multilevel network structure. The unique feature of the multilevel HONET structure here proposed is the separate handling of datagram and real-time traffic.

Multilevel HONET uses physically separate single level HONETs at each hierarchical level which are interconnected via gateways. Figure 5.2 shows a 2-level HONET implemented with the star PONs. The same concept can be extended to build n-level HONETs. Each cluster in the figure is connected to the intercluster PON with one or more gateways. The gateway performs time-slot switching for real-time traffic, which requires electrooptical conversions and buffering. Thus, real-time intercluster traffic is not single-hop anymore. It requires now three hops (one hop per each level it passes through). Clearly, the throughput of real-time intercluster traffic is limited by the throughput of the gateways which cannot
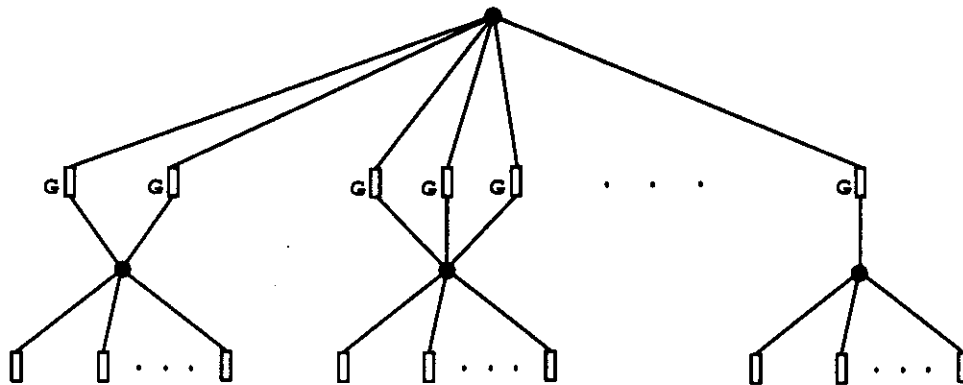
Figure 5.2: A 2-level HONET

exceed the maximum transmission speed that can be supported by electronics. If this throughput is not sufficient, additional gateways must be connected to the clusters.

If there are multiple gateways per cluster, the intercluster network can be implemented using several PONs as illustrated in Figure 5.3. Partitioning of
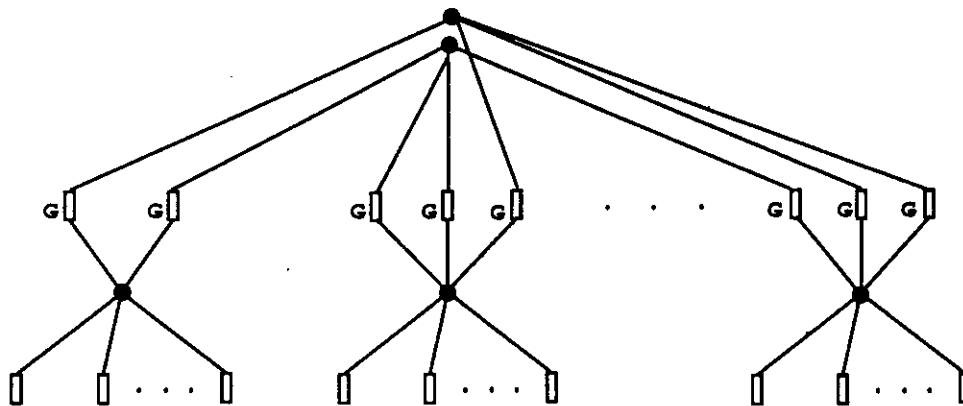
Figure 5.3: Multiple intercluster PONs in 2-level HONET

gateways to intercluster PONs can be arbitrary with the restriction that at least one intercluster PON has at least one gateway from each cluster connected to it (in order to ensure broadcast capability). By doing partitioning, power budget is improved, since smaller number of stations is connected to each PON. The capacity of the intercluster T/WDMA is also improved since each intercluster PON can reuse the same wavelength. Also, the reliability is improved. In case of failure of a gateway or an intercluster PON, an alternative gateway and PON can be used.

The previously described multilevel configuration cannot support single-hop,

"wavelength-on-demand" intercluster connections. It may be possible to achieve the single-hop intercluster communication using the single-hop overlaid networks approach shown in Figure 5.1. For such communication a station needs a slowly tunable transmitter-receiver pair that can tune over all wavebands. Alternatively, for each waveband a separate transmitter-receiver pair can be used. This solution, however, cannot overcome power budget limitations as we already pointed out. Another, more promising approach to provide single-hop intercluster connections is to use a multifiber architecture as we will show soon.

5.2   Multifiber HONET

Another approach to enhance the scalability and increase the capacity of HONET is to use multifiber cables, i.e. cables which contain up to hundreds of fibers. This solution is attractive because, although the multifiber cable is more expensive than single fiber cable, the marginal cost of an additional fiber is low, compared with installation and packaging costs. The payoff of the multifiber cable plant is very high, in that the number of available wavelengths is now amplified by a factor of the hundred. The first step in exploiting multiple fibers in HONET is to implement its circuit-switched and packet-switched network over different fibers. We may further distribute the wavelengths of the circuit-switched network over separate fibers. For example, the wavelengths of T/WDMA and the on-demand wavelengths can be placed on different fibers. Furthermore, we may use a multifiber architecture to implement each of HONET's components: the multihop network, the T/WDMA network and the "wavelength-on-demand" single-hop network, We discuss next two possible multifiber architectures, namely the multifiber tree and the multifiber LLN.

5.2.1   Multifiber tree architecture

In the multifiber tree network [BGK93] the station population is subdivided into clusters. Each cluster is allocated a certain number of "fiber plants", where a fiber plant is a single fiber subtree (i.e., a PON), embedded in the cable plant as illustrated in Figure 5.4. A station connects to one of the allocated fiber plants via a switch. A mechanical [BGK93] as well as a lithium-niobate optical switches can be used to this purpose. We refer a station assigned to a fiber plant as "active" at the fiber plant if it is actually connected to this fiber plant.

