WAVELENGTH DIVISION MULTIPLEXING CROSS CONNECT NETWORKS

BY
NIMA AHMADVAND, B.SC., M.SC.
MAY 1997

A THESIS
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF MCMASTER UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

© Copyright 1997 by Nima Ahmadvand, B.Sc., M.Sc.
All Rights Reserved

WDM CROSS CONNECT NETWORKS

Doctor of Philosophy (1997)
(Electrical and Computer Engineering)

McMaster University
Hamilton, Ontario

TITLE:	Wavelength Division Multiplexing Cross Connect Networks
AUTHOR:	Nima Ahmadvand, B.Sc., M.Sc. (Sharif University of Technology, Tehran, Iran)
SUPERVISOR:	Dr. Terence D. Todd, Associate Professor of Electrical and Computer Engineering, B.A.Sc., M.A.Sc., Ph.D. (University of Waterloo)
NUMBER OF PAGES:	xix,160

Abstract

Fiber optic communication is only 30 years old, but it has already become one of the technologies of choice for current and future high-speed communications. Unfortunately, the enormous bandwidth available in optical fiber cannot be accessed as a single communication channel. This is because each user station typically connects to the network through its electronic interface whose speed is limited to the maximum electronic processing speed. However, access to this bandwidth is possible using parallel channel architectures, where each channel operates at rates which are accessible to electronic processing. Currently, wavelength-division-multiplexing (WDM) is being considered for this purpose. In local and metropolitan area networks, WDM may offer increased transmission concurrency compared with conventional single-channel designs, with only modest increases in station complexity. However, because of device limitations the number of available wavelength channels may initially be less than desired. As a result, spatial reuse of wavelength channels may be required to obtain designs which will support a reasonable number of stations. One option in this case is to perform media access across a spatial wavelength cross connect. In such a system, the total capacity may be n times that of a single passive star network, where n is the number of available wavelengths.

In this thesis, WDM cross connect networks are considered as a specific category of optical networks, which efficiently reuse the available wavelengths. The proposed networks are classified into two classes: single-hop and dual-hop.

In the single-hop case, time slot assignment techniques are considered which can dynamically schedule communications across the wavelength cross connect network. These new scheduling techniques are developed for a high performance dynamic medium access control based on the traffic demands of the network stations. Some simple random access methods are also introduced for the cases where some higher levels of simplicity are sought. Simulation programs are used to compare different scheduling methods. The results also show that the proposed methods perform quite efficiently.

The dual-hop class of WDM cross connect networks proposes a shared buffering scheme which simplifies network operation. These networks consist of two stages. In the first, the wavelength tunability of the user stations is used to route packets from a given local optical network (LON) to a destination LON. When packets arrive at a destination LON, they are buffered and transmitted onto the required destination wavelength. There are a number of advantages to this design including the elimination of protocols which would require both dynamic station transmitter and receiver tunability. This design also takes advantage of increasingly available commercial ATM (Asynchronous Transfer Mode) buffer/switch components. Several hybrid electro-optic options are considered. It is also shown that by using a novel optical multichannel buffering scheme, the number of required buffers can be significantly decreased.

Acknowledgments

I take this opportunity to thank my supervisor, Dr. Terry Todd, who has always treated me as a friend and a colleague, for his unfailing advice and invaluable assistance throughout my doctoral studies. I wish to thank my fellow graduate students in the Optical and Wireless Network Group, Adrian Grah, Mark Janoska, Sabu Joseph, Rupam Sinha, Faisal Shad, and Charbel Sakr, with whom I have shared an office for the past four years, and discussed communication networks in detail. I owe many thanks to my parents, Haji and Aghdass, my brother, Afshin, and my sister, Afrooz, for their enthusiastic support and patience. I would like to thank my lovely wife, Shahla, for her support, love and passion.

I would also like to thank my good friends at McMaster University, Farid Abrari, Mohsen Ghaemian, and Kianoosh Hatami, with whom I have spent a major part of the past four years discussing everything in the world.

Dedicated to my parents,
my brother, my sister,
and my wife

Contents

List of Tables

List of Figures

List of Acronyms

APD	Avalanche Photo Diode
ATM	Asynchronous Transfer Mode
BER	Bit Error Rate
C^3	Cleaved Coupled Cavity
CATV	Community Antenna TeleVision
CC	Control Channel
CDMA	Code Division Multiple Access
CSMA	Carrier Sense Multiple Access
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DBR	Distributed Bragg Reflector
DCCN	Distributed Channel Controller Network
DFB	Distributed FeedBack
DQDB	Distributed Queue Dual Bus
DeMux	Demultiplexer
EDCCN	Extended Distributed Channel Controller Network
EMI	ElectroMagnetic Interference
FBT	Fused Biconical Taper
FDDI	Fiber Distributed Data Interface
FDM	Frequency Division Multiplexing
FDMA	Frequency Division Multiple Access
FM	Frequency Modulation

Fixed Rx	Fixed-tuned Receiver
Fixed Tx	Fixed-tuned Transmitter
Gbps	Giga bits per second
HDTV	High Definition TeleVision
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Standard Organization
ITDMA	Interleaved Time Division Multiple Access
kbps	kilo-bit per second
L-MAC	Local Medium Access Control
LAN	Local Area Networks
LC	Local Controller
LED	Light Emitting Diode
LLC	Logical Link Control
LON	Local Optical Network
MAC	Medium Access Control
MAN	Metropolitan Area Network
MANDALA	Multiple-access-network Architecture employing Network Division And Light-wavelength Assignment
MC	Master Controller
MD	Modal Decomposition
MDLB	Multiple Delay Line Buffer
Mbps	Mega-bit per second
OSI	Open System Interconnection
QDS	Quasi Doubly Stochastic
RFI	Radio Frequency Interference
RJE	Remote Job Entry
Rx	Receiver
SDLB	Shared Delay Line Buffer
SM	Switching Mode
SNR	Signal-to-Noise Ratio

SS/TDMA	Satellite Switched Time Division Multiple Access
TDMA	Time Division Multiple Access
TSA	Time Slot Assignment
Tbps	Tera-bit per second
Tx	Transmitter
VLSI	Very Large Scale Integration
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WDMA	Wavelength Division Multiple Access
WR-ITDMA	Wavelength Routed Interleaved Time Division Multiple Access
WSC	Wavelength Selective Coupler
WXSM	WDM cross(X) connect Switching Mode

Chapter 1

Introduction

1.1 Overview

Fiber optic communications was viewed as a speculative technology as recently as 30 years ago, even though it has now become one of the technologies of choice for high-speed communications. It is likely that in the near future, optical fibers will extend directly into the home as well. This will enable a tremendous variety of broadband services supported by the public telecommunications network. The enormous bandwidth of the single-mode optical fiber in the low-loss wavelength windows at 1.3 μm and 1.5 μm has in theory, enough capacity to carry traffic equivalent to all the telephone traffic in the entire U.S. during the busiest transmission times ($\approx 10^{12}$ bits per second). As the broadband fiber optic network evolves and new broadband applications are developed, new approaches will be required to exploit these large capacities [Ram93, Goo89, Muk92b, Bra90].

Emerging applications such as supercomputer interconnection, digital audio, medical imaging, high definition television (HDTV), distributed processing, supercomputer visualization, and the growth in the Internet require increasingly higher speed communications. There are several aspects regarding the design of multichannel communication networks that should be considered. Type and technology of the optical components, network topology, switching modes (packet switching or circuit switching), single-hop or multi-hop design, and their effects on the network performance and finally the cost-performance trade-off should be investigated.

Wavelength division multiplexing (WDM) is one of the most promising techniques to overcome the electronic/optical conversion bottlenecks at the edges of the network where user stations or network nodes are connected [Gre93, Bra90, Ram93, Goo89]. Since electronic circuits are used to send data through laser transmitters or receive data using optical detectors, the maximum data transmission speed on an all-optical network is still limited to the maximum electronic processing speed, roughly tens of Giga-bits per second. Each communication channel therefore cannot be run faster than this speed. Thus, concurrent parallel communication channels are needed to fully utilize the huge bandwidth in the optical fiber. If it was possible to have 1000 parallel channels each running at 1 Giga-bit per second, an aggregate capacity of 1 Tera-bit per second would be achievable. However, such a technology is not available yet and only a modest number of wavelength channels are currently practical. Using WDM, parallel channel architectures, where each channel operates at rates which are accessible to electronic processing are obtained with only modest increases in station complexity. On the other hand, other multiaccess methods such as time division multiple access (TDMA) and code division multiple access (CDMA) require synchronization within one time-slot and one chip time, respectively, which make them relatively less attractive than wavelength division multiple access (WDMA) [Muk92a].

All optical networks introduce the structure where the optical energy transmitted by a network node traverses the network to the destination without being converted to electronic signals in any part of the path [Gre93]. In this kind of network, wavelengths can be used for routing the optical signals to their final destinations. Thus, wavelength routing together with wavelength conversion can be a great asset in realizing all optical networks. The all optical approach also allows considerable *protocol transparency* and bandwidth-on-demand capability, since once a connection is set up at one wavelength, any bit rate, framing convention or higher level protocol stack may be used independently of what is used on the other connections. Some experiences with prototype all optical networks show that the photonic technology has much simpler structures than the nearest electronic equivalents. Therefore, it can also help to reduce the cost in large extent once the technology is mature. New levels of *reliability* and *availability* appear to be achievable from the all optical approach [G+93].

In the recent past, much effort has been devoted to protocol design for single-hop networks, using a passive optical star implementation. Separate inbound and outbound fibers connect each station to the star coupler. The optical energy received on any input of the star coupler is distributed evenly among and sent to all the output ports of the coupler. The broadcast-and-select nature of these networks brings up the issue of the medium access control (MAC) and contention resolution for a multichannel system. A number of issues regarding single-hop WDM networks are addressed in [Muk92b]. The complexity of the network hardware and related MAC are of great concern. The designer does his best to reduce these complexities to make a cost effective design. Consequently, single transmitter and single receiver designs with one of the devices being tunable is the minimum hardware configuration. In such a system, the total throughput is upper bounded by the number of available WDM channels [HKS87, Meh90, Muk92a, KSG91]. It is apparent that the number of available wavelengths may initially be less than desired. As a result, spatial reuse may be required in order to obtain a local area network which will support a reasonable number of stations. In [JT94a, KFG94, DBAP93], hierarchical single-hop architectures have been considered which permit sets of wavelengths to be reused across a number of local subnetworks. A single set of remote wavelengths is then used to maintain connectivity between them. This scheme can also take advantage of the traffic locality which is often associated with large local area networks (LANs) and metropolitan area networks (MANs). Unfortunately, when traffic flows are more uniform, the performance improvements may be very poor.

In the LAN design considered in [Mat93], media access is performed across a wavelength-routed spatial cross connect. Global coordination is achieved using a control channel implemented using additional optical hardware. This is used to schedule all packet transmissions through the cross connect. In such a system however, the control channel may easily become congested as traffic loading increases. In addition, dynamically tunable station transmitters and receivers are required to operate the system. This drastically complicates the operation of the system compared with that of a single star network.

These designs and other architectures and protocols proposed in the literature attempt to find the best answers for a number of questions that have been raised. These questions cover specifically the areas of network topology, station transmitter and receiver configuration, type of optical components and their limitations. For example, what is the best configuration in a multichannel environment? Should we use a multihop design, as the one used in traditional packet switched networks, or should a single hop design such as the one used in Ethernet be adopted? What are the effects of the network topology on the network performance? What is the optimal number of fixed or tunable transceivers at a station? Although multichannel optical devices and networks are in the early stages of development, many have devoted their efforts to this field to find answers to these questions. These objectives are well justified since multichannel optical networks have the potential to provide enormous increases in transmission capacity.

1.2 Motivations

Multichannel multi-access protocol designs based on WDM technology are proposed in this thesis. The fact that only a modest number of wavelength channels are available leads us to consider designs which spatially reuse the available channels. In the proposed architectures, stations are clustered and interconnected through a WDM cross connect where the packets received from different clusters are routed to their destination cluster according to wavelength. All the designs attempt to decrease the active component count, simplify the station hardware design and at the same time use very simple medium access control (MAC) protocols. Wavelength routing is used to alleviate electronic-optical conversions, and consequently reduce potential bottlenecks throughout the optical path.

The proposed architectures are classified into two main categories. In the first one, all packets traverse the network in one hop from the source to the final destination. Tunability in both transmitter and receiver side is needed. The proposed MAC protocols coordinate and schedule the accesses of different stations to each channel, and resolve contention on both the transmitter and the receiver side. Thus, these MAC protocols describe fast and optimum scheduling methods. A great improvement is achieved in system capacity by wavelength channel reuse.

In the second class of proposed network architectures, shared memories are introduced in the network to dynamically resolve destination contention. These networks consist of two stages. In the first, the wavelength agility of the sources is used to route packet transmissions from a given local optical network (LON) to a destination LON. As in the previous class, a wavelength cross connect is used for this purpose. When packets arrive at a destination LON, they are buffered and subsequently transmitted onto the desired destination wavelength. Each packet experiences two hops from source to destination, hence, these networks are referred to as *dual-hop*. Input access is determined using a local media access control (L-MAC) protocol. Since buffering is introduced into the LONs, slot synchronization is only required between stations associated with the same LON. This simplifies the station design since those stations will be in close physical proximity. Also in such a system, it is difficult to dynamically allocate destination slots because unlike the single passive star case, channel feedback is not available. In the proposed network this problem is alleviated because dynamic contention is resolved remotely, after the buffering stage. In addition, the dual-hop nature of the network eliminates the need for dynamic station receiver agility and the ensuing protocol control complexity. Finally, the design is a hybrid opto-electronic one which exploits the increasing availability of inexpensive commercial ATM (asynchronous transfer mode [SAH94, Onv94]) switch/buffering components.

Throughput and delay results show that the proposed architectures and the corresponding MAC protocols can achieve a noticeable improvement in system performance comparing to the other architectures found in literature.

Dynamic scheduling and MAC protocols are also presented for both classes of proposed architectures.

1.3 Scope of Thesis

The research presented in this thesis falls within the area of multi-channel multi-access data communication networks and medium access control protocol designs. It deals with packet switched network design using optical fiber as the communication medium and capabilities of the optical devices as determined from the current research literature.

Considering the open system interconnection (OSI) model (refer to Section 2.3.1) the proposed designs mainly deal with two lower layers, i.e. physical layer and the data link layer. The main emphasis, however, is on the medium access control (MAC), which is the bottom sub-layer of data link layer.

The proposed designs can also be categorized in terms of many other subdivisions and regions of specialty, such as link configuration, network span, and mode of data exchange. Link configuration refers to the number of transmitters and receivers which can be assigned to a single communication link. The design can be based on a broadcast-and-select topology where several transmitters and receivers are assigned to both transmit or receive from a single channel, or they can be designed such that each transmitter or receiver is pre-assigned to a specific channel. Some of the designs considered in this proposal are fully broadcast-and-select, while others are based on a partially broadcast-and-select topology. In the latter case, stations can transmit onto any channel but receive from a specific home channel.

The network span of the proposals focuses mainly on the local and metropolitan area environments, rather than for wide area networks. Both single-hop and dual-hop packet switched configurations are discussed, and a triple-hop configuration is presented to illustrate the possibility of architectural expansion of the proposed architectures.

Several medium access control protocols, from fully scheduled to random access, from distributed control to centralized control, are presented. Some of the designs are based on existing control channel schemes (or reused control channel) and in some, no feedback is necessary and all control is exercised at the edges of the network in a distributed manner.

1.4 Contributions

This thesis makes advancements in the use of WDM cross connect and wavelength routing schemes in the area of multichannel multiaccess optical networks. It reuses the modest number of available wavelength channels to accommodate relatively large numbers of stations. The proposed architectures and the corresponding MAC protocols are examples of a small number of architectures and protocols of this type.

The main contributions of the research are

- the introduction of dynamic time slot assignment for single-hop WDM cross connect networks,
- the proposal of dual-hop WDM cross connect networks which use cluster shared buffering and switching modules,
- and multichannel optical buffers that drastically decrease the buffering requirements of the network.

The proposed architectures can achieve capacities which may be orders of magnitude greater than that of simple passive star coupled architectures.

Specific accomplishments which were made include proposals of WDM cross connect architectures as a specific and stand-alone category of networks along with medium access control (MAC) protocols for different classes of these networks. The stations are first clustered into local optical networks (LONs). This enables us to partition the control of the system into local (intra-LON) and global (inter-LON) parts. It is shown that with a very modest number of wavelength channels, very high network capacities are achievable. Heuristic methods are considered for dynamic bandwidth allocation in single-hop WDM cross connect networks based on the traffic demands from the stations. The proposed dual-hop networks on the other hand, include a buffering function to reduce the complexities in network control and MAC while maintaining the very high capacity of WDM cross connect networks. At the same time, the shared buffering/switching modules not only decouple the local (intra-LON) control from inter-LON control protocols, but also eliminates the need for station receiver agility. Use of multichannel optical buffers moves the optical networks one step further to approach all-optical networks, and meanwhile achieves a drastic decrease in the buffering requirements. An extremely simple buffering control protocol is proposed that preserves data packet ordering and avoids any potential processing bottleneck in system operation. Network simulators were designed and developed as a mechanism to experiment with the dynamics of the network, and mathematical models were developed as a validation of the simulator operation and to investigate different aspects of the system operation.

1.5 Organization of Thesis

What is presented in this thesis is a concise explanation of the results generated over four years of research in the area of WDM cross connect networks. Some general concepts are combined with more in-depth description so that a casual reader can get an overview of the proposed networks by skipping some of the technical details, while the more interested reader will find in-depth coverage by referring to specific sections dealing with simulation and analytical models. The remainder of the thesis is organized as follows.

- In Chapter 2 an introduction to computer communications networks is presented. The main concepts of networking and communication media are described. It should be noted that we point out some highlights and we do not attempt to describe all the terms in detail, even though the interested reader is provided with several references for more details.

- Multichannel optical networks are discussed in Chapter 3. Since this thesis deals with architecture and protocol designs for optical networks, a brief literature review in this regard can be of great help in understanding the remaining chapters.

- Chapter 4 introduces single-hop WDM cross connect networks and general concepts related to WDM cross connect networks. Medium access control protocols and dynamic time slot assignments for the single-hop cross connect networks are proposed and their performance is investigated. The problem of dynamic time scheduling is investigated in detail and different heuristic methods are presented.

- Chapter 5 covers the proposed dual-hop WDM cross connect architecture and its medium access control protocol. Dynamic medium access control protocols and scheduling methods are also presented for the dual-hop case. Architectural variations and protocol performance are discussed. Different electronic buffering strategies are proposed and multichannel optical buffers are introduced as an approach to all-optical networks. Capacity, packet loss performance and delay models are described and the analytic results are compared with the numerical results from simulations. The performance graphs given in this chapter highlight some of the advantages that WDM cross connect networks can achieve over other architectures proposed in the literature.
- Chapter 6 provides conclusions and summarizes some of the advantages of the proposed designs. Some related research ideas are proposed for future work.

Chapter 2

Computer Communications Networks

Telecommunication technology and communications networks are evolving so rapidly that writing a text to cover the field is an extremely challenging task. These networks currently surround us and they have found applications in all areas of modern society. In this chapter we will present a brief introduction to networking along with references as preparation for the discussion in later chapters.

2.1 Historical Overview

The use of communication links to connect central computers to remote job entry points (RJEs) was a major development in the early 1950s. In the 1960s remote multiplexers and concentrators were developed to collect data from a set of peripherals and send them to the central processor through a common link. With the increasing power and speed of the processors, special processors called *front ends* were used to free the central computer from handling all the communications and control to and from all the peripherals. Although this system could be called a data network, many of the problems associated with data networks do not arise in these systems, so many view them as a computer system with remote peripherals.

In 1970 the ARPANET was introduced. It connected heterogeneous machines at universities and military installations using a new technique known as "packet switching" [refer to Section 2.2.3]. A wide variety of new issues came out of this project, such as layered protocols, mesh networks, flow control, fault-tolerant performance and so on. One of the branches resulting from the ARPANET split in 1983. That was still called ARPANET versus the military network referred to as MILNET, and consisted of approximately 50 nodes across the United States and Western Europe. The ARPANET is now being phased out, with hosts on the network being moved into other domains connected by the DARPA Internet [SHP91].

The major driving force in the rapid advances in computers, communications, and data networks has been solid-state technology and in particular the development of very large scale integration (VLSI). This enabled faster, less expensive processors; faster, larger, less expensive primary memory and bulk storage. This in turn, led to a rapid increase in the number of cost effective applications for computers. On the other hand, the development of more powerful microprocessor chips enabled personal computers that have speed, instruction set, and memory capacities comparable to the most powerful minicomputers of a few years ago. The increasing use of small, single-function systems, such as word-processors and small business computers, and of general purpose microcomputers, such as personal computers and multiuser workstations led to an increased number of systems at a single site. As a result, there was a need to interconnect these systems, to share and exchange data between different systems and to share expensive resources. These were the compelling reasons for the interconnection of computer systems.

In short, a communication network is a system used to connect a set of stations or nodes together and provide them with a mechanism to exchange data [SHP91, BG92, Sta93].

2.1.1 Communication Technology

Communication links usually leased from the voice telephone companies, are the major links used for long-haul communications. There are a number of fixed rates offered. Some of the rates in the newer networks are 64 kbits per second, 1.5 megabits per second and even 45 megabits per second. There are major economies of scale associated with higher link speeds; for example the cost of 1.5 megabits per second is about six times that of a 64 kbits per second, but the data rate is 24 times higher. Therefore, it is economically better to concentrate network traffic on a relatively small set of high-speed links [BG92].

When a user shares one or more high-speed links with other users, he can achieve low cost per unit data over the bulk of the path, and use low-speed links which have high cost per unit data only for local access to the high-speed links. The communication cost estimate is a very complex task, thus we will not pursue this in detail.

For wide area networks, the dominant part of the communication costs is the transmission cost. Thus, it is desirable to use the communication links efficiently. Packet switching networks can be of great help in the efficient use of communication links. Furthermore, optical fiber communication technology resulted in a drastic drop in transmission costs. Capacities on the order of 1 to 10 gigabits per second have been realized and in the future this could rise to Terabits per second rates. Optical fiber is becoming widespread and is expected to be the dominant mode of the transmission in the future [SHP91, BG92].

Transmission cost for LANs however, has not been the dominant factor for traditional network applications. Coaxial cables and twisted pair wires can achieve relatively high-speed communications at modest cost in a small geographic area. However, they are not sufficient to provide the bandwidth for recent applications.