The cable plant is the tree formed by the multifiber cables. A two-level cable plant is illustrated in Figure 5.5. We assume that the cable plant is equipped at
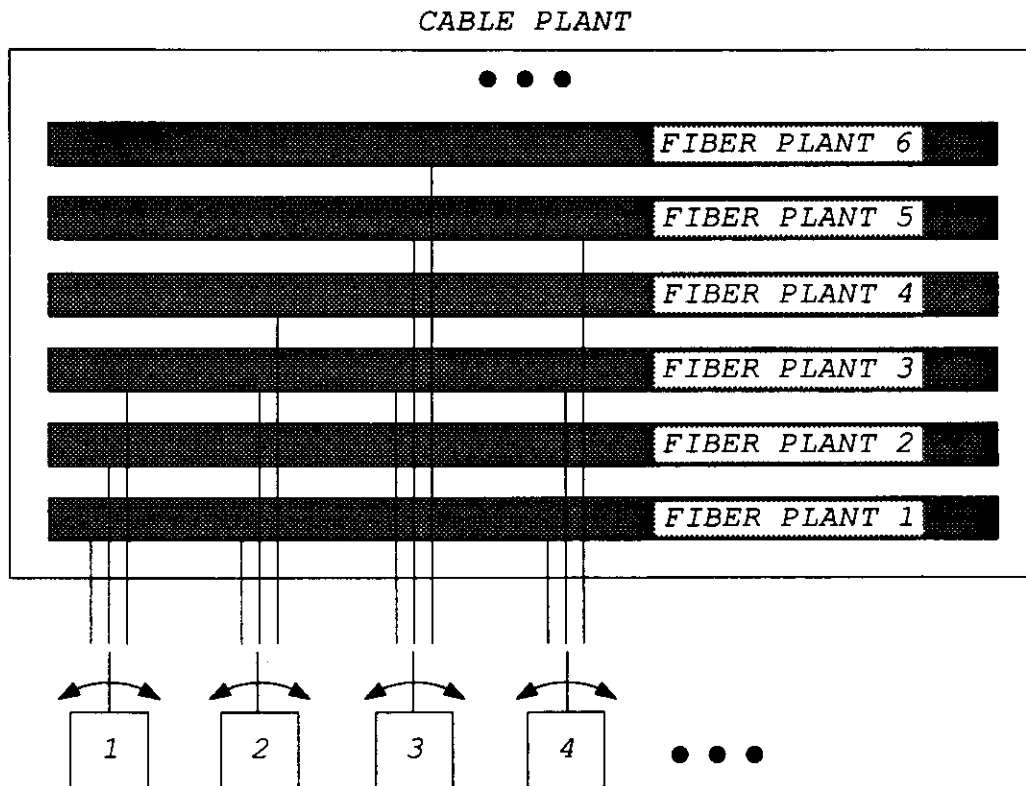
Figure 5.4: Several fiber plants embedded within a cable plant

its internal nodes (i.e., closets) with binary couplers, reflective star couplers and splices. By properly manipulating these components, any arbitrary fiber plant can be constructed within the cable plant. More specifically, we can establish, within the cable plant, a sufficient number of fiber plants so that for any pair of clusters there is at least one fiber plant which interconnects them. One such covering is proposed in [BGK93] which we describe next.

Figure 5.6 shows an example with four clusters and six fiber plants that interconnect them. Each fiber plant covers 2 clusters. If the number of clusters is K, the number of fiber plants L required to interconnect all clusters is

$$L = K(K-1)/2 \qquad (5.1)$$

Let F_{max} be the maximum number of optical fibers per cable in any segment of cable between two closets (i.e., nodes). The cable plant is feasible (i.e., accommodates fiber plants embedding) as long as the fiber count F per cable exceeds F_{max}, i.e. $F \geq F_{max}$. We observe that the maximum number of fibers is required on links connected to the root of the intercluster tree. These links contain $(K/2)^2$
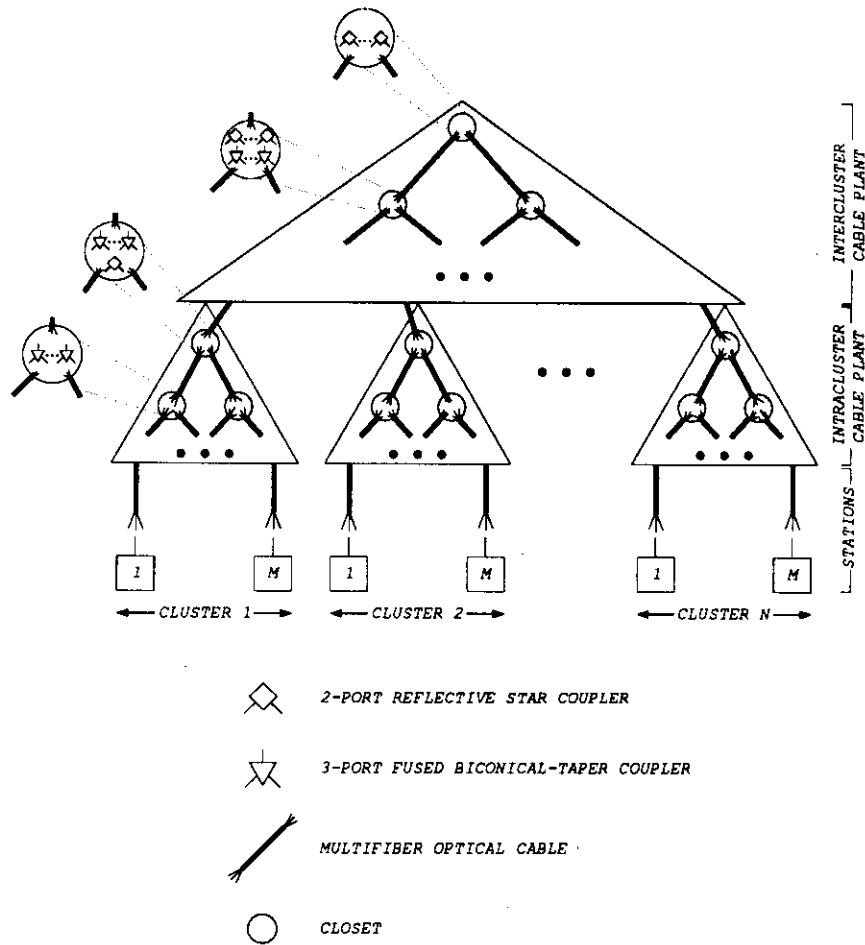
Figure 5.5: Two-level cable plant of the multifiber optical network

fibers that connect $K/2$ clusters from the left subtree with $K/2$ clusters from the right subtree. Thus, we have

$$F_{\mathrm{max}} = (\frac{K}{2})^2 \tag{5.2}$$

We see from expressions (5.1) and (5.2) that with the increase in the number of clusters, and reduction in the number of stations per cluster (as well as per fiber plant) the number of fiber plants and fibers per cable increases. Thus, it would be more cost effective to have a small number of large fiber plants. The rationale for keeping the fiber plants small in size has to do with optical power budget. In fact, an important goal of the proposed solution is to avoid the need for optical amplification of the cable plant), by placing a constraint on the number of stations connected to the fiber plants.

Let us determine the minimum number of fibers per cable necessary to imple-
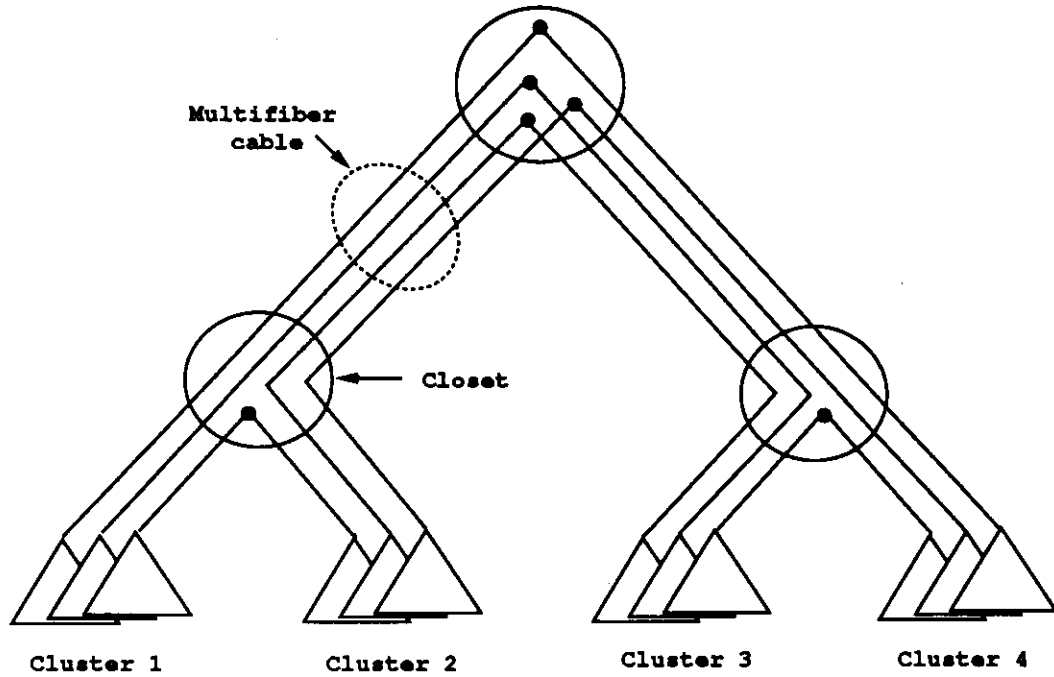
Figure 5.6: Multifiber tree network with four clusters

ment a multifiber tree network with N stations. Let Q be the maximum number of stations that can be connected to a single fiber plant which satisfies power budget constraint. Assuming that each cluster has the same number of stations, we have that the number of stations per cluster is $M \leq Q/2$. The number of clusters is then $K \geq N/M \geq (2N)/Q$. Using this inequality and expression (5.2) we get that the following inequality for the maximum number of fiber plants

$$F_{\max} \geq (\frac{N}{Q})^2 \tag{5.3}$$

From this formula we see that in the multifiber tree network the required number of fibers per cable grows quadratically with the number of stations and that the required number of fibers can become very large. For example, to support a 2048 station network with the maximum number of station per fiber plant $Q = 64$ (i.e., the maximum number of stations in a tree network without amplifiers for power budget of 40 dB) we need 1024 fibers per cable.

Once the intercluster fiber plants have been established (by properly configuring the switches inside the cable plant), a pair of stations that belong to different clusters can establish a single-hop "wavelength-on-demand" connection by connecting to the fiber plant common to their clusters and tuning to the common wavelength. The intercluster "links" of the multihop topology can be established
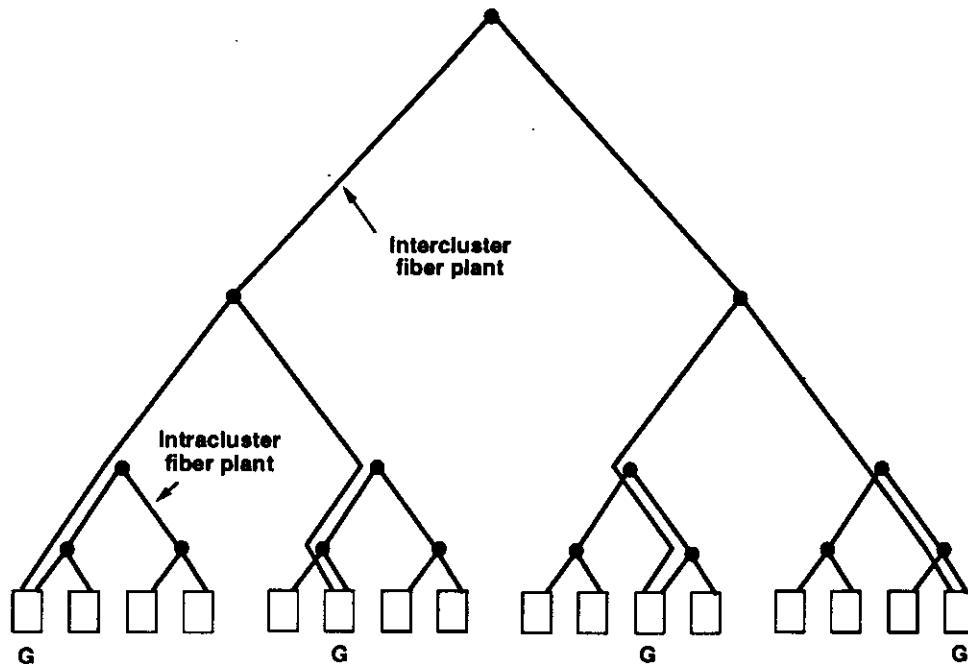
Figure 5.7: Intracluster and intercluster fiber plants for T/WDMA

in the same manner.

In order to implement intercluster T/WDMA, different fiber plants need to be established. Since intercluster T/WDMA covers all clusters, due to power budget limitation only a few stations from each cluster can be connected to the intercluster T/WDMA fiber plant. Figure 5.7 illustrates the implementation of T/WDMA intracluster and intercluster fiber plants. Only stations connected to the intercluster fiber plant can serve as the cluster gateways.

Figure 5.8 illustrates possible communications in a multilevel HONET that uses a large number of fiber plants. Each station is connected to two fiber plants: an intracluster single-hop fiber plant and a multihop fiber plant. A station that has an additional transmitter-receiver pair can be also connected to a set of single-hop "wavelength-on-demand" fiber plants that are used for high data rate point-to-point communication. Packet-switched traffic is established through the multihop fiber plants. Instead of having separate multihop and single-hop "wavelength on-demand" fiber plants, it would be more reasonable to have a single set of fiber plants which can be then shared by both the multihop and the single-hop traffic. For T/WDMA, however, a separate set of fiber plants is required. Intercluster circuit-switched traffic is established through the gateways connected to a T/WDMA intercluster fiber plant.
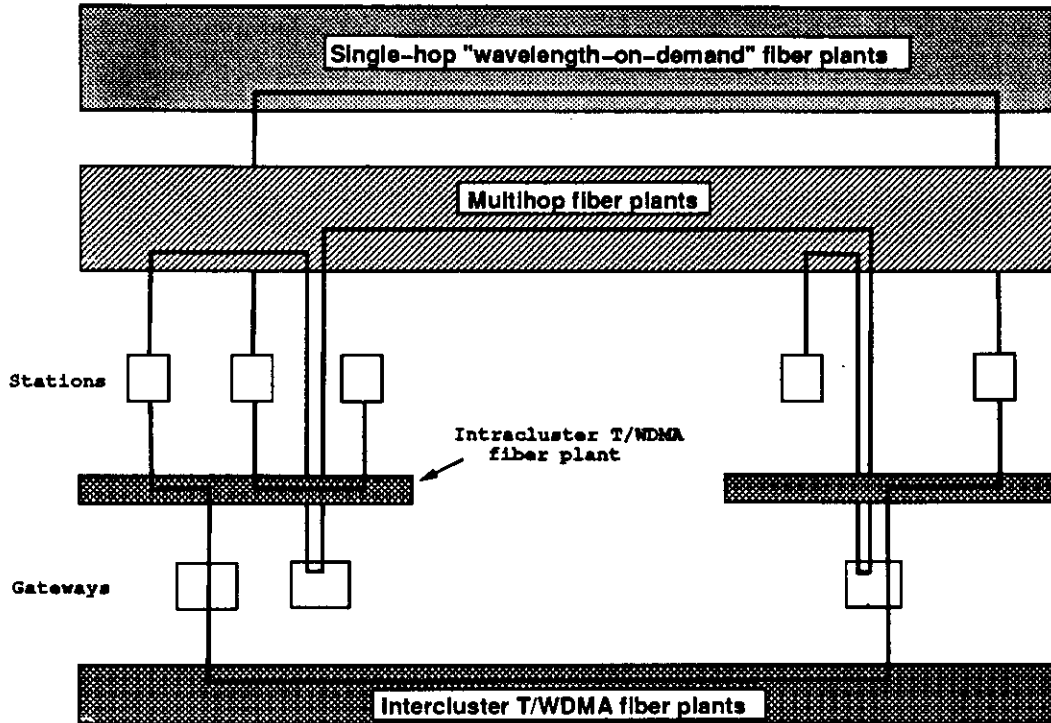
Figure 5.8: Communication in a multilevel and multifiber HONET