2.1.2 Applications of Data Networks

The applications for local area and metropolitan area networks cover a broad range. We discuss a few below just to give a flavor of the types of traffic carried.

Personal Computer Networks: Today's personal computers put processor, file storage, high-level languages, and problem-solving tools in an inexpensive, "user-friendly" package. Having a number of such stand-alone systems will not meet all of an organization's needs; central processing facilities are often still required. There are some programs that are too big to run on a small computer, and there are some corporate-wide data files, such as accounting and payroll that require a centralized facility but should be accessible to a number of users. The need for sharing data and resources is obvious for those people working in the same group on the same project. We should not forget that certain expensive resources, such as disks or printers, can be shared effectively through a local area network.

Office Automation: As the number of staff increases, the information and paperwork volume grows. In many businesses, secretarial and other support functions are heavily labor intensive. At the same time, skilled "information workers", managers, are faced with their own productivity bind. Work needs to be done more quickly with less wait time and waste time between parts of a job. This requires better access to information and better communication and coordination with others.

Factory Local Networks: Automated equipment such as programmable controllers, automated material handling devices, time and attendance stations, machine vision devices, and various forms of robots, are dominating the factory environment. The more a factory is automated, the greater is the need for communications. With proper use of the information, it is possible to improve the manufacturing process.

There are many other applications, such as computerized medical diagnosis, remote computer-aided education, airline reservation systems, automatic teller machines, inventory control systems, weather tracking system, electronic mail, and distributed computing.

2.2 Network Classifications

Communication networks are classified in many ways, including network span, topology, and mode of communication between stations.

2.2.1 Network Span

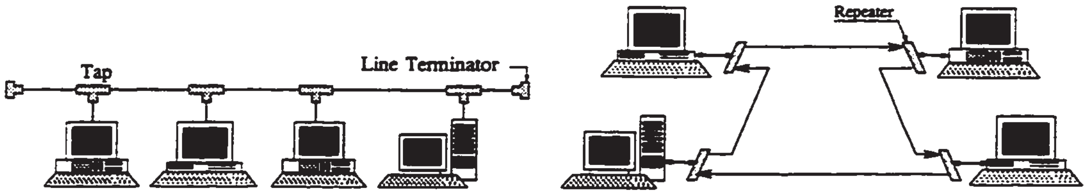
Networks are categorized by geographical coverage as local area networks (LANs), spanning areas with diameters not more than a few kilometers, metropolitan area networks (MANs), spanning areas with diameters up to a few dozen kilometers, and wide area networks (WANs), country-wide and some with world-wide extent [Tan89, BG92, SHP91, Sta93].

2.2.2 Network Topology

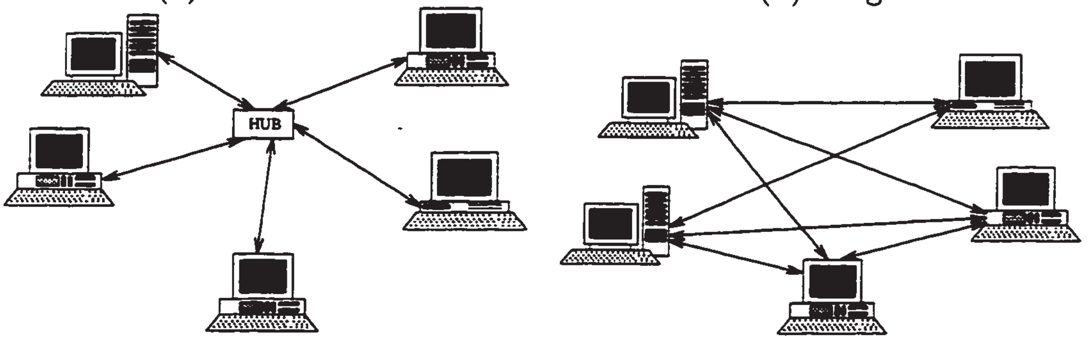
Some common network topologies are shown in Figure 2.1. A passive bus configuration is shown in Figure 2.1(a), where each station (or network node) is connected to a shared medium over which it can both transmit and receive data. Data travels from a given source directly to its destination without being forwarded by any intermediate node, since each station can receive transmissions of all other stations through the shared link. It is apparent that when multiple transmitters access the common channel, collisions can occur when two or more stations transmissions overlap. A medium access control (MAC) protocol is needed to resolve this sort of contention. The ring topology is shown in Figure 2.1(b). The ring consists of unidirectional links where each station receives data from the previous station on the ring and sends data to the next station on the ring. Unlike the passive bus network, data travels around the ring in a multihop manner, since each station is actively coupled to the ring. Information is relayed from one station to the next along the ring.

Figure 2.1(c) shows a star topology. This type of network can be implemented as an active star configuration or a passive one. In the active star case, a special node with multiple transmitters and receivers acts as a hub station. Each station sends data to another station through the central node, or hub. Thus, the active hub can control and resolve contentions in the network by coordinating transmissions. The passive star, however, just distributes the incoming data from any port onto all the output ports. In this case collisions can occur and a MAC protocol is needed to resolve contention.

Finally, Figure 2.1(d) illustrates a general mesh connected network. Each station is equipped with multiple transmitters and receivers that connect it to the other stations through dedicated links. It can be seen that for full pair-wise connectivity, $N(N-1)/2$ links are needed, where N is the number of stations in the network [Sta93].



(a) Passive Bus (b) Ring



(c) Star (d) Mesh

Figure 2.1: Network topologies.

2.2.3 Circuit Switching and Packet Switching

Switching refers to the technique that stations use to share their transmission links. Two common categories exist: circuit switched and packet switched communication. In circuit switching, a circuit, i.e. a dedicated path, is set up between source and destination prior to data communication. This is called the "connection" or "circuit set-up" phase. The circuit or path is released upon completion of the data exchange. Circuit switching is commonly used in telephone networks. This mode of communication guarantees the assigned bandwidth and low fixed transmission delay once the circuit is set up. However, this method is unable to utilize bandwidth on idle circuits. The large overhead due to circuit set up and release phases is another disadvantage.

On the other hand, packet switching is based on sending data from source to destination in self contained units called *packets*. Large messages are broken up into multiple packets and each is sent through the network independently. A packet may be relayed by a number of stations in its path before it arrives at its final destination. This mode of operation is called "store-and-forward". We can easily see that although this method is very efficient in utilizing the communication links effectively, and it does not suffer from circuit set up and release overhead, it is more difficult to guarantee a fixed bandwidth assignment and delay. It is also possible that packets may be received out of sequence. As a result, this type of communication is not ideal for delay sensitive applications, such as real time data processing. Nevertheless, this mode of operation is the most attractive one for computer communications where effective utilization of the communication link is required, and bursty non-delay-sensitive data packets are exchanged. Problems due to increased station complexity due to packet routing protocols, unstable congestion due to bursty traffic, lost or disordered packets, and large variance in packet transmission delay have opened new fields of research to operate packet switching better and more efficiently.

2.3 Layering

A *module* is a device or process within some computer system which performs a given function in support of the overall function of the system [BG92]. Such a function is often called a *service* provided by the module. Although the designers of a module should be intensively aware of the internal details and operation of that module, a user would treat the module as a "black box", and he is concerned only with the inputs, the outputs, and the functional relation of outputs to inputs. Such a module can be used with others to construct a more complex system. This hierarchical design forms a *layered* architecture, where modules appears as black boxes at one layer of the hierarchy, but appears as a system of lower-layer modules at the higher layer of the hierarchy. As an example a computer system could be viewed as a set of processor modules, a set of memory modules, and a bus module. A processor module could, in turn, be viewed as a control unit, an arithmetic unit, an instruction unit, and an input-output unit. Similarly, each of these units could be broken into smaller modules.

The layering or modular design provides simplicity of design; standardization, interchangeability, and widely available modules. All these advantages motivate the idea of a layered architecture in data networks. Layered network architecture is a hierarchy of nested modules, where each given layer regards the next lower layer as one or more modules which provide a set of services. Physically separate modules at a given layer are called *peer processes* or *peer modules.*

In short, the basic idea of a layered architecture is to divide the architecture into small pieces. Each layer adds to services provided by lower layers in such a manner that the highest layer is provided a full set of services to manage communication and run distributed applications. A basic principle is to ensure independence of layers to facilitate interoperability.

2.3.1 OSI Reference Model

The Open System Interconnection (OSI) model developed by the International Standard Organization (ISO) provides a common framework for development of network architectures and protocols. It was first proposed in 1978 and finally adopted by the ISO in 1983 [SHP91, Tan89, Sun89, BG92, Sta93].

In the following we describe some of the basic functions of the OSI reference model. The extended OSI model includes logical link control (LLC) and media access control (MAC) sub-layers for the data link layer to support LANs and MANs as proposed by the institute of electrical and electronics engineers (IEEE) [Tan89, BG92]. The OSI model has been widely accepted as the definitive model for protocol development. Since the model's levels are independent of one another, and if the interfaces between levels is consistent, it is possible to replace any layer with newer versions. Also, in 1988 the U.S. government stated that after 1990 no government office will buy networks that does not conform to the OSI model [SHP91].

The principal functions of the layers of the OSI reference model are shown in Table 2.1.

Despite the momentum of the model, limitations of the OSI reference model are becoming obvious. The model is complex with an excessive number of options. The layering structure is arbitrary, in spite of the attempts of its developers to create a logical layering structure. Appropriate placement of the layers is not always obvious. The most serious weakness of the model results from the fact that its development has been dominated by experts in the communications field rather than in the computer

7	Application Layer	Contains a variety of protocols which are commonly used such as terminal translations and file naming conventions. Electronic mail, remote job entry, directory lookup, and various other general-purpose facilities are also implemented here.
6	Presentation Layer	Concerned with the syntax and semantics of the information transmitted. May provide encryption, compression and/or reformatting.
5	Session Layer	For tying terminals to applications during a session; establishes, manages and terminates connections (sessions).
4	Transport Layer	Provides transparent transfer of data between systems. Provides an error-free point-to-point channel that delivers messages in the order in which they were sent. Includes end to end synchronization, flow control and error recovery.
3	Network Layer	Provides independence from data transfer technology and relaying and routing considerations. Issues include how packets are routed, congestion control, addressing and packetizing.
2	Data Link Layer	Provides functional and procedural means to transfer data between network entities and correct transmission errors. Controls the transmission and reception of blocks with necessary synchronization, error and flow control, and media access control.
1	Physical Layer	Provides electrical (optical), mechanical, functional, and procedural characteristics to activate, maintain and deactivate physical links that transparently pass the bit stream for communication.

Table 2.1: The main functions of the layers of the OSI reference model.

field, hence it is more appropriate for connection-oriented rather than connectionless services [SHP91].

2.4 Transmission Media

The transmission medium is the physical path between transmitter and receiver in a communication network. Transmission media may be classified as guided or unguided. Electromagnetic waves are guided along the physical path in the first case, such as in twisted pair wires, coaxial cable, and optical fibers. In the second case, the medium provides a means for transmitting the electromagnetic waves but do not guide them, such as the atmosphere and outer space.

The choice of the transmission medium is determined by several factors such as, network topology, capacity, reliability, types of data supported, environmental scope, and cost [Sta93, BG92].

2.4.1 Twisted Pair

Twisted pair wires have been the most commonly used transmission medium. Almost all wiring within telephone networks are twisted pair. It consists of two insulated wires in a regular spiral pattern. The wires are usually copper or steel coated with copper. The twisting of each pair minimizes electromagnetic interference between the pair. They may be used to transmit both analog and digital signals. For analog transmission, amplifiers are required about every 5 to 6 km. For digital signals, repeaters are used every 2 or 3 km. Twisted pair has a capacity of up to 24 voice channels using a bandwidth of up to 268 kHz. Digital data can be sent on a wire pair using baseband signaling or using a modem. A data rate of 4 Mbps represents a reasonable upper limit [Sta93]. Twisted pair is limited in distance, bandwidth, and data rate compared to other guided media. Also, the medium is quite susceptible to interference and noise. It is, however, less expensive than either coaxial cable or optical fiber in terms of cost per meter. Because of its connectivity limitations, installation costs may approach that of other media.

2.4.2 Coaxial Cable

Two types of coaxial cables have been widely used, 75-ohm cable, mainly used in community antenna television (CATV), and 50-ohm cable. The latter is typically used for digital signaling, called ***baseband***, while 75-ohm cable is mainly used for analog signaling with frequency division multiplexing (FDM), called ***broadband***, and for high-speed digital signaling and analog signaling without FDM. This is sometimes referred to as ***single-channel broadband***.

Coaxial cable is known as the most versatile transmission medium. It consists of two conductors, where one conductor forms the outer cylindrical shield that surrounds a single inner wire. Some insulating material fills the space between these two conductors and holds them in their place. The diameter of a single coaxial cable is from 0.4 to about 1 inch.

For analog signaling, frequencies up to 400 MHz are possible. For digital signaling, data rates up to 50 Mbps have been achieved [Sta93, SHP91]. The maximum distance in a typical baseband cable is limited to a few kilometers. Broadband networks can span ranges of tens of kilometers. In general, the noise immunity of coaxial cable is superior to that of twisted pair for high frequencies. The cost of installation of coaxial cables falls between that of twisted pair and optical fiber.

2.4.3 Optical Fiber

The development of practical optical fiber communications systems has been a very significant technological breakthrough in data transmission. Fiber has already been widely used in long-haul communications. Technological improvements, together with the inherent advantages of optical fiber, have made it increasingly attractive for local area networking. The tremendous potential bandwidth (around 25 THz), smaller size and lighter weight, significantly lower attenuation (around 0.2 dB/km), and electromagnetic isolation properties of the optical fiber, particularly of single-mode optical fiber, distinguish it from twisted pair or coaxial cable.

An optical fiber is a very thin (2 to 125 μm), flexible medium capable of conducting optical energy. It consists of three concentric cylindrical sections: the core, the cladding, and the jacket. The core, made of glass or plastic, is the innermost section which carries the optical energy. The cladding surrounds the core. It is made of glass or plastic that has optical properties different from the core. The outermost layer, surrounding one or a bundle of cladded fibers, is the jacket. It is composed of plastic or other materials and provides the protection against moisture, crushing, and other environmental dangers.

Light travels along the fiber by means of total internal reflections which happen in any transparent medium that has a higher index of refraction than the surrounding

	Stepped Index Multimode	**Graded Index Multimode**	**Single Mode**
Light Source:	LED or laser	LED or laser	laser
Bandwidth:	wide 200 MHz/km	very wide 200 MHz/km to 3 GHz/km	extremely wide 3 GHz/km to 50 GHz/km
Splicing:	difficult	difficult	difficult
Typical Application:	Computer data link	Moderate length telephone lines	Telecommunication long lines
Cost:	least expensive	more expensive	most expensive
Core Diameter:	50-125 μm	50-125 μm	5-8 μm
Cladding Diameter:	125-440 μm	125-440 μm	15-60 μm

Table 2.2: Comparison of three basic types of optical fibers.

medium, provided that the incident angle is sufficiently high. Optical fibers, form a waveguide for frequencies in the range of 10^{14} to 10^{15} Hz; this mainly covers the visible spectrum and part of the infrared spectrum.

Depending on the number of propagation modes, optical fibers are classified into multimode and single mode fibers. Light from a source enters the cylindrical core and propagates along the fiber as it is reflected. Different angles of reflection result in different modes of propagation. If we reduce the radius of the core to the order of a wavelength, then only a single angle or mode can pass which is the axial ray. In multimode fibers, different modes travel different path lengths and hence take different times to traverse the fiber. This causes signal elements to spread out in time and limits the rate of accurate data transmission. In single mode fiber this phenomenon does not exist and this provides superior performance to multimode fiber. Finally, there exists another type of multimode fiber that is intermediate between single mode and multimode fiber. This type is called graded index multimode fiber, because it has a varying index of refraction in the core. Only a brief discussion and narrative explanations are given in here. Interested readers can find more details in [Sen92, Jon88, Gre93, ST91]. Figure 2.2 illustrates the three types of the fibers and their characteristics, while Table 2.2 compares some of their properties [Sta93].

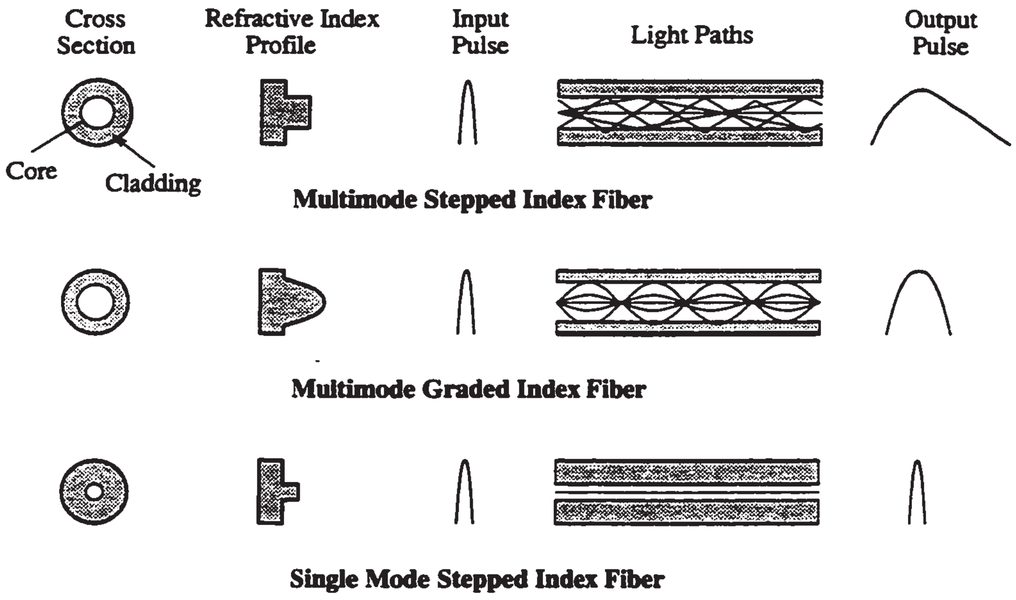


Figure 2.2: Characteristics of three basic types of optical fibers.

2.4.4 Line-of-Sight Media

Microwave, infrared, and laser are three ways in which electromagnetic waves may be transmitted through a line-of-sight medium, i.e., the atmosphere and outer space. The frequency range for microwave is 10^9-10^{10} Hz, 10^{11}-10^{14} Hz for infrared, and 10^{14}-10^{15} Hz for laser [Sta93].

Line-of-sight media can help overcome some of the difficulties in connecting local or metropolitan networks, especially where some buildings or sites cannot be easily connected physically. In line-of-sight communication we require equipment only at each building or site. Installation of such systems is very fast and simple. The only concern is that the transmitter and receiver have to be within line-of-sight of each other.

Among the three techniques, infrared and laser are more susceptible to environmental interference such as rain and fog. Microwave is less directional than the other two methods, and needs to meet frequency licensing regulations. Interested readers can find more detail in [Sta93] and some free space interconnection techniques can be found in [Mid93a].

2.5 Multiaccess Communication

A number of stations sharing a common communication channel form a multiaccess communication network. There are many examples of such systems, satellite systems, radio broadcast, multidrop telephone lines, multitap bus systems, etc. The received signal in each station is a sum of attenuated signals transmitted by a set of other nodes.

In packet switched networks, data packets often include the source station address and destination station address. In multiaccess broadcast-and-select systems the desired destination station can examine the address field and keep those intended for itself and ignore the rest. In multiaccess systems, transmission conflicts may occur when two or more user stations want to transmit simultaneously. Thus, an arbitration protocol is needed to resolve the contention. These protocols can be classified into centralized control protocols and distributed control protocols. These protocols can also subdivided into random access, or "free-for-all", and reservation based, or "scheduled". The medium access control (MAC) sub-layer of the data link layer in the OSI reference model deals with these protocols.

In a centralized MAC protocol, a single node in the network arbitrates access to the shared channel. In such a system, all stations may send their bandwidth and delay requirements to the controlling station. The controller would then execute an arbitration algorithm to decide the appropriate allocations. The result would then be sent to the stations. In a system with distributed control, a similar algorithm may be used in each station, and the central station does not exist.

In random access or contention based systems, user stations attempt to acquire the channel with less global coordination than in the centralized case. A data transmission is successful if it does not collide with others transmitted by other stations. In the popular examples of this, such as Aloha, slotted Aloha, carrier sensing multiple access (CSMA), collided transmissions have to be attempted again.

In a conflict-free protocol, transmissions are coordinated by a common algorithm so that collisions will not occur. Reservation based protocols allocate the time slots to different stations according to their demands, dynamically, or in a fixed manner.

Time division multiple access (TDMA) is an example of a static scheduling protocol, while a token passing ring is a dynamic version.

In the following sections some details are presented for random access and reservation based methods. The interested reader is referred to [SHP91, Tan89, BG92, Gre93, Sun89].

2.5.1 Random Access Methods

One of the main characteristics of computer communication is that data traffic is often bursty. Burstiness is a result of the high degree of randomness seen in the generation time and size of messages. Computer network users often require communication infrequently, but demands a rapid response, resulting in an inherently large peak-to-average ratio in the required data transmission rate. Therefore, when transmission bandwidth is allocated in a fixed fashion, a very low channel utilization may result. Consequently, random access methods are often more efficient for bursty packet switched environments. When dealing with multiaccess communication, we must be prepared to resolve some conflicts, which arise when more than one station transmits at the same time onto the shared channel. Random access methods or *collision-based* protocols deal with the problem of how to control the access to the common channel and resolve contention in simple hardware and software implementations with an acceptable level of performance [Sta93, Sun89, Tan89, SHP91, BG92]. The most common random methods are described below. This class of protocols can be easily implemented through distributed control, since little coordination is required.

Aloha

Aloha in its simplest form, also called *pure* Aloha, has user stations transmit whenever they desire. If a transmitted data packet is not immediately acknowledged by the destination station, then the transmission is not successful and it must be resent. The failure of any transmission can result of two types of collisions: transmitter collision and receiver collision. When more than one transmitter attempts to send data onto the same channel simultaneously, transmitter collision occurs. The second type of collision occurs when the receiver of the destination station is not listening to the transmitting channel. Receiver collision may happen in multichannel environments where each station is equipped with one or more receivers.

In order to avoid continuously repeated collisions, the time before a station retransmission is randomized across colliding stations. Therefore, when an acknowledgment is not received, the transmitting station retransmits the unsuccessful packet after a random backoff time.