5.2.2 Multifiber LLN

The same multifiber idea can be applied to the Linear Lightwave Network (LLN) infrastructure. A multifiber LLN can be built by replacing passive couplers, reflective star couplers and splices at cable closets with Linear Divider-Combiner (LDC) devices. Using LDCs we can take advantage of the fact that each station is not active on all fiber plants to which it is connected at the same time. If each station has only one transceiver, it can be active only on one fiber plant. Thus, by changing power distribution and multiplexing of LDCs, we can change configuration of fiber plants dynamically and distribute optical power only to those stations that are active, which would allow more stations to be assigned to each fiber plant. Thus, the cluster size can be increased and the number of clusters reduced, which, in turn, reduces the required number of fibers to interconnect fiber plants. In case of intercluster T/WDMA fiber plants, dynamic reconfiguration provides more flexibility for gateway allocation. Any station can be connected to the intercluster fiber plant and thus selected as a gateway.

Let us determine now $F_{\max}$ for the multifiber LLN. Each station has access to $K - 1$ intercluster fiber plants. We assume that the station has also access to

an additional intracluster fiber plant. If we assume that each station is active on p fiber plants simultaneously and that each fiber plant has the same number of active stations, we have that the number of active stations per cluster is

$$M_a = \frac{Mp}{K} \tag{5.4}$$

In order to satisfy power budget constraints we have $M_a \leq Q/2$. We assume that Q does not depend on M which is true in the ideal case where no link and LDC excess losses exist. In reality, larger M may require larger fiber plants with more LDCs which results in higher excess losses and therefore smaller Q. We also have that $K \geq N/M$ which gives

$$K \geq \frac{2Np}{QK} \tag{5.5}$$

Solving for K, we get

$$K \geq \sqrt{\frac{2Np}{Q}} \tag{5.6}$$

Using expressions (5.2) and (5.6) we get

$$F_{\text{max}} \geq \frac{Np}{2Q} \tag{5.7}$$

We see that in the multifiber LLN the required number of fibers grows linearly with the number of stations. Thus, the use of the multifiber LLN instead of multifiber tree can significantly improve network scalability. We also see that the number of fibers depends on parameter p where $0 < p \leq K$. When $p = K$, i.e., when each station is active on all fiber plants all the time (which would require that each station has K transceivers), we get the same result for the number of fibers as for the multifiber tree. In practice, we can expect that each station has a small number of transceivers, typically only one or two. Parameter p can be even smaller than one if some of the stations are idle. If $p = 1$, $N = 2048$ and $Q = 64$ the required number of fibers per cable is 16.

In the ideal case where no link and coupling excess losses exists Q for LLN fiber plant is greater or at least equal to the one for the tree fiber plant. For example, when power margin is 40 dB, for a LLN fiber plant Q is equal to 100, as shown in Chapter 4, while for a binary tree fiber plant Q is equal to 64. We must, however, point out that LDC devices have higher losses than passive couplers and splices. This results in a smaller Q for multifiber LLN than for the multifiber tree network. However, we can expect that LDC losses will be reduced in the future with the advances in optical technology. Note that it is possible for the multifiber LLN to outperform the multifiber tree network even with smaller Q.

The cost of the multifiber LLN does not increase significantly compared to the cost of the single fiber LLN. The major cost increase can be attributed to the cost of multiple LDCs. The same network that controls LDCs in a conventional LLN can now control multiple LDCs.

CHAPTER 6

Conclusions and Future Research

The first contribution of this dissertation is the introduction of a new architecture for a local or metropolitan area fiber-optic network. This architecture is based on the concept originally proposed by Kazovsky at al. [KP92] where the multihop approach is used for packet-switched traffic and the single-hop approach for circuit switched traffic. The basic HONET architecture is a generalization of this idea. The HONET multihop network uses time division channel sharing which makes possible implementation of practically any virtual topology even with a single transmitter-receiver pair. We derived an upper bound for the throughput that can be achieved with any virtual topology that can be built with a given degree of channel sharing. In the case of a single transmitter and receiver, we found that a small degree of channel sharing, typically two or three, provides the optimal throughput. Using the ShuffleNet virtual topology it is possible to achieve throughput very close to the obtained upper bound. The throughput is significantly better than the one when no channel sharing is used. We also showed that when each station has more than one transmitter- receiver pair, the highest throughput is obtained when no channel sharing is used. Next, we compared HONET with classic single-hop networks and STARNET and showed how HONET can outperform these networks.

The second major contribution is the introduction of a new time and wavelength division multiaccess scheme for the broadcast star optical network that is used in the single-hop component of the HONET architecture. The proposed scheme is intended for circuit-switched traffic. The scheme can support a large number of small bandwidth applications and can be implemented using tunable devices of relatively slow tuning speed, such as acoustooptic tunable filters. It was demonstrated that the proposed T/WDMA scheme provides high throughput, supports multirate traffic and can be implemented in a cost-effective manner. The performance analysis of this scheme was presented for the point-to-point, single rate traffic case. An approximate analytical model was developed for the uniform traffic case. It was shown that the analytical model matches very closely the simulation results. The model was also employed to determine the optimal operating parameters for the network. A possible extension of the original scheme

103

which can exploit multiwavelength filtering capability of acoustooptic filters was also presented. We also performed simulation study of the basic scheme for non-uniform traffic characterized by strong traffic locality. The results show that both the multiwavelength filtering capability and the traffic locality can be used to improve performance of the proposed scheme.

The third major contribution is the proposed new routing scheme for Linear Lightwave Network (LLN) infrastructure. The new routing scheme called rooted routing overcomes a synchronization problem that exist in LLN that uses originally proposed shortest path routing and thus makes it possible to implement HONET's TDMA and T/WDMA schemes over LLN. An analysis of this new routing scheme with respect to power budget was presented and it was shown that the new routing scheme performs as well as the shortest path routing scheme. Moreover, the rooted routing scheme is more advantageous than the shortest path routing since it can effectively improve the power budget using a single optical amplifier.

The fourth major contribution is the proposed extension of the basic HONET to the multilevel and multifiber architectures in order to overcome scalability limitations and increase network capacity. The multilevel HONET is based on a hierarchical organization and exploits traffic locality and wavelength reuse. It also reduces power budget limitations, thus allowing more stations to be connected to the network. In order to further increase network capacity, we then proposed multifiber HONET. The multifiber HONET can be efficiently implemented using the multifiber tree topology [BGK93]. We also proposed an implementation of the multifiber HONET using the multifiber Linear Lightwave Network. We showed that the multifiber LLN can support a larger number of stations than the multifiber tree with the same number of optical fibers per cable. In fact, the number of fibers per cable in the case of the multifiber tree grows quadratically with the increase in the number of stations in the network while in the case of the multifiber LLN it grows linearly.

Many problems still need to be addressed. One of problems that has been touched upon but not sufficiently explored is related with broadcast and multicast traffic. Effective connection assignment algorithms for such traffic need to be developed and then performance analyzed.