There is another version of Aloha, called *slotted* Aloha. In this case, the time axis is slotted into time slots of duration equal or greater than of the transmission time of a data packet. Each user station must synchronize the start of its transmission to coincide with the slot boundary. This increases the channel efficiency over pure Aloha by decreasing the vulnerable period, which is the time during which another transmitting station would collide with a given transmission.

Pure Aloha is efficient only for light bursty traffic patterns but has several drawbacks for heavier loading. As the loads becomes higher, the number of transmissions and consequently, the number of collisions increases. At some point the number of collisions becomes so high that most of the time is spent to resolve contention and the throughput of the network decreases as the load increases beyond this point.

If a large number of stations are generating data packets under Poisson arrival rates, the throughput for the Aloha protocol can be shown to be

$$S_{Aloha} = Ge^{-2G}, \tag{2.1}$$

where G is the total offered load by all the stations [SHP91, Tan89]. We can easily see that the maximum throughput is about 18 percent of the channel capacity. This occurs when the total offered load (G in Equation 2.1) is 0.5 data units per data time slot.

In the slotted Aloha protocol, the vulnerable period to collisions is half the one for Aloha. This causes an improvement in the throughput for the slotted Aloha over the basic Aloha. With the same assumptions as for the Aloha protocol, the throughput for the slotted Aloha protocol can be written as [SHP91, Tan89, Sun89]

$$S_{Slotted\ Aloha} = Ge^{-G}. \tag{2.2}$$

The maximum throughput in this case is about 36 percent, which is twice that of the Aloha protocol. This happens when the total offered load (G in Equation 2.2) is 1 data unit per data time slot.

Carrier Sensing

If the propagation delay across the network is small, transmitting stations may attempt to avoid collisions by sensing the channel to ensure that it is not in use by other stations. Random access methods that implement this technique are called carrier sense multiple access (CSMA). It is apparent however, that due to non-zero propagation delay, collisions still may occur and collided packets must still be resent. Two main version of CSMA protocols are popular: *non-persistent* and *persistent.*

In non-persistent CSMA, a station with a packet to send first senses the channel and operates as follows.

1. If the channel is idle, transmit.
2. If the channel is busy, backoff and reschedule the transmission for a later time.

A variation of this is the p-persistent case as follows.

1. If the channel is idle, transmit with probability p, and delay one time unit (usually equal to twice the maximum propagation delay) with probability $(1-p)$.
2. If the channel is busy, the station persists in sensing the channel until it is found to be idle and then proceeds from the beginning.

The parameter p can be any number between 0 and 1. The most common choice is 1-persistent. Contrary to the non-persistent case where stations are indifferent, 1-persistent stations are selfish. If more than one station are waiting to transmit, a collision is guaranteed.

An extension to CSMA is CSMA/CD, which stands for carrier sense multiple access with collision detection. Even though CSMA is more efficient than Aloha, it has one important inefficiency. When collisions happen, the transmission channel is wasted for the duration of the transmitted packets. If a station continues to listen to the channel while it is transmitting, this waste can be reduced as follows.

1. The station continues to listen to the channel while it is transmitting. If a collision is detected, it stops transmitting immediately and transmits a brief jamming signal to ensure that all stations know that there has been a collision.
2. The station then waits for a random backoff time, then attempt to transmit again using CSMA.

There is still another question that we need to answer: how long does it take to detect a collision? The question can be rephrased as: how long must a station listen to the channel to make sure that the transmission was successful? In order to answer this question we consider an example shown in Figure 2.3. The passive bus topology is used in this example. The worst case of two stations at either end of the bus is considered. It can be easily seen that the amount of time it takes to detect a collision is twice the end-to-end propagation delay (τ).

An important factor mentioned above is the "backoff" time. It is obvious that collisions generate additional traffic in the system. As the medium becomes busier, it is important not to clog the network with retransmissions, which in turn lead to more collisions. Therefore, it is appropriate that when a station experiences repeated collisions, it backs off for longer periods of time to compensate for the extra load on the network. The generally used rule for this purpose is known as "truncated binary exponential backoff" [Wal91]. This technique is as follows.

- The backoff delay is an integral number of slot times. Each slot time is 512 bit times.
- The number of slot times to back off before the n^{th} retransmission attempt is a random uniformly distributed integer r in the range $0 \leq r < 2^k$, where $k = \min(n, 10)$.

- If the number of unsuccessful attempts reaches a user-defined number (usually 16) an error is signaled.

If we assume a 1-persistent CSMA/CD protocol with binary exponential backoff in a network with a large number of backlogged stations, a fair approximation of the protocol capacity is found as [Wal91]

$$S_{CSMA/CD} \approx \frac{1}{1+2ea} \approx \frac{1}{1+5.4a} , \tag{2.3}$$

where $a = \frac{\tau}{T}$ is the normalized end-to-end propagation delay. τ is the one-way propagation delay and T denotes the transmission time of a data packet (message frame length).

From Equation 2.3, we can easily find that the protocol capacity for CSMA/CD is very close to unity for small values of a and it degrades as a increases. As seen, CSMA and CSMA/CD are very sensitive to propagation delays. If the propagation delays are long compared to the message frame length, the performance of the protocol suffers a lot [SHP91, Sta93, Sun89].

CSMA/CD, named also as Ethernet, is frequently used in LANs and is one of the standards chosen by the IEEE 802 committee. The IEEE 802.3 standard gives precise details of the protocol.

2.5.2 Reservation Based Methods

The first classification that we consider in this case is the *fixed assignment* or *fully scheduled* technique; the second is *on-demand assignment.* The on-demand assignment class can be further categorized as *central control* and *distributed control.* Some representatives of reservation based protocols are discussed in the following sections.

Fully Scheduled

One type of fully scheduled MAC protocol is time division multiple access (TDMA). The time is divided into frames and each in turn is partitioned into time slots. Each time slot is assigned in a static predetermined fashion to each station. Each station

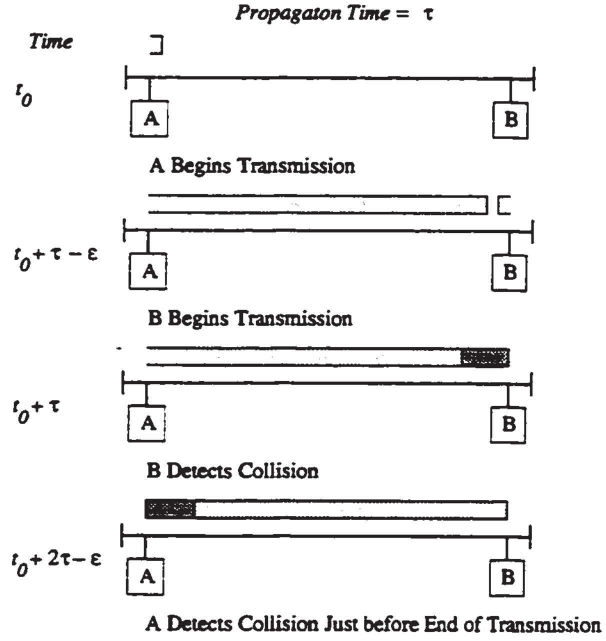


Figure 2.3: Collision detection in baseband passive bus.

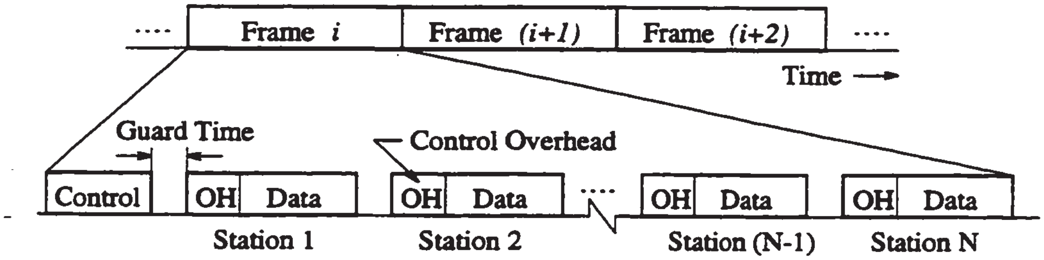


Figure 2.4: TDMA frame structure for N stations in the network.

can send one packet of data as its turn arrives in each frame. Figure 2.4 shows a sample frame structure for TDMA operation.

A simpler frame structure can be used in LANs using such media as coaxial cable. In this case, the guard bands and control overhead might not be necessary for every slot. The frame shown in Figure 2.4 is mainly used on a satellite channel.

The fact that each station is assigned a time slot in each frame is a disadvantage for TDMA in bursty traffic environments. If the stations do not have data ready in a uniform continuous pattern, the channel bandwidth is wasted. However, there exist other environments where all stations have consistent traffic patterns over long time periods. TDMA is an excellent MAC protocol for such cases. Another drawback of TDMA is that all the stations must be synchronized and access the common channel on the predetermined time boundaries. This necessitates additional hardware and allocation of a portion of the bandwidth for synchronization, such as the control overhead and guard time shown in Figure 2.4.

Central Control

A central controller is a specific station in the network. All the communications in the network must be done according to scheduling performed by the central controller. The controller can apply a wide range of static or dynamic methods for bandwidth assignment. One example of this kind is *polling.*

In polling systems the central node checks all the stations one by one to see if they have any data packets ready to send. It starts by sending a polling message to the first station. After receiving the polling message, this station can transmit followed by a *go-ahead message* to denote that the transmission is completed. The central controller then checks the second station and repeats the same procedure, and so on. In more sophisticated methods the bandwidth assignment is done based on the traffic demands from all stations or according to the statistics of the traffic flow.

Distributed Control

In this kind of access control, a central controller does not exist. A common algorithm is executed in all the nodes of the network to determine the access schedule. One of the examples for distributed control is *token passing* method. In typical token passing protocols, all the packets make a complete circuit of the network and return to the originating station to be removed.

Two versions of token passing protocols, token ring and token bus, have been established as standards. IEEE 802.4 and IEEE 802.5 are the standards for token bus and token ring, respectively. The token passing protocol has also been used in another newer standard which is referred to as the fiber distributed data interface (FDDI). Higher data rates than those of token ring and token bus are considered in

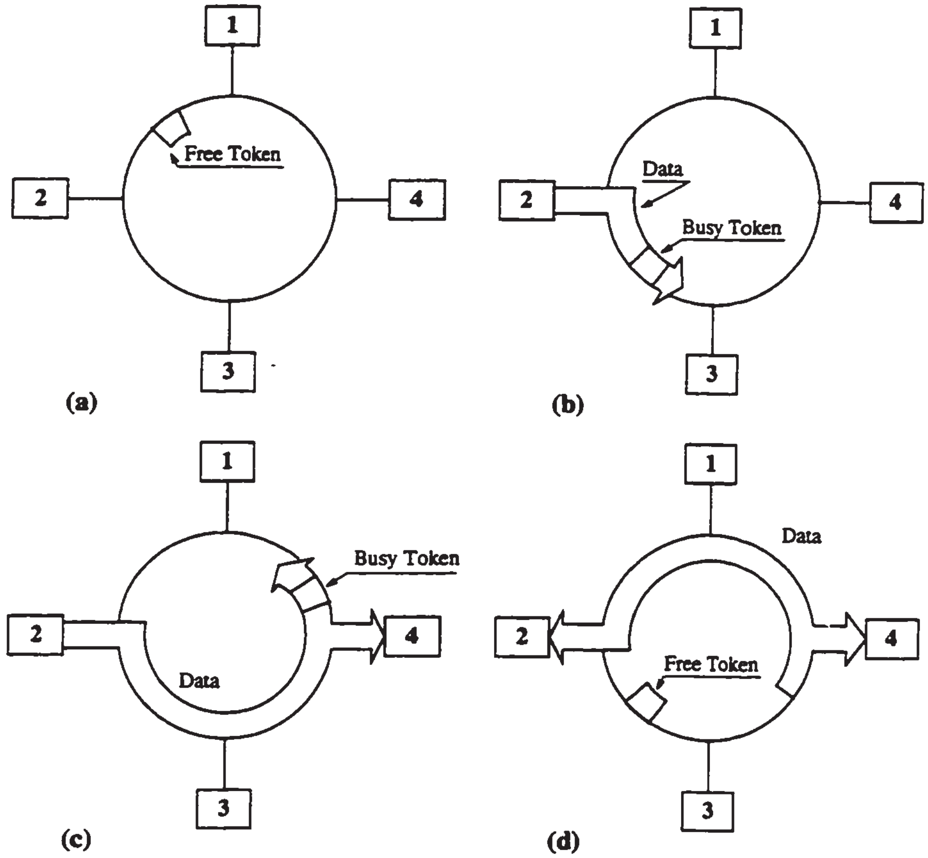


Figure 2.5: Token Ring: Station 2 has data ready to transmit to station 4. (a) Sender looks for free token. (b) Sender changes free token to busy token and appends data. (c) Receiver copies data addressed to it. (d) Sender generates free token upon receipt of its transmitted packet.

this case [Sta93, BG92].

Token ring uses a physical ring topology as shown in Figure 2.1(b). An active interface couples the stations to the ring, so that each station relays the data it receives back onto the ring. This method prevents bit deterioration over large network spans. A short delay is also needed in each interface to allow time for a station to obtain information from an incoming packet and decide whether it needs to remove the packet from the ring or just relay it.

A specific packet known as a *token* is used in token passing protocols. The token has two states: "free" and "busy". When there is no data communication on the ring, a free token is circulating around the ring from one station to another. If a station has data to transmit, it captures the free token, changes it to busy token, and sends it back to the ring followed by its data packets. The transmitting station inserts a new free token when it has completed transmission of its frame and the leading edge of its transmitted frame has returned to the station after a complete circulation of the ring. These steps are shown in Figure 2.5, where station 2 is transmitting data to station 4. This mode of token passing operation is called *transmitter token release*, and only one free token exists in the above scheme. Other variations of the token passing protocols such as *multiple token operation*, and *receiver token release* are used in other token passing systems [SHP91, Wal91, Sta93, Sun89].

Chapter 3

Multichannel Optical Networks

3.1 Overview

This thesis deals with multichannel optical networks. This chapter includes a discussion of their characteristics, the devices used in these networks as building blocks, and their limitations.

As mentioned in Section 2.4.3, optical fiber, as a communication medium, offers a huge bandwidth (around 25 THz) in the low attenuation frequency region (carrier wavelengths of 1.3μm and 1.5μm). Since optical fiber is a very low loss medium and its use enables larger repeater spacings in long-haul links, it has been used as a replacement for copper wires in traditional networks. Optical Fiber has also been used in some newer LAN and MAN designs, such as the fiber distributed data interface (FDDI) and the distributed queue dual bus (DQDB). Employing optical fiber in these networks results in higher data rates, lower error rates, and less electromagnetic interference. However, improvements in performance are limited due to the electronic interfaces at the network nodes. In these networks the potential benefits of the fiber medium are not fully exploited, since their single channel nature is maintained and stations are not capable of communicating at the full bandwidth of the optical fiber. The main design objective in these networks is to increase the distance-bandwidth product of the point-to-point link, so that a number of stations can share the bandwidth of a single channel. A new generation of networks has been proposed in which the optical fiber is used because of its unique properties. Totally new approaches are needed in this generation in order to meet the needs of emerging high-bandwidth applications [Gre93]. In these networks, design emphasis has shifted to the entire network architecture, fast multichannel multiaccess communication protocols, station to station connectivity, and dynamic bandwidth utilization.

The new generation of the networks is "all-optical" in nature, in the sense that once information enters the network, it may remain in the optical domain until it is delivered to its destination. Any optical to electronic conversion is a potential bottleneck in the communication path, since the maximum electronic processing speed is four orders of magnitude less than the potential capacity of the optical fiber [Muk92b]. Some of the design targets in these new networks are simple physical network structure, minimal optical-electronics translations, minimal and fast processing of the transmitted information within the network, and simple medium access control (MAC) protocols.

One of the main strategies to exploit the available bandwidth of the optical fiber has been the multichannel design, where concurrent parallel channels, each operating at the rates available for electronic processing, are multiplexed onto a high rate (high bandwidth) channel. As we will discuss later, wavelength division multiple access is the most promising multiplexing method in this regard.

New emerging applications which demand high bandwidth, drive the research for the new generation of optical networks. These applications demand more bandwidth and data rates than that which traditional networks can offer. Examples of such applications are: "Fiber to the office" and "fiber to the home" to provide multiple connections of high definition television (HDTV), digital audio, networks in universities and research environments that interconnect hundreds or thousands of high performance workstations, medical imaging, distributed software design, more graphics oriented software, and supercomputer interconnections.

In this chapter, we present a quick survey of optical fiber communication systems and their characteristics. Some fundamental components and building blocks of optical networks are discussed as a foundation for the later chapters. Multiplexing methods and their principles are presented, followed by a review of important classes of optical networks.

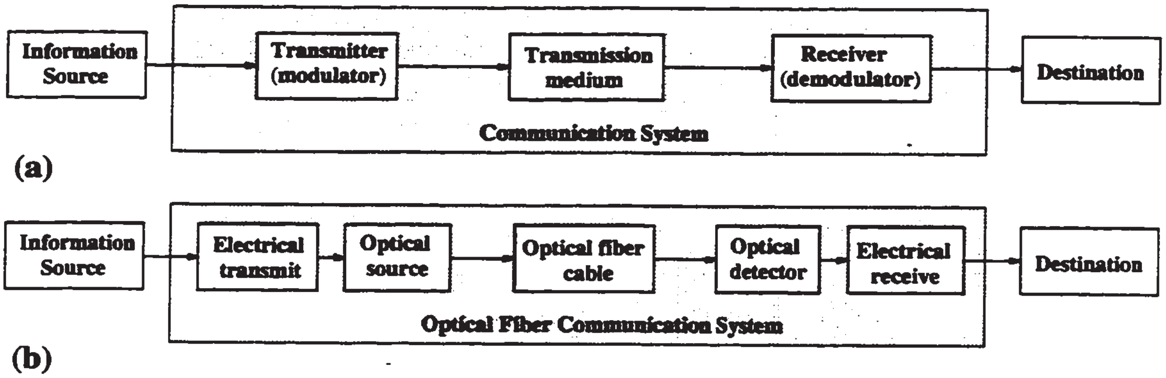


Figure 3.1: (a) The general communication system. (b) The optical fiber communication system.

3.2 Optical Fiber Communication System

The basic concepts in an optical fiber communication system are similar to those in any kind of communication system. The general communication system is shown in Figure 3.1(a). The electrical signal provided by the information source is derived from a message signal (e.g. a sound signal). The transmitter translates (modulates) this signal into a proper form for propagation over the transmission medium. The transmission medium can be a guided medium such as twisted pair, coaxial cable, or optical fiber, or it can be an unguided medium like air or outer space. The transmitted signal is attenuated in any transmission medium, suffers loss, and is subject to degradations due to random signals, noise and distortion. As a result, there is a maximum allowed distance between the transmitting side and the receiving side for any communication system. Beyond this distance the signals at the receiving side are not reliably detectable. The receiver demodulates (detects) the signal and passes it to the destination. The original signal is then extracted from the demodulated electrical signal by the destination. In long haul communications, repeaters or line amplifiers are needed to boost the signal energy and remove signal distortion to allow longer distance communications.

An optical fiber communication system is shown in slightly more detail in Figure 3.1(b). As shown, the electrical stage provides the input signal for the optical source. The optical source modulates the light carrier with the input electrical signal. A light

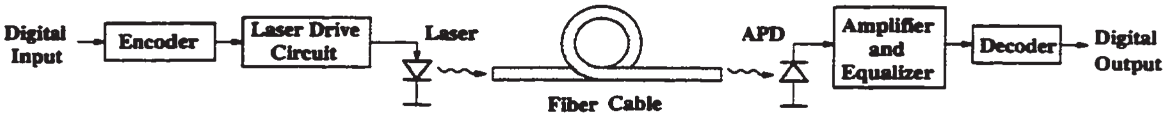


Figure 3.2: A typical digital optical fiber link.

emitting diode (LED) or semiconductor laser can be used as the optical transmitter. The transmission medium is an optical fiber cable which carries the signal to the receiver. The optical detector, usually photodiodes, demodulates the electrical signal from the optical carrier, and provides the input signal for the electrical receiving circuit. The electrical signal usually goes through a signal processing stage and is delivered to the destination, where the original message signal is retrieved from the electrical message signal.

Two methods can be used to modulate the optical carrier, analog or digital. In the analog intensity modulation, the intensity of the light emitted from the optical source varies with the input message signal in continuous fashion. In the digital case, discrete changes in the light intensity are obtained (i.e. on-off pulses). Although the analog technique is much easier to implement, it generally needs much higher signal to noise ratios at the receiver side than that of digital modulation. Besides, semiconductor optical source operation is not linear especially at high resolution frequencies, which is needed for analog modulation. As a consequence, fiber communication links are limited to shorter distances for analog modulation than those for the digital case. The block schematic of a typical digital optical fiber link using a semiconductor laser source and an avalanche photodiode (APD) detector is shown in Figure 3.2. The laser drive circuit directly modulates the intensity of the semiconductor laser with the encoded digital signal. On the receiving side, the APD detects the launched signal. The front-end amplifier and equalizer provide gain and reduce the noise bandwidth. Finally, the decoder extracts the original signal [Sen92].

We will discuss some of the components of the optical fiber communication system later. Let us first consider the advantages of lightwave communications over other techniques, such as radio communications.

3.2.1 Advantages of Optical Fiber Communications

Some of the attractive features and merits of optical fiber communications over more conventional electrical communications are presented here [Sen92]. As optical fiber technology develops, these advantages become more apparent.

- **Enormous potential bandwidth:** Optical carrier frequencies are in the range of 10^{13}-10^{16} Hz. In the near infrared around 10^{14} Hz or 10^5 GHz, optical fiber yields a far greater potential transmission bandwidth than metallic cable systems. The available bandwidth in typical coaxial cable and millimeter radio systems are around 500 MHz and 700 MHz, respectively. In the low attenuation window (carrier wavelengths of 1.3μm and 1.5μm) the available bandwidth in optical fiber is about 25 THz.

- **Small size and weight:** Optical fibers have very small diameters which are often no greater than the diameter of a human hair. Even when such fibers are covered with protective coatings they are far smaller and much lighter than their copper cable counterparts.

- **Electrical isolation:** Optical fibers are electrical insulators and do not exhibit earth loop and interference problems. They are ideal for electrically hazardous environments.

- **Immunity to interference and crosstalk:** Optical fibers form a dielectric waveguide and are therefore free from electromagnetic interference (EMI), radio frequency interference (RFI), and switching transients generating electromagnetic pulses. The operation of an optical system is unaffected by transmission through a electrically noisy environment and the fiber cable requires no shielding from EMI.

- **Signal Security:** The light from optical fibers does not radiate significantly, providing a high degree of signal security. This feature is obviously attractive for military, banking and general data transmission (i.e. computer network) applications.

- **Low Transmission Loss:** Fibers have been fabricated with losses as low as 0.2 dB per kilometer. This facilitates the implementation of communication links with extremely wide repeater spacing.

- **Ruggedness and flexibility:** Although protective coatings are essential, optical fibers may be manufactured with very high tensile strengths.

- **System reliability and ease of maintenance:** The low loss property of optical fiber cables reduces the requirement for intermediate repeaters or line amplifiers. The lifetime of the optical components is around 20 to 30 years [Sen92]. Consequently, they are very reliable. Both of these factors also reduce maintenance time and cost.

- **Potential low cost:** The glass which normally provides the optical transmission medium is made from sand – not a scare source, although the processing cost is still high. This potential has not yet been achieved in all other component areas associated with optical fiber communications. For example, the costs of high performance semiconductor laser and detector photodiodes are still relatively high, as well as some of those concerned with connection technology, such as connectors and couplers, etc.

  Overall system costs of optical fiber communications on long-haul links are essentially less than those for equivalent electrical line links. The lower costs of optical fiber communications has not only provided strong competition for electrical line transmission systems, but also for microwave and millimeter wave radio transmission systems.

3.3 Optical Network Components and Devices

3.3.1 Light Sources

Light emitting diodes (LEDs) and laser diodes are used as sources in optical fiber communication systems. The important characteristics of the light source are as follows.

- *Power:* The transmitting power should be high enough for the signal to be detected at the receiver with the required accuracy.
- *Speed:* It must be possible to modulate the source power at the desired rate.
- *Linewidth:* The spectral linewidth of the source must be narrow to minimize the effect of chromatic dispersion in the fiber.
- *Noise:* Random fluctuations in the source can be treated as noise. It is especially important for coherent communication systems to have a very stable light source.
- *Insensitivity* to environmental changes such as temperature.
- *Reliability.*
- *Low cost.*
- *Long lifetime.*

Laser diodes have high power (tens of mW), high speeds (in the GHz region), and narrow spectral linewidth; but they are sensitive to temperature variations. LEDs, on the other hand, are reliable, have lower cost, long lifetime, and simplicity of design. However, they have broader linewidth (more than 100nm). Their modulation frequency at maximum power is limited to around 100 Mbps. A more complex LED structure called an *edge-emitting diode* can produce more power output with relatively narrower spectral linewidth.

Light sources are available at 0.87 μm, 1.3 μm, and 1.55 μm. AlGaAs LEDs and AlGaAs/GaAs double-heterostructure and quantum-well laser diodes have been used at 0.87 μm. InGaAsP LEDs and InGaAsP/InP double-heterostructure lasers have been used at 1.3 and 1.55 μm. Note that in the case of 1.3 μm, a narrow spectral linewidth is not as crucial as it is for 1.55 μm, since material dispersion is minimal at 1.3 μm.

External-cavity, distributed feedback (DFB) and distributed Bragg-reflector (DBR) lasers are capable of providing spectral linewidths of 5 to 100 MHz at a few mW of

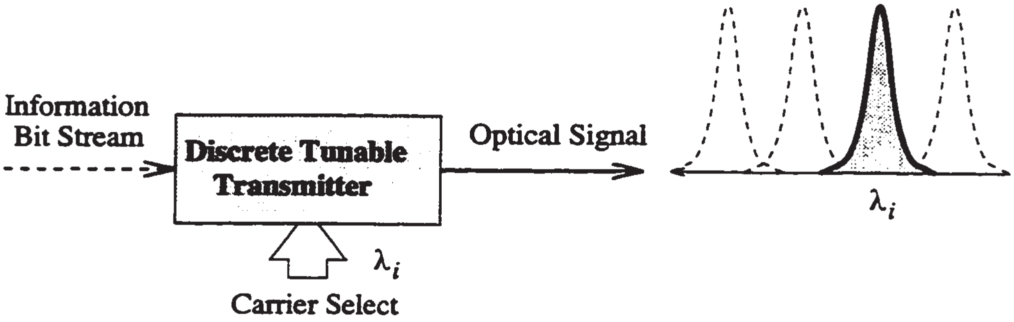


Figure 3.3: Discrete Tunable Light Source.

output power with modulation rates exceeding 20 GHz. Cleaved coupled cavity (C^3) lasers have linewidths as low as 1 MHz, but are subject to thermal drift. The physical aspects and other details can be found in [ST91, Sen92, Jon88, Gre93].

Discrete Tunable Light Source

Wavelength agility of optical sources is a desirable quality for optical communication networks. A discrete tunable source modulates the digital information signal onto one wavelength channel out of a set of available wavelength channels. Figure 3.3 shows a schematic diagram of the tunable source. Two types of electrical input signals exist: the control signal and digital information bit stream. The output is the optical signal containing the information bits modulated by the optical carrier selected by the control signal. As shown in the figure, one channel λ_i is selected from a set of available wavelength channels $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$.

The number of resolvable channels and inter-channel tuning speeds are of great importance. Ideally, a tunable light source has a minimal linewidth and fast tunability over a 200 nm low loss transmission window around either 1.3 μm or 1.55 μm. One option for a tunable light source is laser arrays, which are created by integrating a number of single frequency lasers into a one or two dimensional array on a single substrate and coupling the outputs of all the lasers to a single output fiber [Gre93, Mid93a]. This can provide a very fast tunable optical source, although it is difficult to couple a large number of sources to a single output fiber efficiently. A wideband optical source followed by a tunable optical filter is another method of producing a tunable optical source, which is called spectral splicing [Gre93]. A disadvantage of this scheme is that a large portion of the optical energy would be attenuated due to filtering a small fraction (one channel) of the original spectrum. Mechanical, thermal, or acousto-optic effects are also used to continuously tune the wavelength channel [Bra90]. Thermal tuning of DBR lasers provides a very small tuning range, 1-2 nm, with tuning speeds in the millisecond range. Mechanical tuning techniques, on the other hand, have broader tunability range, 100 nm, with almost the same tuning speeds as thermal tuning. An acoustic wave applied to a DBR laser can change the index of refraction of the material, and consequently, control the lasing wavelength. Acousto-optical tuning range is in the order of 100 nm, but its tuning speed is limited to the microsecond range due to the velocity of acoustic waves. The index of refraction can also be changed by applying an electric field. However, the electro-optical effect is weak in most materials [Gre93]. Therefore, even though the electro-optical tuning method is very fast, in the nanosecond range, the tuning range is very small, about 10 nm.

3.3.2 Light Detectors

There are two common types of light detectors used in optical networks: *p-i-n* photodiodes and avalanche photodiodes (APD). APDs require a high voltage supply and more complicated circuitry to compensate for their sensitivity to temperature fluctuations. The basic operational principles, sensitivity and signal-to-noise ratio (SNR) of receivers using *p-i-n* photodiode and APD are discussed in [Sen92, ST91, Gre93].

Silicon *p-i-n* photodiodes and APDs are available for 0.87 μm wavelength. These APDs have a 10 to 15 dB advantage over *p-i-n* photodiodes. Germanium and InGaAs *p-i-n* photodiodes are used as detectors for 1.3 and 1.55 μm wavelengths. InGaAs has greater thermal stability and lower dark noise. InGaAs APDs operating at speeds of about 2 Gbps are widely available. Some APDs also exist that use a heterostructure of two materials, a small gap material for absorption (such as InGaAs) and a large gap material for multiplication (such as InP). By this method, bandwidth has been extended up to 10 GHz [ST91].

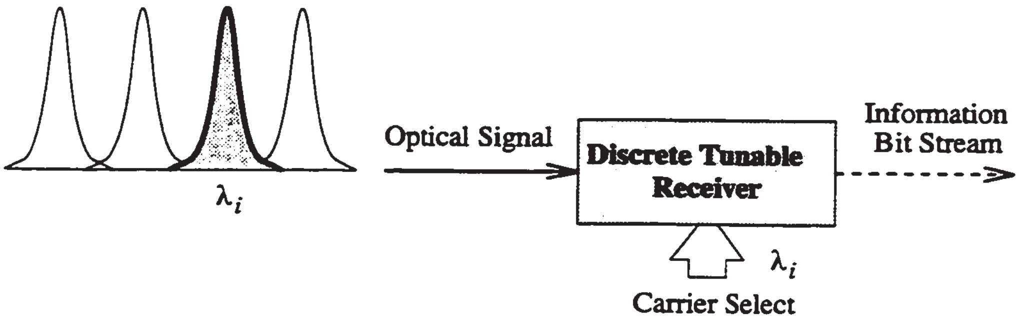


Figure 3.4: Discrete Tunable Light Detector.

Discrete Tunable Light Detector

This device is capable of separating and demodulating an information signal from one of many optical wavelength carrier signals. This operation is shown schematically in Figure 3.4. The control circuit selects one of the available wavelength carriers (channels) for demodulation. The receiver (detector) then filters out that carrier and demodulates the information signal. The wavelength resolution and the tuning speed, as in the case of tunable transmitters, are important parameters.

A tunable optical filter followed by an optical detector is the most common structure for tunable receivers [Sen92, Jon88]. Tuning can be performed passively or actively [Bra90]. In passive techniques, a mechanical feature of the detector structure is varied. Some examples of passive filters are tunable Mach-Zehnder filters, tunable Fabry-Perot filters, and Fabry-Perot etalons. Acousto-optical or electro-optical effects are applied in active filters to change the refractive index of a filter structure. A laser, such as a DFB or DBR laser, operating below threshold can also be used to selectively amplify and pass a wavelength channel.

Passive filters are able to resolve 100 to 1000 wavelength channels, while acousto-optical filters have channel resolutions in the order of 100 channels. Electro-optical techniques can provide channel resolutions in the order of 10 channels. However, they are very fast and have tuning speeds in the nanosecond range, much lower than the microsecond and millisecond ranges for acousto-optical filters and passive filters, respectively.

3.3.3 Lightwave Amplifiers

Line amplifiers are used to boost signal levels so that they can be sent longer distances while still being detectable. In order to be detectable, the signal-to-noise ratio (SNR) of the received signal must be higher than a minimum level, above which reliable detection of the signal is possible. Thousands of cascaded electronic amplifiers are used along a transcontinental communication link to keep the long haul link operating.

In single channel communications where only a single bit stream at rates less than a few Giga-bits per second is being sent, *repeaters* can be used. In these systems, a detector followed by an electronic amplifier can effectively boost the signal energy. In the case of signal reshaping, an equalization stage regenerates the original signal. This form of signal amplification, however, cannot be applied to wavelength multiplexed multichannel communications. In these environments, it is desirable to keep the signal in photonic form throughout the communication path between end nodes. Therefore, wideband purely photonic amplifiers are essential to optical networks.

A lightwave amplifier can be built from any structure that lases. Two main practical approaches exist: *laser diode amplifiers* and *doped fiber amplifiers.* The latter approach has made a revolution in long haul and wavelength multiplexed multichannel communications. Doped fiber amplifiers seem very attractive due to their simplicity of manufacture and coupling into the fiber link, polarization independence, wide bandwidth, and comparative freedom from crosstalk effects. Laser diode amplifiers, on the other hand, are expensive, hard to couple into the fiber link, polarization sensitive, and also have a high level of crosstalk. It should be noted however, that new progresses in heavily injected quantum well laser devices exhibit gain curve widths as great as 240 nm, which is the entire low attenuation band at 1.5 μm. In addition, doped fiber amplifiers are mainly available for 1.5 μm and not for 1.3 μm wavelengths which are favored for long links because of its low dispersion [Gre93, ST91].

There are basically three roles for lightwave amplifiers in optical communication systems.

- *Power Amplifier:* placed immediately after the transmitter laser to boost the transmitted power level.

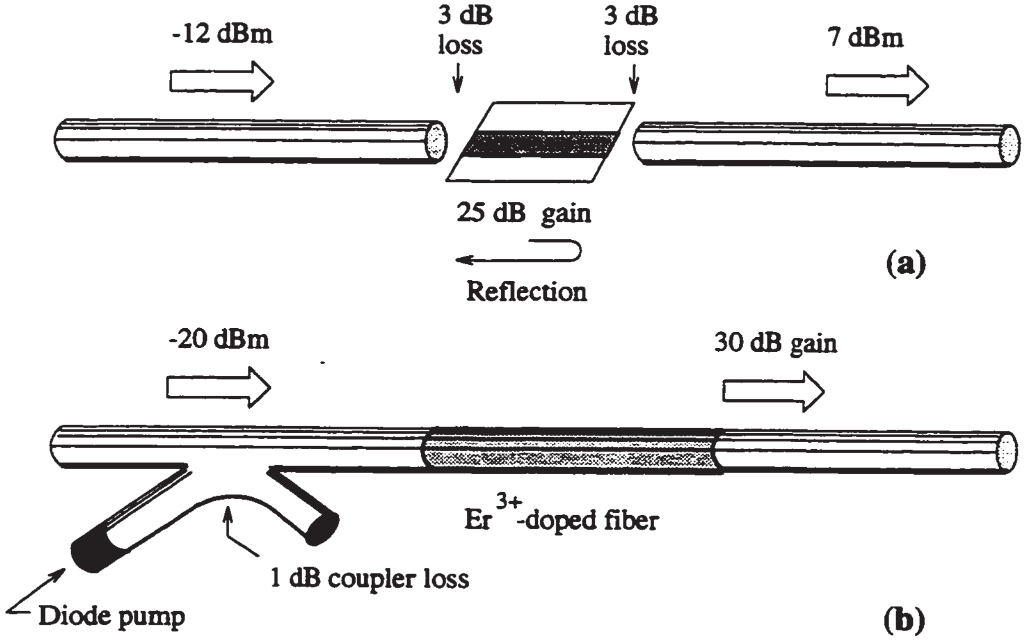


Figure 3.5: Two types of optical amplifiers.(a) Semiconductor laser diode (Fabry-Perot) amplifier. (b) Traveling wave (Erbium doped fiber) amplifier. The values shown are typical values.

- *Line Amplifier:* operates as a repeater at one or more locations between two end points.
- *Receiver Amplifier:* provides the required signal energy level for the photodetector that follows the amplifier at the receiver.

The bandwidth of optical amplifiers is at least 3 orders of magnitude greater than that of electronic amplifiers [Gre93]. Two types of lightwave amplifiers, *Fabry-Perot amplifiers* and *traveling-wave amplifiers* are shown in Figure 3.5. In the first type, the light amplification is obtained by multiple reflections of the light on the sides of the cavity, while in the second case, the reflectivity of both facets of the amplification medium is kept very low and the light signal is amplified in one pass through the medium. The interested reader can find the principles of the operation of photonic amplifiers and other related issues, such as noise, gain saturation, nonlinearity, and cross talk in [Gre93, ST91, Jon88, Sen92, Mid93b, Pal92].

It is worthwhile to add some discussion of doped fiber amplifiers that are very important in long haul communication links and in WDM networks. Erbium doped silica fibers, serving as laser amplifiers, are one of the important components of 1.55 μm fiber optic communication systems. They offer high gain (30 to 45 dB), with low noise, near the wavelength of lowest loss in silica glass fiber. Population inversion, which is an essential ingredient for lasing, is provided by a pump source which excites (pumps) atoms into the higher energy level. Pumping photons have higher energy levels, hence they have shorter wavelengths. Erbium is an excellent material as the dopant, since it has the necessary lifetimes for electron transitions, long for lasing transitions and short for pumping transitions [ST91, Gre93, Sen92]. Furthermore, the pumping wavelengths for Erbium atoms are commercially available, which are 9.8 μm and 1.49 μm.

In short, Erbium doped fiber amplifiers have the following advantages.

- The difficulty and expense of the facet reflectivity control needed in Fabry-Perot amplifiers do not exist.
- Lower passband ripple than that of Fabry-Perot amplifiers.
- Better noise performance than laser diode amplifiers.
- Very low coupling losses.
- It is polarization insensitive.
- It has a very low temperature sensitivity.
- Small high-power pump laser diodes are commercially available.

Doped fiber amplifiers however, have low gain curve width compared to some new laser diode amplifiers and are difficult to build for wavelengths other than 1.5 μm.

3.3.4 Couplers and Splitters

Passive couplers and splitters have been extensively used in optical networks. An optical fiber coupler is a device that distributes light from a main fiber into one or

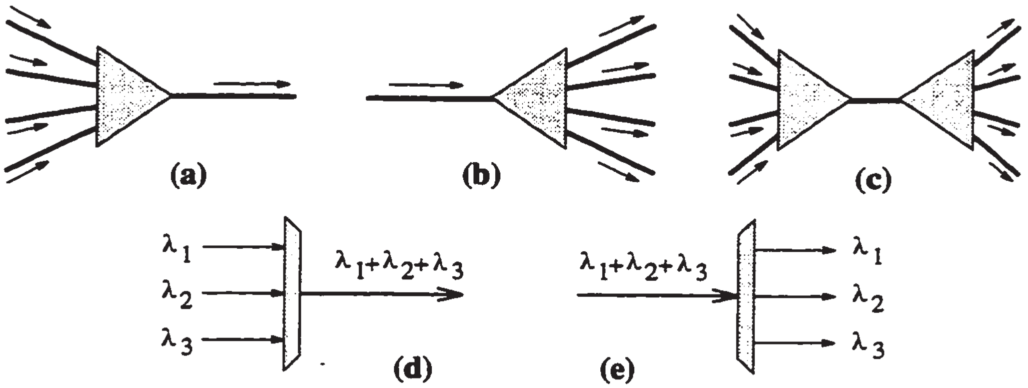


Figure 3.6: Optical couplers and splitters. (a) Combiner. (b) Splitter. (c) Star Coupler. (d) Wavelength Multiplexer. (e) Wavelength Demultiplexer.

more fibers. Optical fiber couplers are classified into two categories according to their interaction type: core interaction type, and surface interaction type. In the first case, the power transfer takes place through the fiber core cross section. In the surface interaction, the power transfers through the fiber surface by converting the guided core modes to both cladding and refracted modes which then enable the power sharing mechanism [Sen92].

Couplers can also be classified according to the number of inputs and outputs and their wavelength channel discrimination capability. All these classes are shown in Figure 3.6. The simplest form is a combiner that mixes and couples the light energy received from the inputs onto one output fiber, see Figure 3.6(a). Figure 3.6(b) shows a device that splits and distributes the received optical energy from a single input fiber to multiple output ports. This device is called a splitter. The next category, shown in Figure 3.6(c), is called *star* coupler which evenly distributes the received energy from each input port to all output ports. The operation of the these devices is insensitive to wavelength. Hence, they are wideband devices.

The next classes of optical couplers and splitters, shown in Figure 3.6(d) and (e), are called wavelength multiplexers and demultiplexers, respectively. A wavelength multiplexer combines different wavelength input signals onto the output fiber, while wavelength demultiplexers separate the wavelength channels of the received optical signals from the input fiber and sends them to different output ports, as shown in the

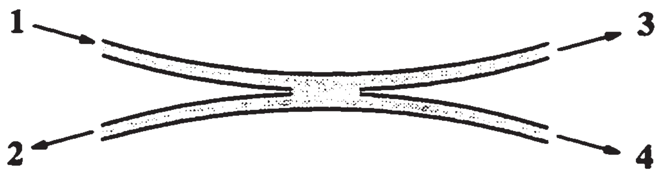


Figure 3.7: Four port fiber fused biconical taper (FBT) coupler.

figure. More details can be found in [Sen92, Gre93, ST91].

Figure 3.7 shows a four port coupler known as a fused biconical taper (FBT) coupler. To build this device, two fibers are twisted together and then spot fused under tension such that the fused section is elongated to form a biconical taper structure. There are some losses associated with non-ideal practical couplers and splitters and only a portion of the total power is coupled between two fibers. We can define these loss parameters with reference to Figure 3.7 as follows [Sen92, Jon88].

$$Excess\ Loss = L_{excess} = 10 \log \frac{P_1}{(P_3 + P_4)}\ (dB) \tag{3.1}$$

$$Insertion\ Loss\ (ports\ 1\ to\ 4) = L_{insertion} = 10 \log \frac{P_1}{P_4}\ (dB) \tag{3.2}$$

Two other measures are also defined that consider the directional isolation and coupling ratio, respectively.

$$Crosstalk\ (4\ port\ coupler) = 10 \log \frac{P_2}{P_1}\ (dB) \tag{3.3}$$

$$Split\ ratio = \frac{P_3}{(P_3 + P_4)} \tag{3.4}$$

In the case of a star coupler with N input and N output ports, the splitting loss and the excess loss are similarly defined as

$$L^{Star}_{splitting} = 10 \log N\ (dB), \tag{3.5}$$

and

$$L^{Star}_{excess} = 10 \log \frac{P_i}{\sum_{j=1}^{N} P_j}\ (dB), \tag{3.6}$$

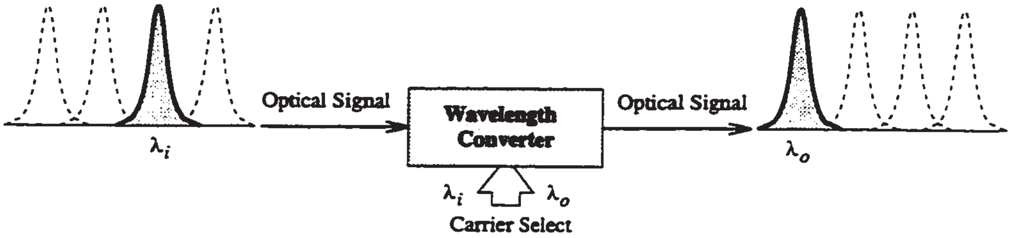


Figure 3.8: Wavelength Converter.

where excess loss is written for input port i, and $i = 1, \cdots, N$.

3.3.5 Wavelength Converters

The device shown in Figure 3.8 changes the carrier wavelength (frequency) of an input signal from λ_i to λ_o, thus it is called *wavelength converter*. The modulated information signal remains unchanged [Mid93a]. There has been an increasing interest in the research and manufacture of these devices recently, since it is an important component in the optical networks. In particular, in order to perform space switching, the wavelength converter can be used in conjunction with fixed optical filters to route the signals to their destinations. A number of networks have been proposed that have employed wavelength converters [AT95, AT97, LL95]. In a tunable wavelength converter, as shown in Figure 3.8, the input and output channels are selectable through an electronic control circuit.