The introduction of multifiber Linear Lightwave Networks opens new interesting and challenging problems. In this dissertation we have addressed one specific multifiber LLN configuration that is equivalent to the one of the multifiber tree network. The multifiber LLN can have, however, a general mesh topology, and it can provide much more potential for fiber reuse than the tree topology. It is

likely that some other configuration can support more stations with a smaller number of fibers. Such an architecture can become a good candidate for a future all-optical wide area network (WAN).

It would be also interesting to expand the single hop architecture with multiple fibers. In such a case the T/WDMA scheme can be expanded into the S/T/WDMA (space, time and wavelength division multiaccess) scheme. If lithium niobate switches are used to connect stations to multiple fibers, switching between fibers can be done almost instantaneously due to the subnanosecond speed of those switches [Hin87]. In such a case both transmitters and receivers would be able to switch between fibers.

APPENDIX A

Upper bound for the maximum propagation delay in LLN with routed routing

Fact: The maximum propagation delay of a LLN with rooted routing is at most twice the maximum propagation delay of the same LLN with shortest path routing provided that the degree of the root node is at least two.

Proof: If the degree of the root node is at least two, and the topology is a tree, at least one pair of stations has the shortest path that goes through the root node. In such a case the maximum propagation delay from a station to the root (or vice versa) cannot be larger than the maximum propagation delay between stations connected by the shortest paths. The maximum propagation delay in a LLN with rooted routing is twice the maximum propagation delay between a station and the root. Therefore, the maximum propagation delay in a LLN with rooted routing is at most twice the maximum propagation delay between stations that are connected by the shortest paths.

APPENDIX B

Power budget optimization for LLN with rooted routing

In this Appendix, we present a solution for the power budget optimization problem in a LLN with rooted routing taking into account all types of losses.

In a LLN with rooted routing an optical signal sent by a station is attenuated on its way to the root due to the combining and excess losses at the nodes and the link losses at the links on the path from the station to the root. Also, an optical signal coming out of the root node is attenuated due to the splitting and excess losses at the nodes and the link losses at the links on the path from the root to a station. In order to minimize the maximum power loss between any two stations, i.e. the power budget, we can partition this problem into three subproblems and solve them separately. We need to

1. minimize the maximum power loss from a station on its path to the root node,

2. minimize the maximum power loss from the root to a station, and

3. given the minimized losses between stations and the root, optimize power distribution /multiplexing at the root node such that the maximum power loss between any two stations is minimal.

In order to find the power transfer matrix for a non-root node, we need to solve the first two problems. In non-root node l signals from all its input links (except from link $r^{(l)}$) are multiplexed to output link $r^{(l)}$. Due to the limitation in combining the power to output port $r^{(l)}$, only a portion of the power from each of the input links is transferred to $r^{(l)}$. The signals coming from input link $r^{(l)}$ to node l are distributed to all its output links except $r^{(l)}$.

The maximum loss of any signal transmitted by station $s \in S_i^{(l)}$ is minimal at input port i of node l if the loss is the same for each of the stations in $S_i^{(l)}$. The maximum attenuation of a signal from output port i of node l at station $s \in S_i^{(l)}$ is minimal if splitting is performed in such a way that each of the stations in $S_i^{(l)}$ receives an equal portion of power. Using these facts, we can derive the

109

expressions for the optimal power loss from station $s \in S_i^{(l)}$ to input port i of node l and from output port i of node l to station $s \in S_i^{(l)}$. Let

$$p = m_i^{(l)}$$

$$Lopt_i^{in(l)} = \begin{cases} D_{pl} & \text{if p is a station} \\ (\sum_{k=1, k \neq r(p)}^{K^{(p)}} Lopt_k^{in(p)}) E^{(p)} D_{pl} & \text{if p is a node} \end{cases} \tag{B.1}$$

$$Lopt_i^{out(l)} = \begin{cases} D_{lp} & \text{if p is a station} \\ (\sum_{k=1, k \neq r(p)}^{K^{(p)}} Lopt_k^{out(p)}) E^{(p)} D_{lp} & \text{if p is a node} \end{cases} \tag{B.2}$$

In order to achieve these power losses, elements of the power transfer matrix for each non-root node l should have the following values:

$$a_{ij}^{(l)} = \delta_{ij}^{(l)} = \frac{Lopt_j^{out(l)}}{\sum_{k=1, \, k \neq i}^{K^{(l)}} Lopt_k^{out(l)}}, \quad i = r^{(l)}, j \neq r^{(l)} \tag{B.3}$$

$$a_{ji}^{(l)} = \sigma_{ji}^{(l)} = \frac{Lopt_j^{in(l)}}{\sum_{k=1, \, k \neq i}^{K^{(l)}} Lopt_k^{in(l)}}, \quad i = r^{(l)}, j \neq r^{(l)} \tag{B.4}$$

$$a_{ii}^{(l)} = 0, \quad i = r^{(l)} \tag{B.5}$$

$$a_{ij}^{(l)} = 0, \quad i, j \neq r^{(l)} \tag{B.6}$$

For the root node we have

$$\delta_{ij}^{(R)} = \frac{Lopt_j^{out(R)}}{\sum_{k=1}^{K^{(R)}} Lopt_k^{out(R)}}$$

$$\sigma_{ij}^{(R)} = \frac{Lopt_i^{in(R)}}{\sum_{k=1}^{K^{(R)}} Lopt_k^{in(R)}}$$

and

$$a_{ij}^{(R)} = \delta_{ij}^{(R)} \sigma_{ij}^{(R)}, \quad i, j \in \{1, \ldots, K^{(R)}\} \tag{B.7}$$

Let C_{ij} be the power loss between stations i and j expressed as ratio between power transmitted by station i and power received by station j. We have that

$$\begin{aligned} C_{ij} &= Lopt_i^{in(R)} (a_{ij}^{(R)})^{-1} E^{(R)} Lopt_j^{out(R)} \\ &= (\sum_{k=1}^{K^{(R)}} Lopt_k^{in(R)}) E^{(R)} (\sum_{k=1}^{K^{(R)}} Lopt_k^{out(R)}), \quad i, j \in \{1, \ldots, N\} \end{aligned} \tag{B.8}$$

110

We see that C_{ij} is independent of i and j. Therefore, the power loss between any pair of stations has the same value which is in fact the power budget. The optimal power budget (expressed in decibels) is thus

$$P[\text{dB}] = 10 \log_{10} C_{ij} = 10 \log_{10} \left(\left(\sum_{k=1}^{K^{(R)}} Lopt_k^{in(R)} \right) E^{(R)} \left(\sum_{k=1}^{K^{(R)}} Lopt_k^{out(R)} \right) \right) \qquad (B.9)$$

If the links have the same losses in both directions (e.g., $D_{ij} = D_{ji}$) we have from (B.1) and (B.2) that

$$Lopt_i^{in(l)} = Lopt_i^{out(l)} = Lopt_i^{(l)} \qquad (B.10)$$

which allows us to simplify the formula for the optimal power budget. In such a case we have

$$P[\text{dB}] = 10 \log_{10} \left(E^{(R)} \left(\sum_{k=1}^{K^{(R)}} Lopt_k^{(R)} \right)^2 \right) \qquad (B.11)$$

111

APPENDIX C

Optimal power budget for LLN with shortest path routing in the ideal case

Fact: The optimal power budget in a LLN with shortest path routing in the ideal case (without the link and excess losses) is

$$P[\text{dB}] = 10 \log_{10}(N-1)^2$$

Proof: Let us consider the power loss on the shortest path from station A to station B as shown in Figure C.1. The nodes on the path from A to B are
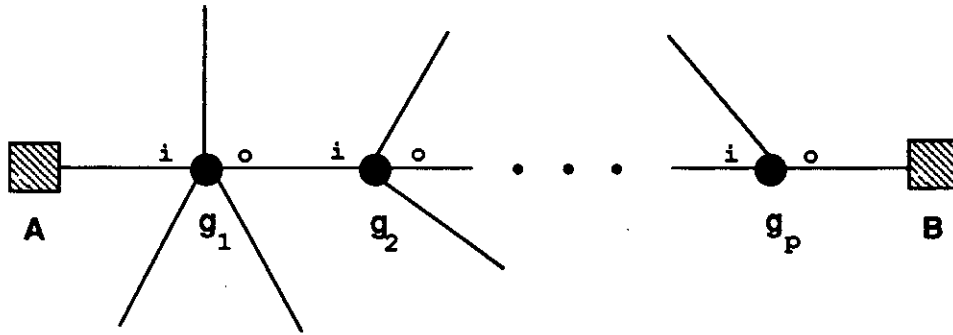
Figure C.1: Path between stations A and B

labeled $g_1, g_2, \ldots, g_p$ where p is the number of nodes on the path. For each of these nodes we label the incoming and the outgoing port on the path as i and o, respectively. Let C_{AB} be the power loss on the path from A to B. We have

$$C_{AB} = (\prod_{l=1}^{p} a_{io}^{(g_l)})^{-1}$$

When power distribution is optimal we have that

$$\delta_{io}^{(g_l)} = \frac{n_o^{(g_l)}}{\sum_{k=1, k \neq i}^{K^{(g_l)}} n_k^{(g_l)}} = \frac{n_o^{(g_l)}}{N - n_i^{(g_l)}}$$

and

$$\sigma_{io}^{(g_l)} = \frac{n_i^{(g_l)}}{\sum_{k=1, k \neq o}^{K^{(g_l)}} n_k^{(g_l)}} = \frac{n_i^{(g_l)}}{N - n_o^{(g_l)}}$$

which gives

$$a_{io}^{(g_l)} = \frac{n_o^{(g_l)}}{N - n_i^{(g_l)}} \frac{n_i^{(g_l)}}{N - n_o^{(g_l)}}$$

and

$$C_{AB} = \frac{N - n_i^{(g_1)}}{n_o^{(g_1)}} \frac{N - n_o^{(g_1)}}{n_i^{(g_1)}} \cdot$$

$$\cdot \frac{N - n_i^{(g_2)}}{n_o^{(g_2)}} \frac{N - n_o^{(g_2)}}{n_i^{(g_2)}} \cdots \frac{N - n_i^{(g_p)}}{n_o^{(g_p)}} \frac{N - n_o^{(g_p)}}{n_i^{(g_p)}}$$

Since

$$n_o^{(g_l)} = N - n_i^{(g_{l+1})}$$

the expression for C_{AB} can be simplified. Thus,

$$C_{AB} = \frac{N - n_i^{(g_1)}}{n_i^{(g_1)}} \frac{N - n_o^{(g_p)}}{n_o^{(g_p)}}$$

Because $n_i^{(g_1)} = n_o^{(g_p)} = 1$, we have

$$C_{AB} = (N - 1)^2$$

Since the power losses are the same for all pairs of stations, we have the following result for the power budget

$$P[\text{dB}] = 10 \log_{10} C_{AB} = 10 \log_{10}(N - 1)^2 \qquad (\text{C.1})$$

APPENDIX D

Power budget optimization for LLN with rooted routing and optical amplification at the root

In this appendix we consider the power optimization in a LLN with rooted routing that has a self-loop link at the root node. An optical amplifier with amplification gain $G[\text{dB}]$ is installed at this link.

Introduction of the loop link does not change the optimal power distribution and multiplexing at the non-root nodes. Consequently, the power transfer matrices for the non-root nodes are also defined by the expressions (B.3), (B.4), (B.5) and (B.6). The optimal power distribution and multiplexing at the root node is now done in the same manner as for the non-root nodes. Thus, the elements of the power transfer matrix of the root node have the following values:

$$a_{ij}^{(R)} = \delta_{ij}^{(R)} = \frac{Lopt_j^{out(R)}}{\sum_{k=1,\, k\neq i}^{K^{(R)}} Lopt_k^{out(R)}}, \quad i = r^{(R)}, j \neq r^{(R)} \tag{D.1}$$

$$a_{ji}^{(R)} = \sigma_{ji}^{(R)} = \frac{Lopt_j^{in(R)}}{\sum_{k=1,\, k\neq i}^{K^{(R)}} Lopt_k^{in(R)}}, \quad i = r^{(R)}, j \neq r^{(R)} \tag{D.2}$$

$$a_{ij}^{(R)} = 0, \quad i = j = r^{(R)} \tag{D.3}$$

$$a_{ij}^{(R)} = 0, \quad i, j \neq r^{(R)} \tag{D.4}$$

The optimal power budget is

$$
\begin{aligned}
P[\text{dB}] = {} & 10\log_{10}(\sum_{k=1}^{K^{(R)}} Lopt_k^{in(R)} E^{(R)} D_{RR} \cdot \\
& \cdot E^{(R)} \sum_{k=1}^{K^{(R)}} Lopt_k^{out(R)}) = \\
= {} & 10\log_{10}((E^{(R)} \sum_{k=1}^{K^{(R)}} Lopt_k^{(R)})^2 D_{RR})
\end{aligned}
\tag{D.5}
$$

115

where D_{RR} represents the power loss at the loop link. Since amplification is performed on the link, this loss is negative (i.e., $G[\text{dB}] = -10 \log_{10} D_{RR}$). Thus, we have

$$P[\text{dB}] = 10 \log_{10}(E^{(R)} \sum_{k=1}^{K^{(R)}} Lopt_k^{(R)})^2 - G[\text{dB}] \qquad \text{(D.6)}$$

REFERENCES

[ACG86] E. Arthurs, J. M. Cooper, M. S. Goodman, H. Kobrinski, M. Tur, and M. P. Vecchi. "Multiwavelength Optical Crossconnect for Parallel-processing Computers." *Electronic Letters*, **24**:119–120, 1986.

[AGK88] E. Arthurs, M. S. Goodman, H. Kobrinski, and M. P. Vecchi. "HY-PASS: An optoelectronic hybrid packet-switching system." *IEEE Journal on Selected Areas in Communications*, **6**(9):1500–1510, December 1988.

[AKB92] R.C. Alferness, U. Koren, L. L. Buhl, M. G. Young, T. L. Koch, G. Raybon, and C. A. Burrus. "Broadly Tunable InGaAsP/InP Laser Based on a Vertical Coupler Filter with 57-nm Tuning Range." *Applied Physics Letters*, **60**(26):3209–3211, June 1992.

[AKH87] A. S. Acampora, M. J. Karol, and M. G. Hluchyj. "Terabit Lightwave Networks: The Multihop Approach." *AT&T Technical Journal*, **66**(6):21–34, Nov./Dec. 1987.

[AS91] A. S. Acampora and S. I. A. Shah. "Multihop Lightwave Networks: A Comparison of Store-and-Forward and Hot-Potato Routing." In *Proceedings of IEEE INFOCOM '91*, volume 1, pp. 10–19, Bal Harbour, Florida, April 1991.