Wavelength conversion can be obtained using several mechanisms, such as four wave mixing in semiconductor optical amplifiers, absorptive bistability in DBR and Y lasers, as well as gain saturation in semiconductor optical amplifiers [DPM+93]. A wavelength conversion technique using optical depletion of the carrier density in the gain region of a DBR laser operating above threshold is presented in [DPM+93]. This technique can obtain a BER (bit error rate) penalty-free wavelength conversion over 18 nm at a bit rate of 2.5 Giga-bits per second. A wavelength converter that has unidirectional output and converted light wavelength tuning capability has been demonstrated in [TKY+94]. The device emits converted light only when light input is injected (i.e. optically triggered conversion). In this device, optical signals are unidirectionally transmitted, such that the output power of the light from the light input-end facet of the device is 30 dB smaller than that from the output-end facet.

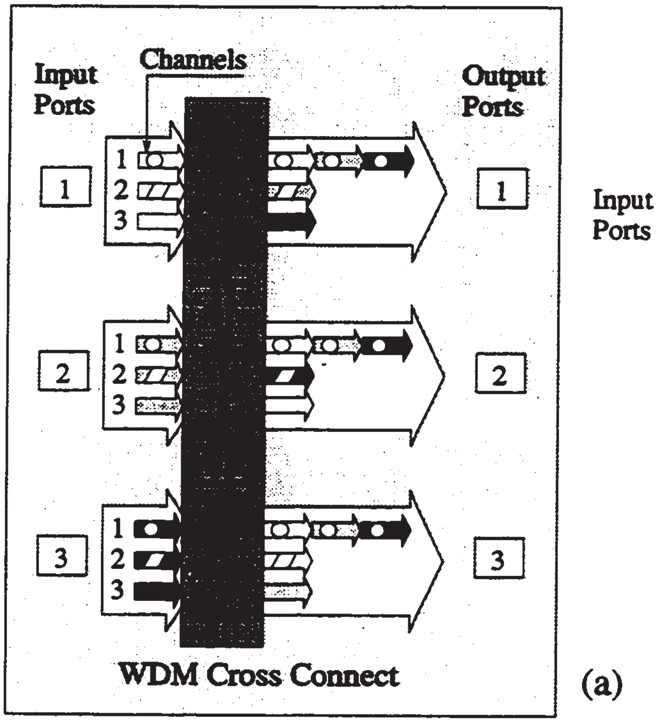
Very fast wavelength conversion is possible when semiconductor laser amplifiers are used. Tunable wavelength conversion using single mode semiconductor lasers can easily be performed and has been achieved by using a multi-electrode DFB laser or a DBR laser. Repeated wavelength conversion of 10 Gbps pseudo-random signals is demonstrated in [YIYO95, YIT+94], using super-structure grating DBR lasers operating in the 1.55 μm wavelength region. Error-free wavelength conversion with and very low power penalty (less than 1 dB) can be achieved over a broad wavelength range from 1.486 to 1.573 μm (about 90 nm wide) periodically at a frequency resolution of 40 MHz. The wavelength switching time is estimated to be about 5 ns.

Wavelength conversion of an 18 Gbps data signal over 15 nm (corresponding to about 1.9 THz) has been presented in [SPE+93]. This technique has the potential of operating at extremely high bit rates as the characteristic time for the operation speed is measured to be 200 fs. In [ZPD+94], a theoretical analysis and experimental measurements of broadband optical wavelength conversion by four-way mixing in semiconductor traveling wave amplifiers is presented. Two other experimental wavelength conversion techniques are presented in [KKYW92] and [GWK+92].

3.3.6 Wavelength Cross Connect

A wavelength cross connect or wavelength router is an advanced wavelength routing splitter-coupler device [Mid93a]. This device in its complex form can selectively route packets from different channels on each input port to a different output port. It is in fact a fast multichannel rearrangeable optical switch. Figure 3.9 illustrates the wavelength router operation and its corresponding wavelength interconnection matrix for a three wavelength system. In this figure, λ_{ij} denotes the wavelength j of input port i. It can easily be seen that while wavelength λ_1 has been designated as a fully broadcast channel, the other channels are selectively routed based on the source input port.

Wavelength cross connected networks are considered in this thesis. The WDM cross connect used in most of the proposed architectures is a static wavelength router



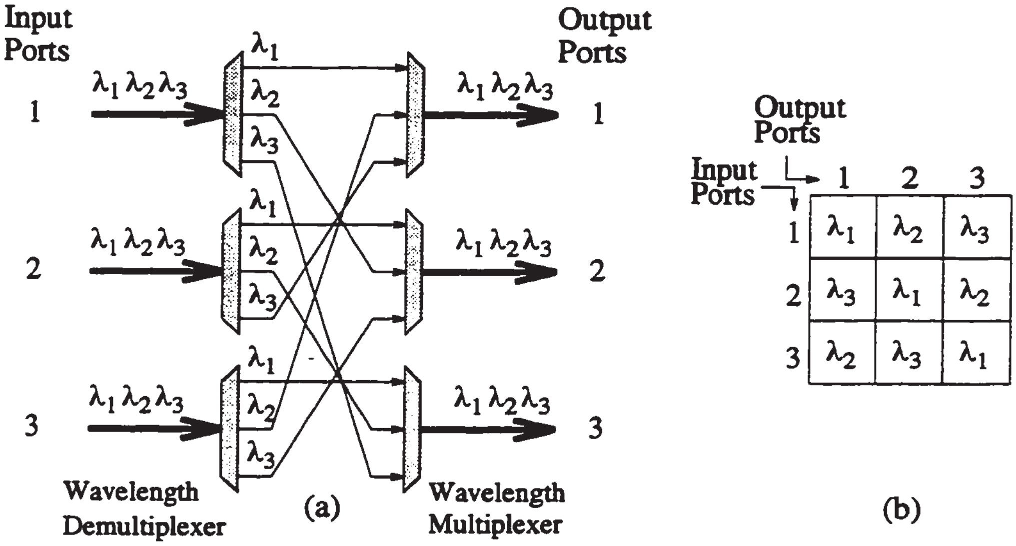
Input Ports ↓ / Output Ports →	1	2	3
1	λ_{11}	λ_{11} λ_{13}	λ_{11} λ_{12}
2	λ_{21} λ_{22}	λ_{21}	λ_{21} λ_{23}
3	λ_{31} λ_{33}	λ_{31} λ_{32}	λ_{31}

(b)

Figure 3.9: Wavelength Cross Connect (a) Schematic diagram. (b) Wavelength Interconnection Matrix

where the wavelength channel interconnection matrix remains fixed. An example is shown in Figure 3.10. As shown in the corresponding connection table, full connectivity is assured and wavelength channels are reused. Other more complex connection patterns can easily be incorporated in the WDM cross connect. [Mid93a] shows a wavelength routed perfect shuffle interconnection that requires only two wavelengths, independent of the number of nodes in the network.

Wavelength cross connects are key components in optical networks which implement spatial wavelength channel reuse. The number of available input and output ports, wavelength channel resolution, and reconfiguration speed are parameters of interest for wavelength routers.



Input Ports \ Output Ports	1	2	3
1	λ_1	λ_2	λ_3
2	λ_3	λ_1	λ_2
3	λ_2	λ_3	λ_1

Figure 3.10: Static Wavelength Cross Connect (a) Architecture. (b) Wavelength Interconnection Matrix

3.3.7 Optical Switches

An optical switch is a device that directly controls the optical signal path. Although the control section of an optical switch may be totally or partly electronic, these devices are optical elements. Nonlinear optical effects can be employed in order to switch or route optical signals. The switching control in these devices is performed by a separate control beam, hence these are truly *all-optical* switches. Materials whose optical properties, specifically, the refractive index, vary with the intensity of the control beam are crucial to the ultimate implementation of all-optical switches. The interested reader can find details of the characteristics of all-optical switching devices and the relevant technologies in [Mid93b].

Spatial optical switching is very essential to avoid electro-optical conversion bottlenecks throughout the communication path. Unfortunately, no commercially viable switch architecture has been found that can provide low-attenuation low-crosstalk purely photonic switching of size greater than 16 × 16. However, a fully nonblocking lightwave switch can be built by cascading several smaller size switches in a number of stages (layers) [Gre93].

Optoelectronic hybrid switches control the route of the signal information, but do not retain this information as light. Such switches are transducers that convert information between electronic and optical form. In fact switching is obtained by controlling the conversion process. The conversion process speed also imposes a limit on the switching speed.

Switching of data rates up to 2.5 Gbps in integrated matrix switches with dimensions up to 16 × 16 are available, and optoelectronic space-division matrix switches operating at 10 Gbps are currently foreseeable [Mid93b].

Much work has been devoted to photonic fast packet switching architectures [JM95]. A variety of internal multiplexing methods and their combinations, including wavelength, time, code, and space division multiplexing, are used in these architectures. In some of them, such as the FOX [Goo89] and the HYPASS [AGKV88] networks, an optical passive star coupler is used as the optical switching fabric controlled by electronic devices. Contention resolution schemes applied to electronic switches are used in some switches as well. A knockout controller is used in the photonic knockout switch [EKY92], and a Batcher-banyan network is employed in BHYPASS [Goo89], while the Star-Track switch [Goo89] uses an electronic token ring type control network. A thorough survey of photonic switching is presented in [JM95]. This review considers various forms of optical fast packet switching techniques, including single-hop shared media, multi-hop, time and code division, space-division, self-routing, and free space switching methods.

3.4 Multiplexing

As mentioned before, the maximum electronic speed at the edges of the network prevents us from using the entire bandwidth available as one communication channel. Thus, the key in lightwave network design is to introduce concurrency among multiple user transmissions into the network. This can be provided based on different multiplexing methods. If the available bandwidth is divided into different frequency channels and then different users are assigned to different channels, frequency division multiple access (FDMA), or equivalently wavelength division multiple access (WDMA) would result. Time division multiple access (TDMA) provides concurrency

	Frame *i*						Frame *i+1*						Frame *i+2*						Frame *i+3*						
Channel 1	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	
Channel 2	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	.....
Channel 3	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	
Channel 4	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	Time →

Table 3.1: An example of ITDMA frame. It is assumed that there are 6 stations and 4 channels.

among several users by assigning different time slots to each user. Distinct orthogonal wave shapes or codes are used in code division multiple access (CDMA) to distinguish different communications [Muk92b, Ram93].

3.4.1 Time Division Multiplexing

Predetermined channel time slots are assigned to each user in TDMA. The user has access to the entire channel bandwidth, but only during its allocated time slots. In multichannel multiaccess environments where stations need to access all the channels, an extension to the regular TDMA time slot assignment is used to ensure that all the stations have been assigned time slots on all the channels. This type of multichannel TDMA is called interleaved TDMA (ITDMA). Figure 3.1 shows an example of ITDMA time slot assignment.

As shown in Table 3.1, all the time slots on the channels are synchronized. In a specific time slot, each station has access to only one channel and there is no possibility of collision on any. ITDMA, like ordinary TDMA, is mostly suitable for the environments where all the stations are sending data uniformly and constantly to all the channels. However, it is not efficient for bursty or focused traffic patterns. In these cases many of the time slots may remain idle when there is not any data ready for transmission in some of the stations. Even those nodes which have data ready, have to send their data in the form of one packet per frame. This can cause long delays.

The need for stations to acquire frequency, phase, bit timing, and bit framing synchronization makes TDMA more complex than WDMA (FDMA) [Muk92b].

3.4.2 Wavelength Division Multiplexing

Wavelength division multiplexing (WDM) is based on the division of the available optical bandwidth of a single fiber into a number of lower bandwidth wavelength channels. The intent is to provide concurrent parallel channels each operating at speeds available to electronic circuits [Bra90, SKG91, JM93, Gre93, Muk92b, Ram93]. There are a number of advantages associated with this method. First, the bandwidth of the optical fiber is more easily accessed in the wavelength domain directly, rather than in the time domain. Second, wavelengths can be used to perform network and system oriented functions, such as routing and switching. Third, the transparency to data communication format and speed is obtained. This means that different data, signaling format, and communication speed can be used in each channel independently of those used on the other channels. Fourth, there is no need for synchronization of transmissions on different channels. Finally, the electronic interfaces only need to run at electronic speeds.

The proposed WDM networks in the literature can be classified into two main groups based on the number of wavelength channels assumed to be available. Those which assume availability of hundreds or even thousands of channels are referred to as *dense* WDM networks, while those that assume only a moderate number of available wavelength channels, less than a hundred channels, are called *coarse* WDM networks [Bra90, Ram93].

3.4.3 Code Division Multiplexing

Spread spectrum techniques such as code division multiple access (CDMA) have been used in military communications for a long time. Its use has also been investigated for satellite and mobile radio communications. Currently, CDMA is a standard communication protocol in wireless mobile communications (IS-95) [Vit95].

Spread spectrum methods offer several potential advantages. Spread spectrum makes efficient use of the channel by providing communications among a large community of relatively uncoordinated users. It also permits multiple users to access the channel with no waiting time. Furthermore, spread spectrum can overcome the effects of strong interference, or it can be used to hide (encrypt) the signal [PP90, Vit95].

Spread spectrum, or CDMA, is based on the assignment of orthogonal codes to the address of each user, which substantially increases the bandwidth occupied by the transmitted signal, or equivalently spreads its spectrum. CDMA capacity is only interference limited. Therefore, any reduction in interference converts directly and linearly into an increase in capacity. In such environments as wireless mobile communications, spatial isolation through the use of multibeam or multisectored antennas can largely decrease the interference. Suppressing quiet periods during voice communications also reduces the interference, and consequently provides a proportional increase in capacity [GJP+91]. These two factors are shown to be sufficient to render the capacity of CDMA at least double that of FDMA and TDMA under similar assumptions [GJPW90]. [GJP+91] shows that a power-controlled CDMA system used in cellular systems can exhibit a net improvement in capacity over digital TDMA on the order of 4 to 6, and on the order of 20 over current FM/FDMA.

CDMA has also been considered as a communication protocol for optical networks [Van89, FV88, PP90, CSW89, TNO85], although it has not been found to be as attractive as in wireless mobile communications. Some hybrid solutions such as FDMA/CDMA can combine two multiple access methods, so that the beneficial aspects of each technique mitigate the shortcoming of the other [Van89]. Thus, a random-carrier CDMA was introduced that does not require stable tunable lasers with accurate wavelength control. [FV88] shows that random-carrier CDMA can help overcome laser frequency drift and laser linewidth. It should be noted, however, that a fundamental difference exists between some optical processing techniques, especially noncoherent processing using intensity modulation, and non-optical conventional processing methods. Digital intensity modulated optical signals are "unipolar" consisting of sequences of 0's and 1's. As a result, the "bipolar" codes that have values of +1 and -1, used in conventional systems, cannot be applied to optical systems. New optical orthogonal codes are needed. Optical orthogonal code design methods are presented in [CSW89]. [TNO85] and [PP90] consider the use of maximal length codes and Gold codes for this purpose.

3.5 Broadcast-and-Select Networks

In broadcast-and-select architectures, the transmission from each station is broadcast to all of the stations in the network. The desired signal is then extracted from all the signals at the receiver. It is apparent that tunability is needed in the transmitting or receiving side. For example, in the case of tunable transmitters and fixed-tuned receivers, all the stations select their data communication channels with respect to the channels that their destination stations are tuned to. It is obvious that if two or more stations attempt to send data to the same destination station, a collision occurs. Alternatively, we could have fixed-tuned transmitters and tunable receivers, or both transmitters and receivers tunable. The network fabric is totally passive in broadcast-and-select networks, consisting of passive combiners and splitters. Optical amplifiers may be used in these networks to boost the optical energy.

Star and bus topologies are two of the most common broadcast-and-select optical network architectures. Optical star and bus topologies are shown in Figure 3.11. The optical power received from any input port of the star coupler will be distributed evenly to all the output ports according to a specified splitting ratio. In the bus architecture, each station transmits onto the bus through a coupler and receives from the bus via another coupler.

It easily be seen that if each station transmits with a power of P and there are N stations in the network, then, in a star topology, each station receives at most a power of P/N and the overall splitting loss is $10 \log N$ dB. In a bus topology, each station transmits and receives via a pair of 2×2 couplers. Each coupler distributes the power from the input ports to the output ports according to a specified splitting ratio. If this ratio is optimized, only P/N^2 is received by each station and the overall splitting loss is $20 \log N$ dB. As a consequence, the bus architecture can support less stations than the star. The star architecture is preferred to the bus in this regard. We should note, however, that optical amplifiers can be used to increase the number of supported stations [Gre93, Ram93].

There are a number of problems with the broadcast-and-select approach. These networks require a large number of wavelengths, typically on the order of the number

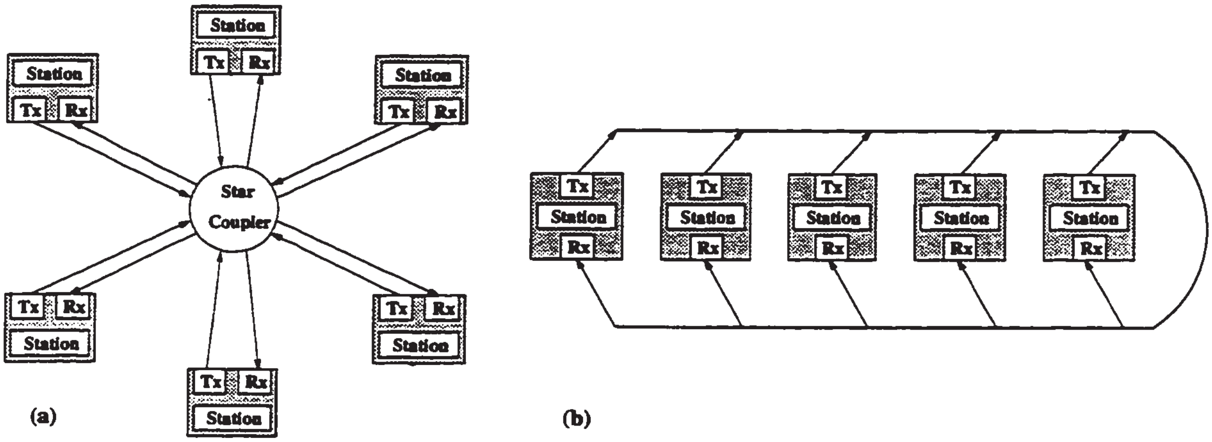


Figure 3.11: Broadcast-and-select architectures. (a) Star architecture. (b) Bus architecture.

of stations in the network. When we increase the number of channels, other problems such as the stability of the lasers and filters become important. The second issue is the splitting loss. Since the transmitted power from a station is broadcast to all the stations in the network, a small fraction of the transmitted power will reach each station, and this becomes smaller as the number of station increases. Consequently, optical amplifiers would be required beyond a certain point. Finally, it is difficult to scale broadcast-and-select networks to WANs. It is very obvious that a countrywide network with million users cannot use a broadcast-and-select network using a single star coupler.

3.6 Single-Hop and Multi-Hop Networks

In single-hop networks, information reaches its final destination directly without being converted to electronic form in between. As mentioned before, for broadcast-and-select single-hop networks, tunability is needed on the transmitting and/or receiving side. In order to avoid tuning altogether, multi-hop networks have been proposed. In these networks, each station is equipped with a small number of fixed-tuned transmitters and receivers. Each transmitter and receiver operates at a different wavelength. Even though the physical topology of the network can be any broadcast topology

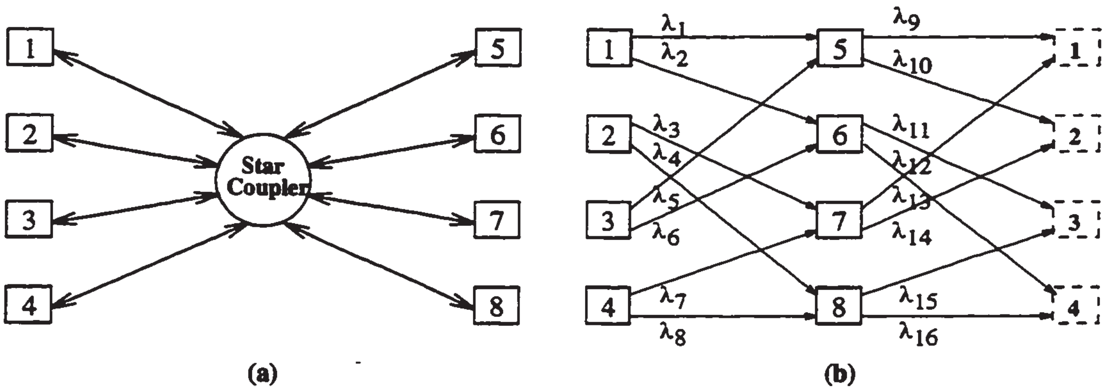


Figure 3.12: Multi-hop architectures. (a) Physical topology: Star. (b) Logical Topology: Shufflenet.

such as star or bus, each station can directly transmit to the stations that have their receivers tuned to its transmit channel. Packets intended for other stations have to be routed through intermediate nodes. Each intermediate station converts the received optical signal to electronic form, then the packets are retransmitted to the next node according to their destination addresses. Obviously, there exists a *logical topology* on top of the physical topology that determines the actual connectivity between the stations in the network. An example is shown in Figure 3.12 where the physical topology is a star, and the logical topology is called *Shufflenet.* Each station, in this example, is equipped with two fixed-tuned transmitters and two fixed-tuned receivers. As shown in the figure, Station 2, for example, can directly transmit to Stations 7 and 8, but it has to reach the other stations through intermediate nodes. A good review of single-hop networks and architectures can be found in [Muk92b], and for multi-hop networks in [Muk92c].

3.7 Wavelength Routing

Wavelength routing is an efficient way to avoid the problems of broadcast-and-select networks, i.e. lack of wavelength reuse, splitting loss and lack of scalability. Figure 3.13 shows an example of wavelength routing in a network with only two wavelength channels. The routing nodes redirect the incoming channels to one of the outgoing

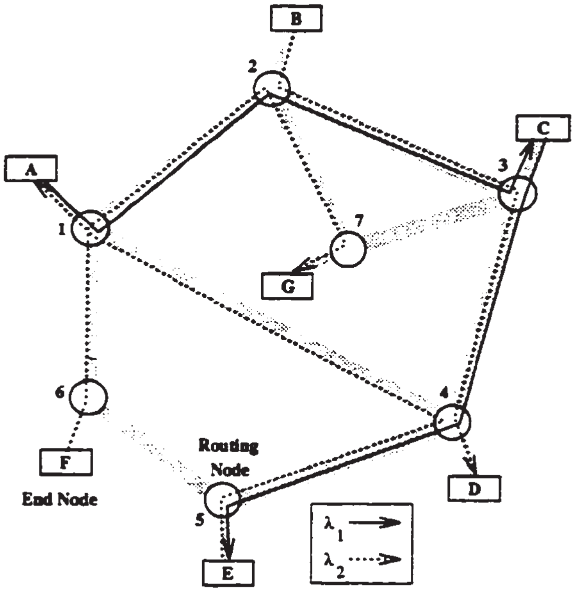


Figure 3.13: Wavelength reuse in a wavelength routing network.

links based on their wavelengths. This redirection can be based on a fixed routing table, or it can be performed according to a dynamic schedule based on traffic demands. In the figure only unidirectional routing nodes and links are assumed. It is apparent that with more wavelength channels and bidirectional links, more complicated scenarios and more efficient connections are possible.

Finding the minimum number of wavelengths required to support a given set of connections is an NP-complete problem [Ram93]. Simulation studies, however, can help us to find the best number of wavelengths for a specific topology. Routing, network control, and management for wavelength routing networks have been extensively studied in telecommunications and computer networks [Ram93, Gre93].

3.8 Network Architectures with Wavelength Reuse

These architectures incorporate network clustering techniques to achieve spatial wavelength reuse. Many of these networks use a traditional multihop topology with several optical-electronic conversions. Recently, a new trend in research has begun to investigate the possibility of incorporating wavelength routing and wavelength reuse

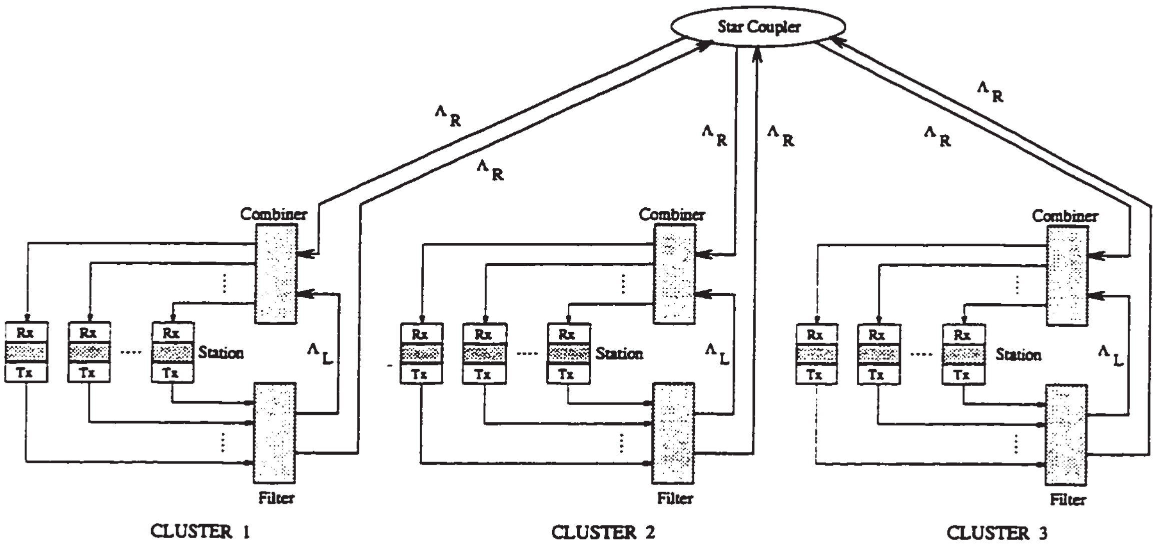


Figure 3.14: Hierarchical 2 level wavelength reuse.

into multichannel single-hop designs. A partial broadcast-and-select architecture is formed where a subset of stations receive data transmissions based on the location of the transmitting station and the channel used. In such a network the total capacity is no longer upper bounded by the number of available channels. It should be noted that network clustering is another area of research that looks for topologies with optimum network performance for some specific traffic patterns and communication technology [KK80, LA93].

There are two approaches for wavelength reuse depending on the type of reuse scheme used. We discuss them in the following sections.

3.8.1 Hierarchical Wavelength Reuse

The basic architecture of networks with hierarchical wavelength reuse is formed by partitioning the stations into L levels of station clusters [HWM94, GG92, DS94, KFG94, JT94a]. Figure 3.14 shows an example of a two level hierarchical topology with 3 clusters that reuses a set of local wavelengths, denoted by Λ_L, for intra-cluster communications and uses another set of remote wavelengths, shown by Λ_R, for interconnection of the clusters, i.e. inter-cluster communications. The optical

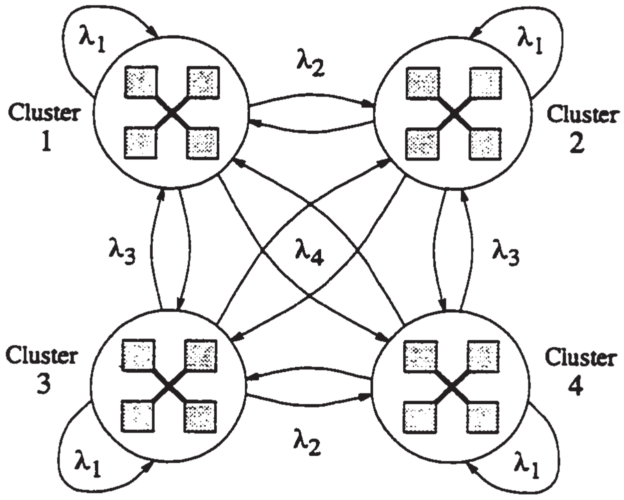


Figure 3.15: Flat wavelength reuse with 4 clusters and 4 wavelengths.

filter separates the local wavelengths from the remote ones and resends them back to the same cluster while it sends the remote wavelengths to a passive star coupler. This star coupler distributes the remote wavelengths to all clusters. Each station is equipped with two receivers to access the local and remote wavelengths and one tunable transmitter to send data packets.

In the hierarchical architectures, the corresponding MAC has to be aware of the two different types of communication.

3.8.2 Flat Wavelength Reuse

Wavelength spatial reuse can also be provided with a flat topology [Mat93, AT95, AT97]. In a flat topology, network clustering is used but the wavelength channels do not go through different routing levels as seen in the hierarchical case. The main idea of the flat architecture with 4 wavelengths and 4 clusters is shown in Figure 3.15. As shown, λ_1 is reused for intra-cluster (local) communications. λ_2 is used for inter-cluster communication between clusters 1 and 2, and between clusters 3 and 4. λ_3 is the communication channel for clusters 1 and 3, as well as clusters 2 and 4, and so on.

A network introduced in [Mat93], named MANDALA, uses the same wavelength

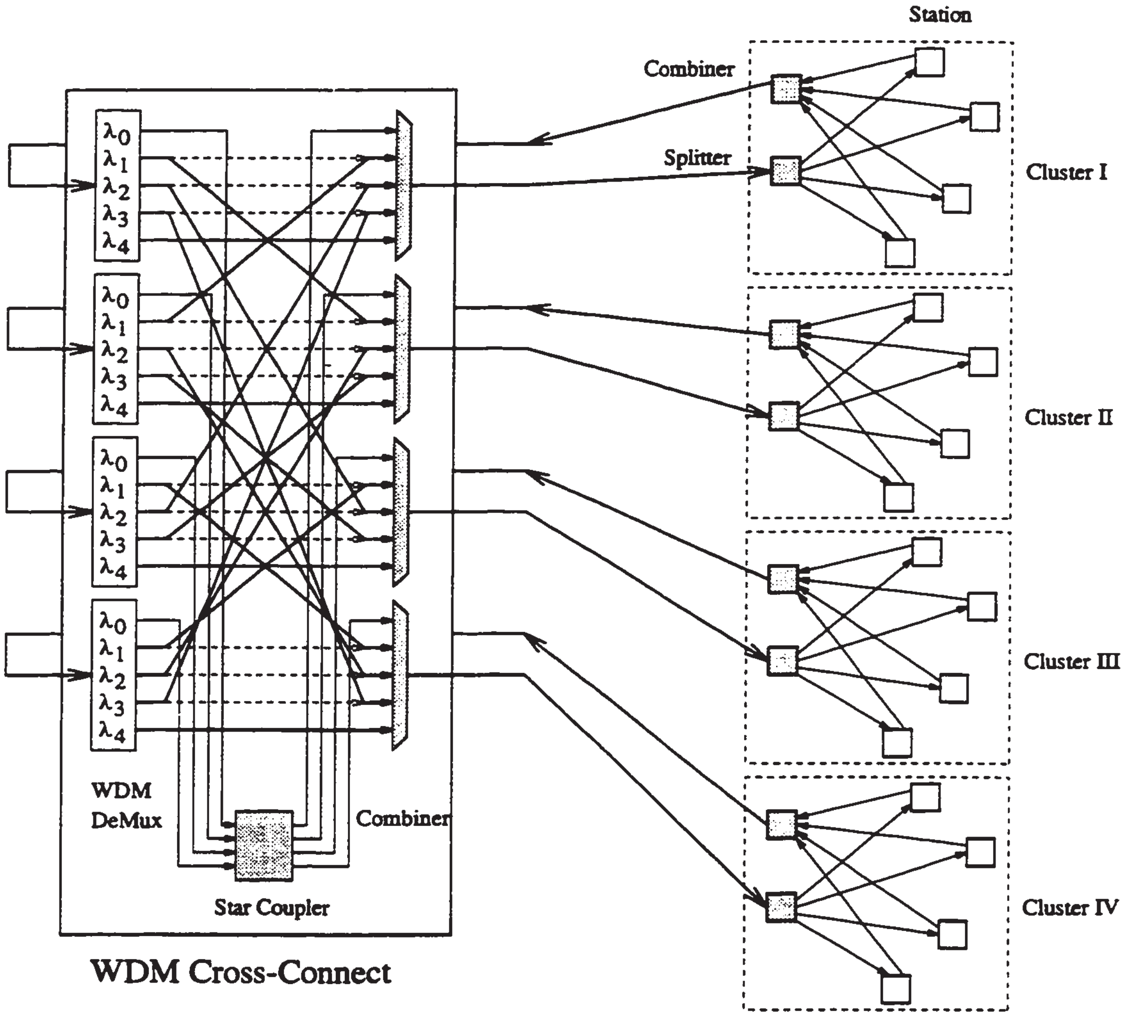


Figure 3.16: Flat wavelength reuse in MANDALA.

reuse shown in Figure 3.15. The network consists of n wavelength channels and N stations divided into subnetworks. One wavelength channel, λ_0 is dedicated as a global control channel. Each station is equipped with one tunable transmitter and one tunable receiver. Figure 3.16 shows this architecture with 4 data channels and one control channel. Stations use a form of the slotted ALOHA protocol to make their reservations on the control channel.

Chapter 4

Single-Hop WDM Cross Connect Networks

4.1 Prelude

The enormous bandwidth of optical fiber cannot be accessed as a single communication channel, since each user station is connected through its electronic interface whose speed is limited to the maximum electronic processing speed. The key in the design of optical networks is to introduce concurrency among multiple user transmissions [Muk92b]. As mentioned before, wavelength division multiplexing (WDM) offers the most promising technique for achieving this. The state-of-the-art of the optical technology permits a modest number of wavelength channels that is unfortunately less than that desired. As a result, channel reuse is necessary to accommodate a reasonable number of stations. Wavelength routing, on the other hand, can help by using optical processing to direct data packets without going through optical/electronic conversions.

WDM cross connect architectures combine wavelength channel reuse with wavelength routing. User stations are assumed to have access to all or a set of the available channels through the tunability of their transmitters and receivers. Some form of medium access control (MAC) protocol is needed to coordinate multiple user accesses

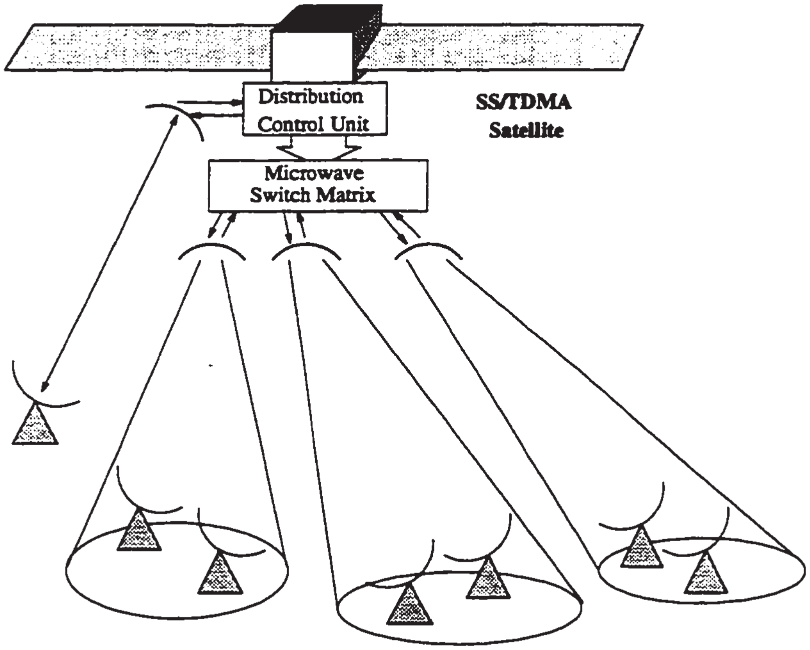


Figure 4.1: SS/TDMA system.

to the network. In this chapter, we propose single-hop WDM cross connect architectures which perform medium access control across a spatial WDM cross connect. A flat wavelength reuse scheme is used in these architectures. Pre-reservation based and random MAC protocols are introduced for the proposed systems.

In pre-reservation based control, all communications go through a reservation stage. Channel allocations can then be obtained based on the traffic demands of all active stations. A central controller may be used to perform the allocations and report the results to all the stations. In distributed control, all the stations execute the same algorithm to determine the appropriate allocations. Many pre-reservation techniques assume the time axis is partitioned into frames and each frame in turn, divided into time slots.

The time slot scheduling problem has been extensively investigated for satellite switched-time division multiple access (SS/TDMA) systems [IUMY77, Inu78, Inu79, BCW81, Gan92]. The basic idea of SS/TDMA is shown in Figure 4.1. The satellite interconnects several different zones. The signals from each zone are cyclically interconnected to other beams or zones so that the set of transponders appears to have beam-hopping capability. The SS/TDMA technique utilizes the merits of high gain spot-beam antennas along with the efficient TDMA method of providing complete connectivity of the coverage area. A TDMA frame is divided into a number of switching modes, and each mode is assigned a fixed on-board switch connection so that the traffic from various regions is routed to designated regions without conflict.

We propose new scheduling methods for WDM cross connect networks based on the extensions of the schemes used in SS/TDMA system. Full connectivity can be obtained by applying a wavelength routed-interleaved time division multiple access (WR-ITDMA) method. WR-ITDMA uses a fixed time slot assignment (TSA) which is an extension to basic TDMA. Since the TSA is static in WR-ITDMA, the performance of the system is poor in the case of non-uniform traffic patterns. Dynamic TSA, proposed in this thesis, tries to dynamically change the time slot connection table so that it can be configured according to the traffic requirements of the network. If there are N stations, the traffic matrix is an $N \times N$ matrix $\mathbf{D} = [d_{ij}]$ where element d_{ij} $(i, j = 1, \cdots, N)$ gives the number of packets (each packet is one time slot in duration) queued in station i and destined to station j.

The main idea in dynamic scheduling, either in SS/TDMA or in the proposed dynamic TSA, is to analyze the traffic matrix in terms of switching mode matrices. The result is the decomposition of the traffic matrix to a sum of switching modes (matrices), which we call "modal decomposition" (MD) of the traffic matrix. In an SS/TDMA system the switching modes (SMs) are scalar multiples of permutation matrices, where a permutation matrix is a matrix obtained by permuting the rows or columns of the identity matrix. In the following we will show that the dynamic TSA in WDM cross connect networks is a more complicated problem than that of SS/TDMA. We will see that a small subset of WDM cross connect networks, which we call the "trivial case" is the exact analogy of SS/TDMA. The TSA methods used in SS/TDMA with some generalizations and modifications can be used in the case of WDM cross connect networks. In the following, we assume each station is equipped with one tunable transmitter and one tunable receiver. Thus, the dynamic TSA searches for the tuning time tables for the transmitters and receivers of all the stations in the network, such that it can keep a good level of the system performance.

Again, this TSA may be obtained in either a centralized or decentralized fashion using the traffic demands of all the stations. Each station then uses the connection time table to tune its transmitter and receiver in the proper time slots. Note that the decomposition procedure must be very fast to avoid bottlenecks in system operation.

4.2 Network Architecture

A WDM cross connect routes incoming packets to different output ports according to their transmitted wavelength. All-optical routing is thus possible with the use of proper connections inside the cross connect. In the following sections we introduce optical networks which use static WDM cross connects. In the following, N denotes the total number of stations in the network and n gives the number of available wavelength channels.

4.2.1 Trivial Case: $N = n + 1$

This case is similar to an SS/TDMA system. However, in this case the switching is being done in the stations instead of the central satellite switch module. Figure 4.2 shows the "unfolded" architecture, so that all the stations transmitters are shown on the left, and the receivers on the right. All stations are equipped with one tunable transmitter and one tunable receiver. Both the transmitter and the receiver are assumed to have fast wavelength agility across the entire set of wavelengths, $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. As shown in the figure, station outbound and inbound fibers are directly connected to the WDM cross connect. The WDM cross connect wavelength routes the incoming packets to the proper output port according to the connection table. An example of a connection table is shown in Table 4.1. Dynamic TSA in this case can apply the same schemes used in SS/TDMA.

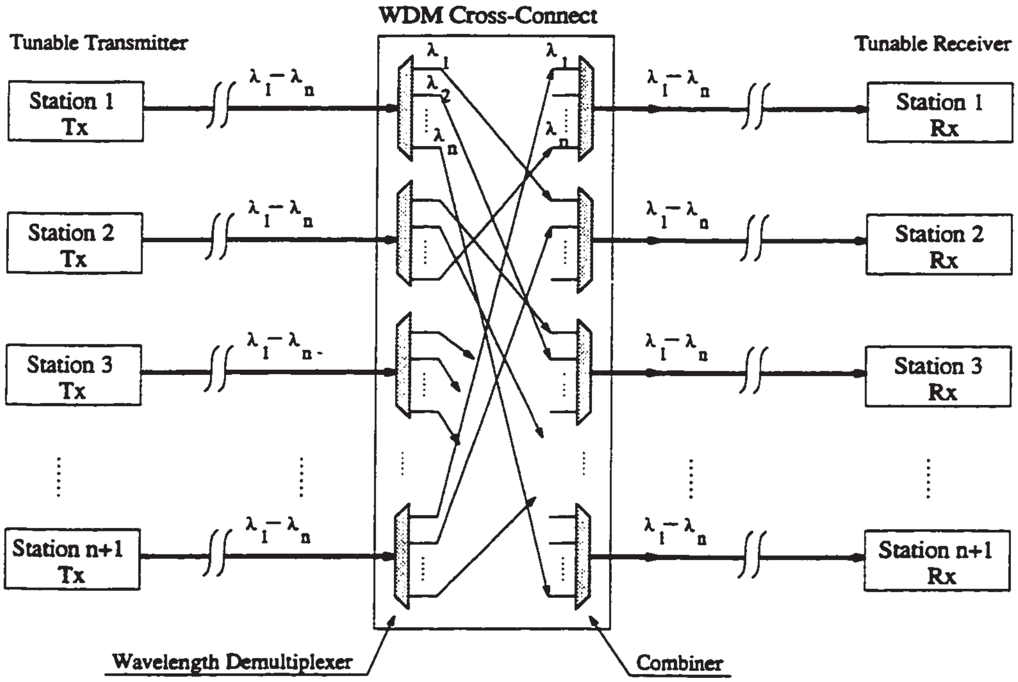


Figure 4.2: A WDM cross connect network: "Trivial Case".

An example of a traffic matrix for such a system is

$$\mathbf{D} = \begin{array}{|c|c|c|c|}\hline 0 & 4 & 6 & 4 \\ \hline 3 & 0 & 1 & 2 \\ \hline 3 & 4 & 0 & 7 \\ \hline 5 & 1 & 2 & 0 \\ \hline \end{array}$$

for $N = 4$. It can be seen that each station does not send any data to itself, so the traffic is zero on the main diagonal.

Source Station	*Destination Station*					
	1	2	3	$\cdots$	n	$n+1$
1	•	λ_1	λ_2	$\cdots$	λ_{n-1}	λ_n
2	λ_n	•	λ_1	$\cdots$	λ_{n-2}	λ_{n-1}
3	λ_{n-1}	λ_n	•	$\cdots$	λ_{n-3}	λ_{n-2}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
n	λ_2	λ_3	λ_4	$\cdots$	•	λ_1
$n+1$	λ_1	λ_2	λ_3	$\cdots$	λ_n	•

Table 4.1: Wavelength Routing Connection Table for the "Trivial Case".

Later we will show how we expand this matrix to its switching modes, i.e.

$$\mathbf{D} = \begin{array}{|c|c|c|c|}\hline \; & \; & 5 & \; \\ \hline \; & 0 & \; & \; \\ \hline \; & \; & \; & 5 \\ \hline 5 & \; & \; & \; \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & 2 & \; & \; \\ \hline 2 & \; & \; & \; \\ \hline \; & \; & \; & 2 \\ \hline \; & \; & 2 & \; \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & \; & \; & 4 \\ \hline 1 & \; & \; & \; \\ \hline \; & 4 & \; & \; \\ \hline \; & \; & 0 & \; \\ \hline \end{array}$$
$$+ \begin{array}{|c|c|c|c|}\hline \; & 1 & \; & \; \\ \hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \; & \; & 0 & \; \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & 1 & \; & \; \\ \hline \; & \; & 1 & \; \\ \hline 1 & \; & \; & \; \\ \hline \; & \; & \; & 0 \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & \; & 1 & \; \\ \hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \; & 1 & \; & \; \\ \hline \end{array} .$$

As shown each mode has only one non-zero element in each row and column. This is because of the fact that the elements in each row (column) share the same transmitter (receiver). Each switching mode is in effect for a period equal to the largest number in each mode matrix. As a result, the other transmitters and receivers that have less traffic to be sent in this mode stay idle for the rest of the duration of that mode. If we denote the MD of the traffic matrix $\mathbf{D}$ as $\mathcal{MD}\{\mathbf{D}\}$, then we can write

$$\mathcal{MD}\{\mathbf{D}\} = 5 \begin{array}{|c|c|c|c|}\hline \; & \; & 1 & \; \\ \hline \; & 1 & \; & \; \\ \hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \end{array} + 2 \begin{array}{|c|c|c|c|}\hline \; & 1 & \; & \; \\ \hline 1 & \; & \; & \; \\ \hline \; & \; & \; & 1 \\ \hline \; & \; & 1 & \; \\ \hline \end{array} + 4 \begin{array}{|c|c|c|c|}\hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \; & 1 & \; & \; \\ \hline \; & \; & 1 & \; \\ \hline \end{array}$$
$$+ \begin{array}{|c|c|c|c|}\hline \; & 1 & \; & \; \\ \hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \; & \; & 1 & \; \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & 1 & \; & \; \\ \hline \; & \; & 1 & \; \\ \hline 1 & \; & \; & \; \\ \hline \; & \; & \; & 1 \\ \hline \end{array} + \begin{array}{|c|c|c|c|}\hline \; & \; & 1 & \; \\ \hline \; & \; & \; & 1 \\ \hline 1 & \; & \; & \; \\ \hline \; & 1 & \; & \; \\ \hline \end{array} .$$

It can be seen that the SM matrices are integer multiples of permutation matrices. The durations of the modes in the above example are 5, 2, 4, 1, 1, and 1, respectively. The numbers not shown are 0.