[ASZ92] A. S. Acampora, S. I. A. Shah, and Z. Zhang. "Performance Analysis of Hot Potato Routing for Multiclass Traffic in Multihop Lightwave Networks." In *Proceedings of IEEE INFOCOM '92*, volume 2, pp. 644–655, Florence, Italy, May 1992.

[Aya89] E. Ayanoğlu. "Signal Flow Graphs for Path Enumeration and Deflection Routing Analysis in Multihop Networks." In *Proceedings of GLOBECOM '89*, pp. 1022–1029, Dallas, Texas, November 1989.

[BC89] R. Ballart and Y. Ching. "SONET: Now It's the Standard Optical Network." *IEEE Communications Magazine*, **27**(3):8–15, March 1989.

[Ber92] L. Bergman et al. "CASA Gigabit Network Testbed: 1992 Annual Report." Technical Report CCSF-20-92, California Institute of Technology, Pasadena, California, April 1992.

117

[BFG90] J. A. Bannister, L. Fratta, and M. Gerla. "Topological Design of the Wavelength-Division Optical Network." In *Proceedings of IEEE INFOCOM '90*, volume 3, pp. 1005–1013, San Francisco, California, June 1990.

[BFG91] J. A. Bannister, L. Fratta, and M. Gerla. "Detour Routing in High-Speed Multichannel Networks." In Marjory J. Johnson, editor, *Protocols for High-Speed Networks, II: Proceedings of the IFIP WG 6.1/WG 6.4 Second International Workshop on Protocols for High-Speed Networks*, pp. 75–90. Elsevier Science Publishers, Amsterdam, The Netherlands, 1991.

[BFT91] F. Borgonovo, L. Fratta, and F. Tonelli. "Circuit Service in Deflection Networks." In *Proceedings of IEEE INFOCOM '91*, volume 1, pp. 69–76, Bal Harbour, Florida, April 1991.

[BG90] J. A. Bannister and M. Gerla. "Design of the Wavelength-Division Optical Network." In *Proceedings of IEEE ICC '90*, pp. 962–967, Atlanta, Georgia, April 1990.

[BGK93] J. Bannister, M. Gerla, and M. Kovačević. "An All-Optical Multifiber Tree Network." *IEEE Journal of Lightwave Technology*, 1993. to appear in special issue on Broadband Optical Networks.

[BMS92] S. Banerjee, B. Mukherjee, and D. Sarkar. "Heuristic Algorithms for Constructing Near-Optimal Structures of Linear Multihop Lightwave Networks." In *Proceedings of IEEE INFOCOM '92*, volume 2, pp. 671–680, Florence, Italy, May 1992.

[Bra90] C. A. Brackett. "Dense Wavelength Division Multiplexing Networks: Principles and Applications." *IEEE Journal on Selected Areas in Communications*, 8(6):948–964, August 1990.

[BSB91] K. Bala, T. E. Stern, and K. Bala. "Algorithms for Routing in a Linear Lightwave Network." In *Proceedings of IEEE INFOCOM '91*, volume 1, pp. 1–9, Bal Harbour, Florida, April 1991.

[BSC90] M. Bagley, G. Sherlock, D.M. Cooper, L.D. Westbrook, et al. "Broadband Operation of InGaAsP-InGaAs GRIN-SC-MQW BH Amplifiers with 115mW Output Power." *Electronic Letters*, 26(8):512–513, April 1990.

[CCC88]   K. W. Cheung, G. Coquin and M. Choy. "Single- and Multiple-wavelength Operation of Acousto-optically tuned lasers at 1.3 microns." In *Proceedings of 11th IEEE International Semiconductor Laser Conference*, pp. 130–131, Boston, Massachusetts, 1988.

[CDR90]   M. Chen, N.R. Dono, and R. Ramaswami. "A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks." *IEEE Journal on Selected Areas in Communications*, 8:1048–1057, August 1990.

[CG88]   I. Chlamtac and A. Ganz. "Channel allocation protocols in frequency-time controled high speed networks." *IEEE Transactions on Communications*, **36**(4):430–440, April 1988.

[Che90]   K.W. Cheung. "Acoustooptic tunable filters in narrowband WDM networks: System issues and network applications." *IEEE Journal on Selected Areas in Communications*, 8:1015–1024, August 1990.

[CLS89]   K.W. Cheung, S.C. Liew, D.A. Smith, C.N. Lo, J.E. Baran, and J.J. Johnson. "Simultaneous five-wavelength filtering at 2.2 nm wavelength separation using an integrated-optic acoustic-optic tunable filter with subcarrier detection." *Electronic Letters*, **25**(10):636–637, 1989.

[CY91]   M. Chen and T.S. Yum. "A conflict-free protocol for optical WDMA networks." In *Proceedings of IEEE GLOBECOM '91*, pp. 1276–1281, Phoenix, Arizona, December 1991.

[Dar87]   T. E. Darcie. "Subcarrier Multiplexing for Multiple-Access Lightwave Networks." *Journal of Lightwave Technology*, **LT-5**(8):1103–1110, August 1987.

[DGL90]   N. R. Dono, P. E. Green, K. Liu, R. Ramaswami, and F. F. Tong. "A wavelength division multiple access network for computer communication." *IEEE Journal on Selected Areas in Communications*, 8:983–994, August 1990.

[Dow91]   P. W. Dowd. "Random Access Protocols for High-Speed Interprocessor Communication Based on an Optical Passive Star Topology." *Journal of Lightwave Technology*, 9(6):799–808, June 1991.

[Dow92]   P. W. Dowd. "Wavelength Division Multiple Access Channel Hypercube Processor Interconnection." *IEEE Transactions on Computers*, **41**(10):1223–1241, October 1992.

[DQD88] "Distributed Queue Dual Bus (DQDB) Metropolitan Area Network (MAN)." Proposed Standard IEEE P802.6/D6-88/105, November 1988. Institute of Electrical and Electronics Engineers.

[Dra89] C. Dragone. "Efficient $N \times N$ Star Couplers Using Fourier Optics." *Journal of Lightwave Technology*, **LT-7**(3):479–489, March 1989.

[GF88] M. Gerla and L. Fratta. "Tree Structured Fiber Optic MAN's." *IEEE Journal on Selected Areas in Communications*, **SAC-6**(6):934–943, July 1988.

[GK91] A. Ganz and Z. Koren. "WDM passive star - protocols and performance analysis." In *Proceedings of IEEE INFOCOM '91*, pp. 991–1000, Bal Harbour, Florida, April 1991.

[GKB92] M. Gerla, M. Kovačević, and J. Bannister. "Multilevel Optical Networks." In *Proceedings of IEEE ICC '92*, volume 3, pp. 1168–1172, Chicago, Illinois, June 1992.

[GKV90] M.S. Goodman, H. Kobrinski, M. Vecchi, R.M. Bulley, and J.L. Gimlett. "The LAMBDANET multiwavelength network: Architecture, applications, and demonstrations." *IEEE Journal on Selected Areas in Communications*, **8**(6):995–1004, August 1990.

[GL92] A. Ganz and B. Li. "Broadcast-Wavelength Architecture for a WDM Passive Star-Based Local Area Network." In *Proceedings of IEEE ICC '92*, volume 2, pp. 837–842, Chicago, Illinois, June 1992.

[Gre93] P. E. Green. *Fiber Optic Networks*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[Hen89] P. S. Henry. "High-capacity Lightwave Local Area Networks." *IEEE Communications Magazine*, **27**(10):20–26, October 1989.