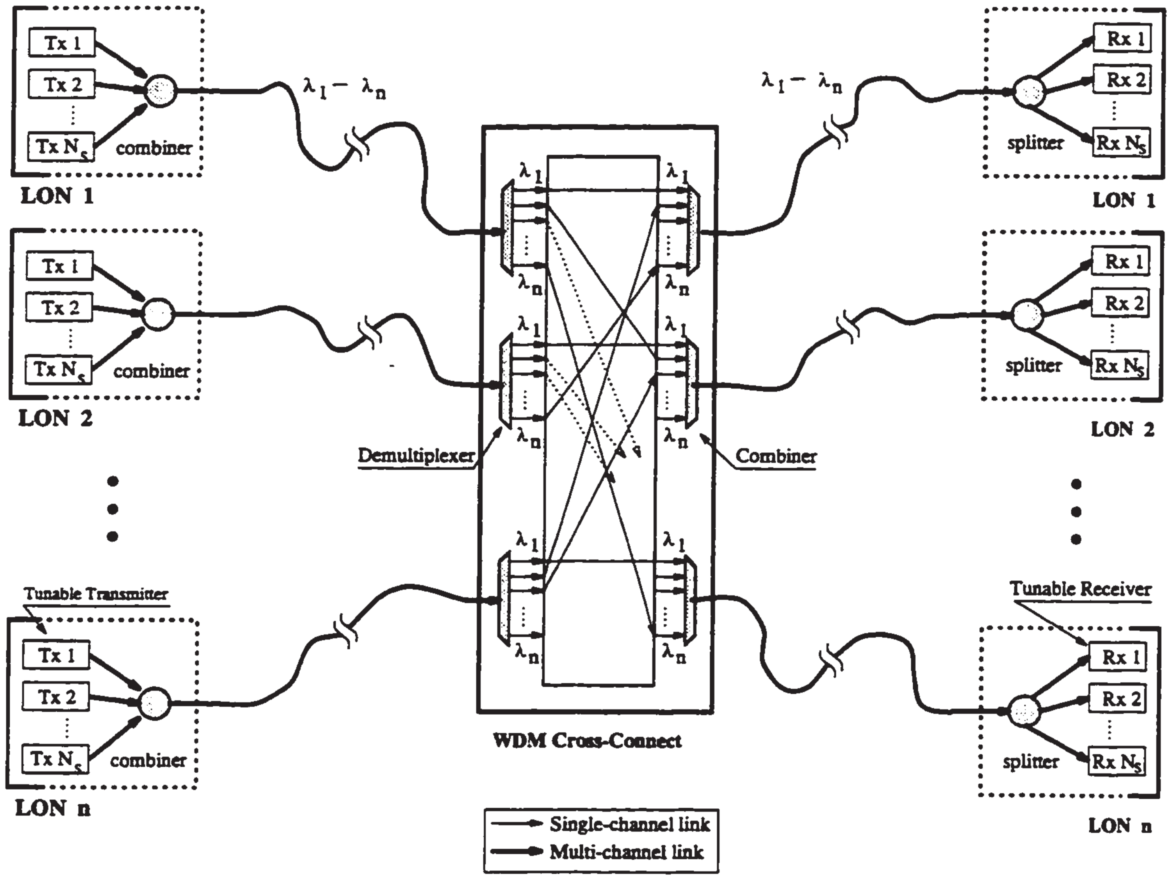


Figure 4.3: WDM cross connect network.

4.2.2 General Case: $N > n + 1$

Normally the number of stations in a network is much greater than the number of available channels. Therefore, $N_s = \lceil N/n \rceil$ stations have to use the same connection pattern as far as the WDM cross connect is concerned, since the routing is based only on wavelength. In order to reduce the component counts in the central cross connect, the network is clustered into local optical networks (LONs) with N_s stations in each. The intent is that sets of stations in close physical proximity would be configured as a LON during system setup. Inside each LON, all station transmit fibers are combined and applied to one input of a spatial WDM cross connect. The result system is shown in Figure 4.3, again unfolded.

Source LON	*Destination LON*					
	1	2	3	···	$n-1$	n
1	λ_1	λ_2	λ_3	···	λ_{n-1}	λ_n
2	λ_n	λ_1	λ_2	···	λ_{n-2}	λ_{n-1}
3	λ_{n-1}	λ_n	λ_1	···	λ_{n-3}	λ_{n-2}
···	···	···	···	···	···	···
$n-1$	λ_3	λ_4	λ_5	···	λ_1	λ_2
n	λ_2	λ_3	λ_4	···	λ_n	λ_1

Table 4.2: Wavelength Routing Connection Table for the "General Case".

Stations can dynamically wavelength route packets between LONs using the wavelength agility of their transmitters. Station transmitters are assumed to have fast wavelength agility across the entire set of wavelengths, $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Each channel is thus dedicated for inter-LON communications between two specific LONs. For example, a station in LON 1 would use λ_3 to send packets to all stations in LON 3. A routing table is shown in Table 4.2 which indicates how connections might be made inside the cross connect. Since the number of channels which can be reached by a given transmitter is given by n, and since the total number of destination channels is n^2 (considering spatial reuse), station receiver agility is clearly required to achieve full station to station connectivity. The alternative, of course, is to permit multiple receivers per station or to restrict the channels that are used. The destination station thus uses its receiver agility to select the packet destined to itself.

Wavelength routed-interleaved TDMA (WR-ITDMA) can be used in this case to schedule the transmission and reception of data packets among all the stations. The connection table for one frame ($N = nN_s = 9$ consecutive time slots) and a transmitter/receiver tuning time table for the case of $n = N_s = 3$ are shown in Figures 4.4 and 4.5, respectively.

In the general case of WR-ITDMA, there are n channels, N_s stations per LON and n LONs. If we let $n_m = max(n, N_s)$, the tuning wavelength of the transmitter at station i in time slot j, which will be denoted by Tx_i^j, is

$$Tx_i^j = \begin{cases} \lambda_t & \text{if } t \leq n, \\ idle & \text{otherwise,} \end{cases} \tag{4.1}$$

in which

$$t = (j-i) \bmod n_m + 1. \tag{4.2}$$

Similarly, the tuning wavelength of the receiver at station i in time slot j, denoted by Rx_i^j, can be shown to be

$$Rx_i^j = \begin{cases} \lambda_r & \text{if } r \leq n, \\ idle & \text{otherwise,} \end{cases} \tag{4.3}$$

where

$$r = (\lceil j/n_m \rceil - i) \bmod n_m + 1. \tag{4.4}$$

If we consider the connection table shown in Table 4.2 for the WDM cross connect, we can show that if station f in LON s wants to transmit to station g in LON d, it uses wavelength channel λ_k in time slot j, where

$$k = (d-s) \bmod n + 1, \tag{4.5}$$

and

$$j = n_m[(k+g-2) \bmod n_m] + (k+f-2) \bmod n_m + 1. \tag{4.6}$$

The tuning wavelength for this time slot can be easily verified from Equations 4.1 and 4.3. In the general case, it can be seen that the capacity of this system is given by $\min(N_s^2, n^2)$.

Each station has one assigned time slot in each frame on each of the wavelengths to insure full connectivity. This can also be seen in the example shown in Figures 4.4 and 4.5. In the case of non-uniform traffic, if there is no traffic going from one station to another station in a given frame, the corresponding time slot is wasted. We can define the efficiency of the scheduling as the ratio of the total traffic routed in each scheduling frame over the total frame length. The goal of dynamic scheduling is to maximize this performance measure.

As mentioned before, the connection time schedule for WR-ITDMA is fixed and it is not suitable for non-uniform or focused traffic. The idea in dynamic TSA is to change the connection time table according to the network traffic pattern on a longer frame basis. Gathering all station traffic requirements, each station (or equivalently,

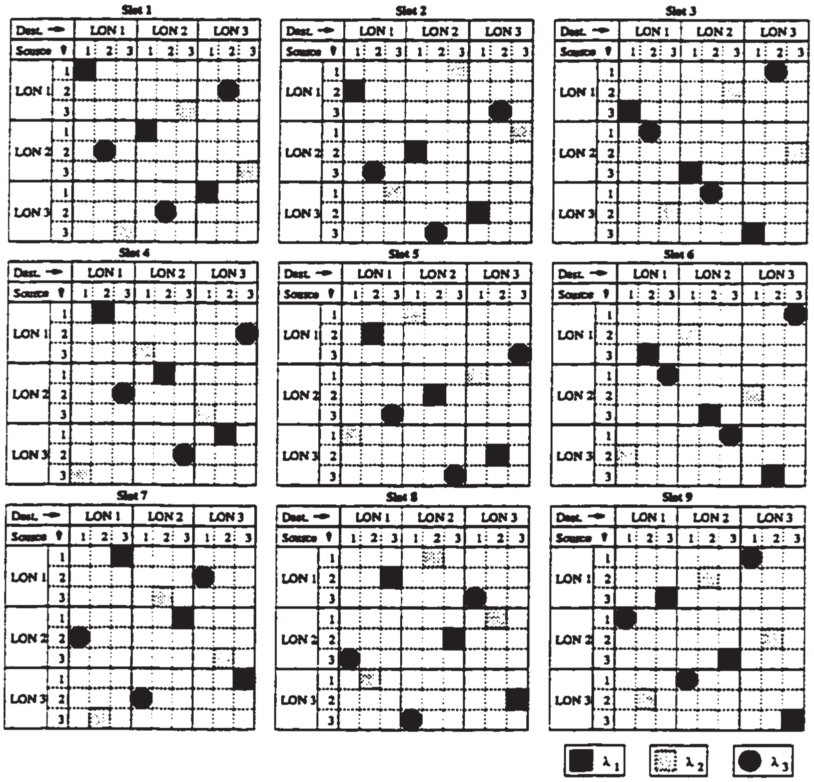


Figure 4.4: Connection time table for WR-ITDMA system.

a central controller) can form the traffic matrix. Each station (or central controller) then, will use the MD of the traffic matrix to tune its transmitter and receiver accordingly. In the case of a centralized control, the central controller must send the resulting MD to all the stations in the network.

4.3 Switching Modes of WDM Cross Connect Networks

By closer inspection of the scheduling table shown in Figure 4.4, we see that the SM of a WDM cross connect network, has the form shown below.

(a) RECEIVERS

Time Slot		1	2	3	4	5	6	7	8	9
Station	1	λ_1	λ_1	λ_1	λ_2	λ_2	λ_2	λ_3	λ_3	λ_3
	2	λ_3	λ_3	λ_3	λ_1	λ_1	λ_1	λ_2	λ_2	λ_2
	3	λ_2	λ_2	λ_2	λ_3	λ_3	λ_3	λ_1	λ_1	λ_1

(b) TRANSMITTERS

Time Slot		1	2	3	4	5	6	7	8	9
Station	1	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3
	2	λ_3	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3	λ_1	λ_2
	3	λ_2	λ_3	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3	λ_1

Figure 4.5: Tuning table of WR-ITDMA for receivers and transmitters of the stations in each LON.

$$\mathbf{M} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} & \cdots & \mathbf{S}_{1n} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} & \cdots & \mathbf{S}_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{S}_{n1} & \mathbf{S}_{n2} & \mathbf{S}_{n3} & \cdots & \mathbf{S}_{nn} \end{bmatrix} \tag{4.7}$$

M is the SM, which we will call the Wavelength cross(X) connect Switching Mode (WXSM) for easier reference. $\mathbf{S}_{ij}$ is an $N_s \times N_s$ matrix which has only one non-zero element that is equal to unity. We call this matrix a "singleton". A singleton can be represented completely by its size (N_s) and the location of its unit (non-zero) element, denoted by $\mathbf{S}^1_{ij} = (k, l)$, which means that in singleton $\mathbf{S}_{ij}$ the unit element is located in row k and column l ($i, j = 1, \cdots, n$; and $k, l = 1, \cdots, N_s$). In the SM of a WDM cross connect network, singletons in the same row (column) cannot have their non-zero elements in the same row (column). To write this formally, let

$$\mathbf{S}^1_{ij} = (r_1, c_1), \tag{4.8}$$

and

$$\mathbf{S}^1_{lm} = (r_2, c_2), \tag{4.9}$$

then,

$$\begin{aligned} i \neq l \ \ and \ \ j = m \ &\Longrightarrow\ c_1 \neq c_2, \\ i = l \ \ and \ \ j \neq m \ &\Longrightarrow\ r_1 \neq r_2, \end{aligned} \tag{4.10}$$

for $i, j, l, m = 1, 2, \cdots, n$. Otherwise, there are conflicts among a number of stations in access to the same transmitting or receiving channel.

In each $N_s \times N_s$ block of the traffic matrix represented by a singleton, only one connection can be scheduled. This is because each block corresponds to an inter-LON communication, and all the stations in one LON use one specific channel to connect to all stations in another LON. Considering the above results and definitions, we will present some methods to decompose the traffic matrix to its WXSMs. First we consider the following example.

Example 1: The following traffic matrix is given for a 2-channel WDM cross connect network with 2 stations per LON. By applying dynamic TSA, the matrix $\mathbf{D}$ is decomposed to WXSMs, as shown [1].

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 0 & 2 & 6 & 4 \\ \hline 4 & 0 & 1 & 5 \\ \hline\hline 4 & 3 & 0 & 2 \\ \hline 2 & 2 & 1 & 0 \\ \hline \end{array}$$

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline & & 5 & \\ \hline 4 & & & \\ \hline\hline & 3 & & \\ \hline & & & 0 \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & 2 & & \\ \hline & & & 5 \\ \hline\hline 4 & & & \\ \hline & & 1 & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & & 4 \\ \hline & 0 & & \\ \hline\hline & & 0 & \\ \hline 2 & & & \\ \hline \end{array}$$

$$+ \begin{array}{|c|c||c|c|} \hline & & 1 & \\ \hline & & & \\ \hline\hline 0 & & & 1 \\ \hline & 1 & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline 0 & & & \\ \hline & & 1 & \\ \hline\hline & & & 1 \\ \hline & 1 & & \\ \hline \end{array}$$

Finally, the MD of the traffic matrix is,

---

[1] Double lines are used to separate stations in different LONs. In the example, the first two rows and columns correspond to the stations in LON 1 and the third and fourth rows and columns correspond to the stations in LON 2.

	LON	1			⋯	n		
LON	Station	1	⋯	N_s	⋯	1	⋯	N_s
1	1							
	⋮							
	N_s							
⋮	⋮							
n	1							
	⋮							
	N_s							

Table 4.3: General format of the traffic matrix for the WDM cross connect network.

$$\mathcal{MD}\{\mathbf{D}\} = 5\begin{array}{|c|c||c|c|}\hline & & 1 & \\ \hline 1 & & & \\ \hline\hline & 1 & & \\ \hline & & & 1 \\ \hline\end{array} + 5\begin{array}{|c|c||c|c|}\hline & 1 & & \\ \hline & & & 1 \\ \hline\hline 1 & & & \\ \hline & & 1 & \\ \hline\end{array} + 4\begin{array}{|c|c||c|c|}\hline & & & 1 \\ \hline & 1 & & \\ \hline\hline & & 1 & \\ \hline 1 & & & \\ \hline\end{array}$$

$$+\begin{array}{|c|c||c|c|}\hline & & 1 & \\ \hline & & & \\ \hline\hline 1 & & & 1 \\ \hline & 1 & & \\ \hline\end{array} + \begin{array}{|c|c||c|c|}\hline 1 & & & \\ \hline & & 1 & \\ \hline\hline & & & 1 \\ \hline & 1 & & \\ \hline\end{array}.$$

4.4 Decomposition into Switching Modes

In general we can write each traffic matrix $\mathbf{D} = [d_{ij}]$, where $i, j = 1, 2, \cdots, nN_s$, as

$$\mathbf{D} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} & \cdots & \mathbf{B}_{1n} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} & \cdots & \mathbf{B}_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{B}_{n1} & \mathbf{B}_{n2} & \mathbf{B}_{n3} & \cdots & \mathbf{B}_{nn} \end{bmatrix}, \tag{4.11}$$

where $\mathbf{B}_{ij} = [b_{kl}]$, ($i, j = 1, \cdots, n$; and $k, l = 1, \cdots, N_s$), is an $N_s \times N_s$ submatrix (block). Implicitly, we assume that the traffic matrix of a WDM cross connect network is formed according to the general format shown in Table 4.3.

In turn, each element b_{kl} of a submatrix $\mathbf{B}_{ij}$, denoted by b^{ij}_{kl}, can be shown to be

$$b^{ij}_{kl} = d_{(i-1)N_s+k,\ (j-1)N_s+l} \ . \tag{4.12}$$

For the traffic matrix $\mathbf{D} = [d_{ij}]$, we define S^R_i, S^C_j and S^B_{kl} as the row i sum, column j sum and block kl sum, respectively. We also refer to them as a "line sum". These sums can be written as

$$S^R_i = \sum_{j=1}^{nN_s} d_{ij}, \tag{4.13}$$

$$S^C_j = \sum_{i=1}^{nN_s} d_{ij}, \tag{4.14}$$

and

$$S^B_{kl} = \sum_{i=(k-1)N_s+1}^{kN_s} \ \sum_{j=(l-1)N_s+1}^{lN_s} d_{ij}. \tag{4.15}$$

4.4.1 SS/TDMA and Trivial Case of WDM Cross Connect Networks

A matrix with all the row and column sums equal to 1 is called a "doubly stochastic" matrix. A matrix with all equal row and column sums (T_s), but different from 1, can be written as a multiple of a doubly stochastic matrix, simply by factoring T_s out of the matrix. This matrix is called "quasi-doubly stochastic", QDS for short.

In the following, we first quote some of the known characteristics of the MD for SS/TDMA that can be directly applied to the trivial case of WDM cross connect networks, and then we will propose extensions to the general case of WDM cross connect networks.

Theorem 1 *Given any $N \times N$ traffic matrix $\mathbf{D}$ with total traffic T and maximum row and column sum of T_m, ($\max\{S^R_i, S^C_j\} = T_m;\ i,j = 1,\cdots,N$), it is always possible to add appropriate nonnegative integers to the entries of $\mathbf{D}$ so as to obtain a new $N \times N$ traffic matrix $\mathbf{D_q}$ which is QDS with row and column sum of T_m.*

Proof: Refer to [IUMY77], [Inu79].

Theorem 2 *Let* $\mathbf{D}$ *be an* $N \times N$ *QDS matrix. Then* $\mathbf{D}$ *is a sum of integer multiples of permutation matrices.*

Proof: Refer to [Hal67], [Bal95].

Theorem 3 *The necessary and sufficient time to communicate all of the traffic in the traffic matrix with maximum row and column sum of* T_m*,* $(\max\{S_i^R, S_j^C\} = T_m;\ i,j = 1,\cdots,N)$*, in an SS/TDMA system with* N *transponders, or for the trivial case of a WDM cross connect network with* $N-1$ *channels is* T_m*.*

Proof: Refer to [IUMY77], [Hal67].

Theorem 4 *The* $N \times N$ *traffic matrix* $\mathbf{D}$ *of the SS/TDMA system with* N *transponders or that of the trivial case of a WDM cross connect network with* $N-1$ *channels is always decomposable to at most* N *switching modes.*

Proof: An ITDMA standard scheduling scheme can cover a traffic matrix with N SMs. Since each SM can include only one element of each row or column, then for the general case of a traffic matrix containing non-zero elements, we need exactly N modes. The duration of each mode is equal to the maximum element of that mode matrix. The duration of one or more modes might be zero for some of traffic matrices. Consequently, in these cases the number of SMs is less than N.

Theorem 5 *The minimum number of switching modes for decomposition of an* $N \times N$ *traffic matrix* $\mathbf{D}$ *of the SS/TDMA system with* N *transponders or that of the trivial case of a WDM cross connect network with* $N-1$ *channels is* $\max\{\max_i\{\alpha_i\}, \max_j\{\beta_j\}\}$*, where* α_i *is the number of non-zero entries on row* i*,* β_j *is the number of non-zero entries on column* j*,* $(i,j = 1,\cdots,N)$*.*

Proof: Each SM can contain one element from each row or column. Therefore, the minimum number of modes is equal to the maximum number of non-zero elements in all the rows or columns.

From Theorems 1-5 we can conclude that it is always possible to decompose a traffic matrix to its SMs. A given traffic matrix can always be converted to a QDS matrix by adding some dummy traffic (Theorem 1), which according to Theorem 2 can be written as a sum of integer multiples of permutation matrices. Finally, we can apply scheduling methods to minimize either the transmission duration according to Theorem 3 or the number of switching modes according to Theorem 5. Theorem 4, on the other hand, shows that ITDMA can be used as a basis for a fast and fixed scheduling method. This is done by extension of the duration of each basis SM of ITDMA to the length of the maximum traffic demand in that mode.

4.4.2 WDM Cross Connect Network

The following theorems and results are applicable to the general case of WDM cross connect networks.