[Hin87] H. S. Hinton. "Photonic Switching Using Directional Couplers." *IEEE Communications Magazine*, **25**(5):16–26, May 1987.

[HK88] M. G. Hluchyj and M. J. Karol. "ShuffleNet: An Application of Generalized Perfect Shuffles to Multihop Lightwave Networks." In *Proceedings of IEEE INFOCOM '88*, pp. 379–390, New Orleans, Louisiana, March 1988.

[HKS87] I.M.I. Habbab, M. Kavehrad, and C. W. Sundberg. "Protocols for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology." *Journal of Lightwave Technology*, **LT-5**(12):1782–1794, December 1987.

[Hos90] R. J. Hoss. *Fiber Optic Communication Design Handbook*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.

[HRS92] P. A. Humblett, R. Ramaswami, and K. N. Sivarajan. "An Efficient Communication Protocol for High-speed Packet-switched Multichannel Networks." In *Proceedings of the ACM SIGCOMM '92 Symposium*, pp. 2–13, Baltimore, Maryland, August 1992.

[JCC92] V. Jayaraman, D. A. Cohen, and L. A. Coldren. "Demonstration of Broadband Tunability in a Semiconductor Laser Using Sampled Gratings." *Applied Physics Letters*, **60**(19):2321–2323, May 1992.

[JM92] F. Jia and B. Mukherjee. "The Receiver Collision Avoidance (RCA) Protocol For A Single-Hop WDM Lightwave Network." In *Proceedings of IEEE ICC '92*, volume 1, pp. 6–10, Chicago, Illinois, June 1992.

[KA88] S. K. Korotky and R. C. Alferness. "Waveguide electrooptic devices for optical fiber communication." In S.E. Miller and I. Kaminow, editors, *Ch. 11 of Optical Fiber Telecommunications II*. Academic Press, 1988.

[Kle76] L. Kleinrock. *Queueing Systems, Volume II: Computer Applications*. John Wiley and Sons, New York, New York, 1976.

[KP92] L. Kazovsky and P. Poggiolini. "STARNET: a multiGbit/s optical local area network with both packet and circuit switching." In *Proceedings of OFC '92*, pp. 297–298, San Jose, California, January 1992.

[LA91] J. F. Labourdette and A. S. Acampora. "Wavelength Agility in Multihop Lightwave Networks." *IEEE Transactions on Communications*, **29**(8):1223–1230, August 1991.

[LG92] B. Li and A. Ganz. "Virtual Topologies for WDM Star LANs - The Regular structure approach." In *Proceedings of IEEE INFOCOM '92*, volume 3, pp. 2134–2143, Florence, Italy, May 1992.

[LK92a] J. Lu and L. Kleinrock. "On the Performance of Wavelength Division Multiple Access Networks." In *Proceedings of IEEE ICC '92*, volume 3, pp. 1151–1157, Chicago, Illinois, June 1992.

121

[LK92b]   J. Lu and L. Kleinrock. "A Wavelength Division Multiple Access Protocol for High-speed Local Area Networks with a Passive Star Topology." *Performance Evaluation*, **16**(1-3):223–239, November 1992.

[Max87]   N. F. Maxemchuk. "Routing in the Manhattan Street Network." *IEEE Transactions on Communications*, **COM-35**(5):503–512, September 1987.

[Meh90]   N. Mehravari. "Performance and Protocol Improvements for Very High Speed Optical Fiber Local Area Networks Using a Passive Star Topology." *Journal of Lightwave Technology*, **8**(4):520–530, April 1990.

[Min89]   S. Minzer. "Broadband ISDN and Asynchronous Transfer Mode (ATM)." *IEEE Communications Magazine*, **27**(9):17–24, September 1989.

[Min91]   W. J. Miniscalco. "Erbium-doped Glasses for Fiber Amplifiers at 1500nm." *Journal of Lightwave Technology*, **9**(2):234–250, February 1991.

[MMK87]   S. Murata, I. Mito, and K. Kobayashi. "Over 720 GHz (5.8nm) Frequency Tuning by a 1.5 Micrometer DBR Laser and Bragg Wavelength Control Regions." *Electronic Letters*, **23**(8):403–405, April 1987.

[NNA91]   K. Nakagawa, S. Nishi, K. Aida, and E. Yoneda. "Trunk and Distribution Network Application of Erbium-doped Fiber Amplifier." *Journal of Lightwave Technology*, **9**(2), February 1991.

[RH90]   R. Ramaswami and P. A. Humblett. "Amplifier Induced Crosstalk in Multi-channel Optical Networks." *IEEE Journal of Lightwave Technology*, **8**(12):1882–1896, December 1990.

[Ros86]   F. E. Ross. "FDDI—A Tutorial." *IEEE Communications Magazine*, **24**(5):10–17, May 1986.

[RS93]   R. Ramaswami and K. N. Sivarajan. "A Packet-Switched Multihop Lightwave Network Using Subcarrier and Wavelength Division Multiplexing." *IEEE Transactions on Communications*, 1993. to appear.

[SEL91]   N. K. Shankaranarayanan, S. D. Elby, and K. Y. Lau. "WDMA/Subcarrier-FDMA Lightwave Networks: Limitations due to Optical Beat Interference." *Journal of Lightwave Technology*, **9**(7):931–943, August 1991.

[SGK91]   G.N.M. Sudhakar, N.D. Georganas, and M. Kavehrad. "Slotted Aloha and Reservation Aloha Protocols for Very High-speed Optical Fiber Local Area Networks Using Passive Star Topology." *IEEE Journal of Lightwave Technology*, **9**(10):1411–1422, November 1991.

[SH93]   G. Semaan and P. Humblet. "Timing and Dispersion in WDM Optical Star Networks." In *Proceedings of IEEE INFOCOM '93*, volume 2, pp. 573–577, San Francisco, California, March 1993.

[SR91]   K. Sivarajan and R. Ramaswami. "Multihop Lightwave Networks Based on de Bruijn Graphs." In *Proceedings of IEEE INFOCOM '91*, volume 3, pp. 1001–1011, Bal Harbour, Florida, April 1991.

[Sta93]   W. Stallings. *Local and Metropolitan Area Networks*. Macmillan Publishing Company, New York, New York, fourth edition, 1993.

[Ste90]   T. E. Stern. "Linear Lightwave Networks: How Far They Can Go?" In *Proceedings of GLOBECOM '90*, pp. 1866–1872, San Diego, California, December 1990.

[Ste91]   T. E. Stern. "A Linear Lightwave MAN Architecture." In Guy Pujolle, editor, *High-Capacity Local and Metropolitan Area Networks*, pp. 161–179. Springer-Verlag, Berlin, Germany, 1991.

[Tek90]   V. J. Tekippe. "Passive Fiber-Optic Components Made by the Fused Biconical Taper Process." *Fiber and Integrated Optics*, **9**:97–123, 1990.

[Vaz91]   A. Vaziri. "Scientific Visualization in High-speed Network Environments." *Computer Networks and ISDN Systems*, **22**(2):111–130, September 1991.

[WD83]   R. Wyatt and W. J. Delvin. "10kHz Linewidth $1.5\mu m$ InGaAsP External Cavity Laser with 55 nm Tuning Range." *Electronic Letters*, **19**:110–112, 1983.

[WHA88]   W. Warzanskyj, F. Heisman, and R. C. Alferness. "Polarization-independent Electro-optically Tunable Narrow-band Wavelength Filter." *Applied Physics Letters*, **53**(1):1151–1159, July 1988.