Theorem 6 *Let* **D** *be a traffic matrix for a WDM cross connect network. Then* **D** *can be decomposed to a sum of integer multiples of WXSM matrices.*

Proof: We show a method that can always decompose the traffic matrix of a WDM cross connect network to its WXSMs. Although the proof presented is not using a rigorous method, it gives us some insight into the problem and it is useful later. Referring to WR-ITDMA (see Equations 4.1, 4.3, and 4.6) we can form standard WXSMs that always cover all the elements of a traffic matrix. These are WXSMs that can systematically be obtained regardless of the traffic matrix. This method is not efficient in finding the optimum scheduling. However, it shows that a given traffic matrix can always be written as a sum of integer multiples of WXSM matrices. Each mode will be in effect for the duration of the maximum traffic demand in that mode. For more detail refer to the WR-ITDMA method and the example in Section 4.5.3.

Theorem 7 *The necessary (but not always sufficient) time to communicate all of the traffic in the traffic matrix with maximum line sum of T_s, $(\max\{S_i^R, S_j^C, S_{kl}^B\} = T_s;\ i,j = 1,\cdots,nN_s,\ \ k,l = 1,\cdots,n)$, in a WDM cross connect network with n channels and N_s stations per LON is T_s.*

Proof: The elements in the same row of the traffic matrix are sourced by a common station (transmitter), those in the same column are destined to the same station (receiver), and the elements in the same block have to be sent through the same channel. It is obvious that all the elements in the same line, i.e. row, column or block, have to be sent serially, since the common resource they are sharing can only serve one traffic entity at a time. Consequently, the line (row/column/block) sums show the minimum time duration needed to schedule those traffic elements from/to/onto the same source/destination/channel. As a result, the minimum necessary time to communicate all of the traffic in the traffic matrix is the maximum of all the line sums, which is T_s.

As mentioned in Theorem 7, T_s is the lower bound for the length of the MD of a traffic matrix in a WDM cross connect network. Unlike the SS/TDMA case, there are some instances where we cannot achieve the lower bound. As an example consider the matrix

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ \hline\hline 1 & 0 & 0 & 1 \\ \hline 0 & 1 & 1 & 0 \\ \hline \end{array}\,.$$

Although all line sums are 2, the minimum time duration of the MD of this traffic matrix is 3, which is shown below.

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline & 1 & & \\ \hline & & 1 & \\ \hline\hline & & & 1 \\ \hline 0 & & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & & 1 \\ \hline 1 & & & \\ \hline\hline & & 0 & \\ \hline & 1 & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & & 0 \\ \hline & 0 & & \\ \hline\hline 1 & & & \\ \hline & & 1 & \\ \hline \end{array}$$

Theorem 8 *The traffic matrix* $\mathbf{D}$ *of a WDM cross connect network with* n *channels and* N_s *stations in each LON can always be decomposed to sum of at most* nN_s *switching modes.*

Proof: Again we use the WR-ITDMA method. We know that we can cover a given traffic matrix with nN_s switching modes that can be found through Equations 4.1, 4.3, and 4.6. Each mode duration is the same as the maximum element in that mode. When there are instances where the durations of some of the modes are 0, then the number of WXSMs would be less than nN_s.

Theorem 9 *The lower bound for the the number of WXSMs in a WDM cross connect network with* n *channels and* N_s *stations per LON is* $\max\{\max_i\{\alpha_i\}, \max_j\{\beta_j\}, \max_{kl}\{\gamma_{kl}\}\}$, *where* α_i *is the number of non-zero elements in row* i $(i = 1, \cdots, nN_s)$, β_j *is the number of non-zero elements in column* j $(j = 1, \cdots, nN_s)$, *and* γ_{kl} *is the number of non-zero elements in block* kl $(k, l = 1, \cdots, n)$.

Proof: Since the elements of the same row are sharing the same transmitter and the elements in the same column have the same destination, they have to be sent serially, i.e. in different switching modes. Similarly, the elements of the same block share one wavelength channel, and they have to be sent in different modes. Thus, we need at least as many modes as the maximum number of non-zero elements in each of these sets. Unfortunately, it is not always the minimum number and similar to Theorem 7, is just a lower bound. The example shown in the proof of Theorem 7 can also be used in this case. In the example, the lower bound is 2, but the minimum achievable is 3 WXSMs.

Theorem 6 shows that we can always decompose a traffic matrix of a WDM cross connect network to its WXSMs. The lower bound of the transmission duration can be found using Theorem 7, while Theorem 9 shows the minimum number of switching modes. Theorem 8, on the other hand, proposes a fixed fast method for the decomposition of the traffic matrices of the WDM cross connect network. This method, however, does not consider an optimal transmission time or number of switching modes.

4.5 Time Slot Assignment Algorithms

The TSA by using the MD of a traffic matrix **D** will result in the decomposition of **D** into N_m switching matrices. The objective function is to minimize the overall transmission duration T_{total} given by

$$T_{total} = \tau N_m + T_D, \tag{4.16}$$

where N_m is the number of switching matrices, τ is the overhead time spent per switching and T_D is the time spent in transmission of the traffic. T_D in turn can be written in terms of the durations of the switching modes. If we decompose the traffic matrix **D** to its switching modes M_i as

$$\mathbf{D} = \sum_{i=1}^{N_m} \mathbf{M}_i, \tag{4.17}$$

then we can write T_D as

$$T_D = \sum_{i=1}^{N_m} |\mathbf{M}_i|, \tag{4.18}$$

where $|\mathbf{M}_i|$ denotes the maximum entry in $\mathbf{M}_i$, or equivalently the duration of $\mathbf{M}_i$.

In [GW85, Pr88, Gan92], it has been shown that the problem given in Equation 4.16 is computationally intractable (NP-complete) for the case of SS/TDMA systems. In the WDM cross connect network, a more general case exists. As we saw before, a specific case of the WDM cross connect networks, i.e. the trivial case, is the same as SS/TDMA. Consequently, the problem of TSA for general WDM cross connect networks is much more complicated than that for SS/TDMA, and is also NP-complete. As in SS/TDMA, the problem is reduced to the two following cases, depending on the relative magnitude of τ with respect to the duration of a time slot:

1. **minimal transmission duration:** minimize N_m subject to the constraint that T_D is minimum,
2. **minimal number of switching matrices:** minimize T_D subject to the constraint that N_m is minimum. This case has also been shown to be an NP-complete problem [GW85].

Heuristics are needed to find suboptimal solutions for each case. Many of the heuristics proposed for the MD of the original SS/TDMA use bipartite matching or the "marriage problem", since each SS/TDMA (or the trivial case of WDM cross connect networks) can be represented by a bipartite graph [Bal95, Hal67, Law76]. Unfortunately, the graph representation of the general case is not a simple bipartite graph and many of the previously proposed algorithms are not applicable. The matching problem of the WDM cross connect network is similar to the generalized marriage problem or the *3-dimensional matching* problem. In this problem, the sets W, X, and Y correspond to *three* different sexes, and each triple in the matching correspond to a *3-way marriage* that would be acceptable to all three participants. We can easily see that we can assume W as the set of available channels (considering spatial reuse for a network with n channels, there are n^2 channels), which correspond to the blocks in the traffic matrix. X can be assumed to be the set of the transmitters or sources, which correspond to the rows of the traffic matrix. Finally, we can assume Y as the set of receivers or sinks, which correspond to the columns of the traffic matrix. Therefore, the MD of the traffic matrix is the same as 3-dimensional matching (3-way marriage problem). 3-dimensional matching is shown to be an *NP-complete* problem, refer to [GJ79] page 50. Thus, the TSA problem in the case of the WDM cross connect network is a much more complicated problem than that of SS/TDMA. Because of the NP-completeness of 3-dimensional matching, a direct application of many of the heuristics proposed for the SS/TDMA system to the new problem is not possible.

In the following we propose three TSA heuristics. We assume that there are n channels and N_s stations per LON, so that the total number of stations is $N = nN_s$.

The proposed TSA methods are also selected to be fast. It should be noted that even though very fast computers with plenty of memory space exist, it is not reasonable to waste their power on exhaustive searches which result in marginal enhancement of the TSA. On the other hand, if distributed control is used in the network, a fast algorithm can be used on all the network stations with different speeds, while a complex method is feasible only for very fast and powerful stations.

A general assignment algorithm includes an iterative application of the following matrix decomposition step.

$$\mathbf{D}^{(k)} = \mathbf{M}^{(k)} + \mathbf{D}^{(k+1)}, \quad k = 1, 2, 3, \cdots, \tag{4.19}$$

where $\mathbf{D}^{(k)}$ and $\mathbf{D}^{(k+1)}$ are the traffic matrices in the k-th and $(k+1)$-st iteration, and $\mathbf{M}$ is the WXSM matrix in the k-th iteration. The iteration ends when $\mathbf{D}^{(k+1)} = 0$.

4.5.1 Minimizing N_m subject to minimum T_D

The assignment scheme presented here is a modified version of a simple computational method known as the "greedy algorithm" [IUMY77], which we call the "modified greedy algorithm". In the greedy algorithm, there are some instances that the algorithm stays in a loop forever ("deadlock") [Inu79]. In the modified greedy algorithm presented below, this problem is fixed, while the fast greedy algorithm is in effect most of the time. This method is also modified for the case of cross connect networks in which line sums are row, column and block sums. This procedure tries to achieve an assignment plan with minimum T_D, while incorporating heuristics that tend to minimize the number of switching matrices N_m. We assume that the transmitter and receiver tuning times are much shorter than a time slot.

Modified Greedy Algorithm for minimum transmission duration time (Method 1)

The modified greedy algorithm consists of the following steps.

Step 1: Find the line (row, column or block) that has the maximum traffic in a given matrix $\mathbf{D}$, which is an $N \times N$ matrix. This line is called the "critical" line, l_c, with line sum S_c.

Step 2: Find the maximum element, "critical" element, in the critical line. Call it d_c. Check for the modification flag set by step 9.2. If it is set, then find the second (third, etc.) maximum element instead.

Step 3: Remove the lines (row, column and block) which contain the critical element. Find the maximum element in the residual elements. Extract at most d_c units of traffic from this element.

Step 4: Repeat the same procedure as Step 3 ($n^2 - 1$ times), again eliminating the lines containing the elements selected in the previous steps.

Step 5: Extract n^2 numbers found in the former steps, thereby dividing $\mathbf{D}$ into two parts $\mathbf{M}$ and $\mathbf{C}$, so that $\mathbf{D} = \mathbf{M} + \mathbf{C}$. $\mathbf{M}$ is the WXSM matrix found from the previous steps.

Step 6: If the line l_c of the matrix $\mathbf{C}$ is also a critical line, go to Step 8. Otherwise, continue.

Step 7: If the line l_c of $\dot{\mathbf{C}}$ with new line sum S'_c is not critical, and another line with line sum S_1 is critical, put $S_1 - S'_c$ units of the traffic from critical element of $\mathbf{M}$ back to $\mathbf{C}$, so as to maintain the l_c critical, and update the other elements of $\mathbf{M}$ accordingly.

Step 8: If $\mathbf{M}$ is non-zero, go to step 10. Otherwise continue.

Step 9: In this case matrix $\mathbf{C}$ is the same as $\mathbf{D}$. This happens only when there are more than one critical lines. For the case of having more than one critical line, if we do not select any element from one of the critical lines at all, we have to return the entire traffic selected on l_c back to the original matrix. In order to overcome this problem, we must select one element from each critical line in each of the decomposition stages. Marking all the critical lines, we start by finding a system of distinct representatives (SDR)[2] such that each critical line has one element in this set. It is obvious that some of the elements can be in the intersection of two or three critical sets. For example, if the row i and column j are critical, then element $d_{ij} (\neq 0)$ is on two critical lines. Let us call the minimum of these critical elements $d_{c(min)}$. Now we can extract $d_{c(min)}$ traffic units from those critical positions, and decompose the matrix to $\mathbf{M}$ and $\mathbf{C}$ as before. We can minimize the number of switching matrices by maximizing $d_{c(min)}$. If we implement an algorithm that finds the SDR, we can enhance it by searching for the set with the maximum possible $d_{c(min)}$. Unfortunately, this algorithm is not as simple as the one for SS/TDMA used in [Inu78]. Therefore, we use a simpler but

---

[2] Let $S_1, S_2, \cdots, S_n$ be subsets of a finite set E. If each subset S_i contains a distinct element x_i so that $x_i \neq x_j$ for $i \neq j$, then $x_1, x_2, \cdots, x_n$ are called a ***system of distinct representatives*** (SDR) of $S_1, S_2, \cdots, S_n$. (In SS/TDMA or the trivial case of WDM cross connect networks, E and S_i may be interpreted as the column numbers and the i-th row set of an $n \times n$ traffic matrix. In the MD of WDM cross connect networks we are searching for distinct representatives of each row so that they are different in column numbers and block numbers at the same time.)

more exhaustive search to find the SDR. First, we call the very first critical line as l_{c0}.

Step 9.1: Replace the critical line with the next critical line. If the critical line is different from l_{c0}, go to step 10. Otherwise continue.

Step 9.2: Set the modification flag, which denotes that we need to select the second (third, and so on) maximum as the critical element in Step 2. If all the elements are tested before, it means that the lower bound cannot be reached. Add one unit of dummy traffic to the critical element and continue.

Step 10: Replace **D** with **C**, and repeat all the steps from Step 1. This process ends when $\mathbf{C} = 0$.

Example 2: A traffic matrix is given as

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 5 & 4 & 4 & 2 \\ \hline 6 & 1 & 3 & 2 \\ \hline\hline 7 & 1 & 5 & 1 \\ \hline 1 & 0 & 3 & 2 \\ \hline \end{array}$$

for a WDM cross connect network with 2 channels and 2 stations per LON. In this example, we assumed that each station is sending data to itself as well. Using the above method, we can decompose this traffic matrix as

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline & 4 & & \\ \hline & & 3 & \\ \hline\hline 7 & & & \\ \hline & & & 2 \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & & 2 \\ \hline 3 & & & \\ \hline\hline & & 3 & \\ \hline & 0 & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline 3 & & & \\ \hline & & & 2 \\ \hline\hline & 1 & & \\ \hline & & 3 & \\ \hline \end{array}$$

$$+ \begin{array}{|c|c||c|c|} \hline & & 3 & \\ \hline 3 & & & \\ \hline\hline & & & 1 \\ \hline & 0 & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline 2 & & & \\ \hline & & & 0 \\ \hline\hline & & 2 & \\ \hline & 0 & & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & 1 & \\ \hline & 1 & & \\ \hline\hline & & & 0 \\ \hline 1 & & & \\ \hline \end{array}.$$

In this example, the number of WXSMs is 6 ($N_m = 6$), while the minimum number of switching modes (the lower bound) is 4, and $T_D = 7+3+3+3+2+1 = 19$, which is equal to the lower bound.

4.5.2 Minimizing T_D subject to minimum N_m

If the tuning time of the transmitters and receivers is large compared to each time slot in a frame, then the dominant parameter in the transmission time in Equation 4.16 is τN_m. In this case, we maintain N_m or the number of switchings as the minimum value (close to the lower bound) and minimize the time spent in traffic transmission or T_D. As Theorem 9 indicates, $N_{m,min}$ is greater than or equal to the maximum number of non-zero elements in all the lines (rows, columns, and blocks).

The modified greedy algorithm in this case searches for the maximum elements all the time. Simulations show that the method performs very well and fast. The number of switching modes is also close to the lower bound. Note that this case is an NP-complete problem.

Modified Greedy Algorithm for minimum number of switchings (Method 2)

Step 1: Extract the maximum element in the traffic matrix.

Step 2: Extract the next maximum element in the residual matrix found by removing row, column and block elements of the previous maximum element. Continue the same step for $(n^2 - 1)$ times.

Step 3: Now we can decompose the traffic matrix as $\mathbf{D} = \mathbf{M} + \mathbf{C}$, where $\mathbf{M}$ is the WXSM formed by the elements extracted from the traffic matrix in the previous steps.

Step 4: Replace $\mathbf{D}$ with $\mathbf{C}$, and repeat all the steps from Step 1. This process ends when $\mathbf{C} = 0$.

The algorithm tries to achieve the minimum transmission duration by accommodating the connections with (almost) the same durations in one frame as much as possible. This method, however, is not as sophisticated and efficient as the previous one. It is a fast algorithm that can give a good number of switching modes close to the lower bound. An example of this simple algorithm is shown below.

Example 3: The same traffic matrix as in the previous example is given, i.e.

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 5 & 4 & 4 & 2 \\ \hline 6 & 1 & 3 & 2 \\ \hline\hline 7 & 1 & 5 & 1 \\ \hline 1 & 0 & 3 & 2 \\ \hline \end{array} .$$

Using the above procedure, we can decompose $\mathbf{D}$ to its WXSMs as

$$\mathbf{D} \; = \begin{array}{|c|c||c|c|} \hline & 4 & & \\ \hline & & 3 & \\ \hline\hline 7 & & & \\ \hline & & & 2 \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & & 2 \\ \hline 6 & & & \\ \hline\hline & & 5 & \\ \hline & 0 & & \\ \hline \end{array}$$

$$+ \begin{array}{|c|c||c|c|} \hline 5 & & & \\ \hline & & & 2 \\ \hline\hline & 1 & & \\ \hline & & 3 & \\ \hline \end{array} + \begin{array}{|c|c||c|c|} \hline & & 4 & \\ \hline & 1 & & \\ \hline\hline & & & 1 \\ \hline 1 & & & \\ \hline \end{array} .$$

In this example, the number of WXSMs $N_m = 4$, that is equal to the minimum number of switching modes and $T_D = 7 + 6 + 5 + 4 = 22$, while the lower bound for T_D is 19.

4.5.3 Fixed Number of Switchings (Method 3)

As we have seen in Theorem 8, it is always possible to decompose a traffic matrix of a WDM cross connect network with n channels and N_s stations per LON to at most nN_s switching modes. The fixed scheduling of WR-ITDMA is used in this case.

Example 4: Assume the same traffic matrix as in the previous examples.

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 5 & 4 & 4 & 2 \\ \hline 6 & 1 & 3 & 2 \\ \hline\hline 7 & 1 & 5 & 1 \\ \hline 1 & 0 & 3 & 2 \\ \hline \end{array} .$$

Using the WXSMs of WR-ITDMA, we can decompose the traffic matrix to the WXSMs as follows.

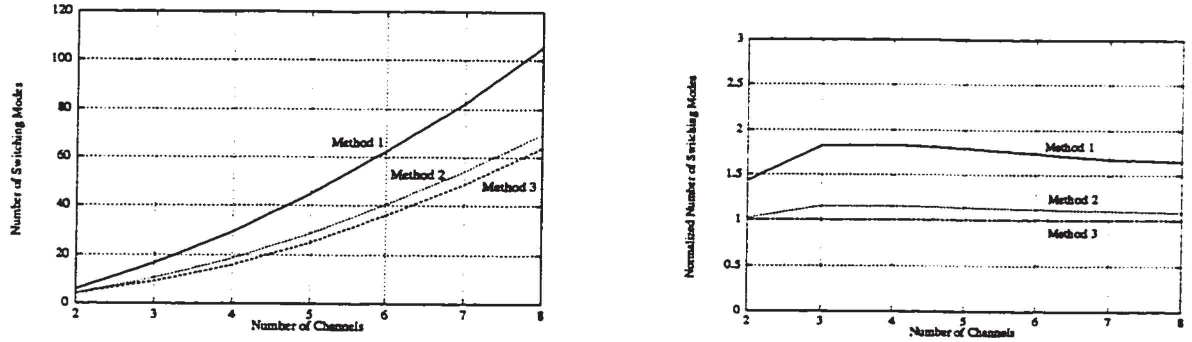
$$\mathcal{MD}\{\mathbf{D}\} = 5\begin{array}{|c|c||c|c|}\hline 1 & & & \\ \hline & & & 1 \\ \hline\hline & & 1 & \\ \hline & 1 & & \\ \hline\end{array} + 6\begin{array}{|c|c||c|c|}\hline & & & 1 \\ \hline 1 & & & \\ \hline\hline & 1 & & \\ \hline & & 1 & \\ \hline\end{array}$$
$$+ 4\begin{array}{|c|c||c|c|}\hline & 1 & & \\ \hline & & 1 & \\ \hline\hline & & & 1 \\ \hline 1 & & & \\ \hline\end{array} + 7\begin{array}{|c|c||c|c|}\hline & & 1 & \\ \hline & 1 & & \\ \hline\hline 1 & & & \\ \hline & & & 1 \\ \hline\end{array}$$

As shown, we decompose the traffic matrix to 4 WXSMs with total transmission duration of $5 + 6 + 4 + 7 = 22$. The lower bound for T_D is 19 and the lower bound for N_m is 4.

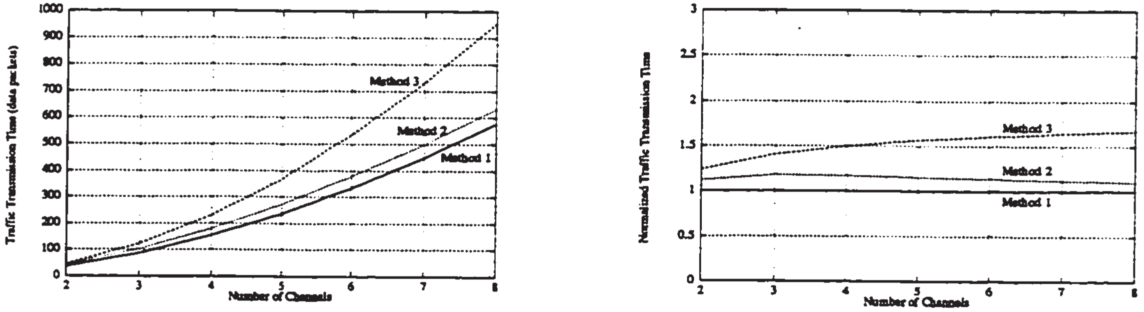
In this scheme, the WXSMs are fixed and they always cover the entire traffic matrix. However, there is no attempt to minimize the transmission duration.

4.6 Efficiency of Time Slot Assignment Algorithm

The TSA efficiency is investigated in this section using computer simulation. In the simulations, the traffic matrices are produced randomly, where elements of the matrices are uniformly distributed between 0 and a maximum number. The above three scheduling methods are then used to schedule the given traffic matrices. This procedure is repeated for a large number of random matrices (around 100,000) and the average values for the number of switching modes (N_m), the traffic transmission time (T_D) and the total transmission time (T_{total}) are found for 5 different values of switching penalties. The switching penalties are assumed to be from 0 (negligible penalty) to a fairly large value of 4 data packet long. The results are shown in Figures 4.6 to 4.8. In these figures, Method 1 refers to the modified greedy algorithm for minimum transmission duration described in Section 4.5.1. Method 2 is the modified



(a) N_m (b) N_m/N



(c) T_D (d) T_D/T_s

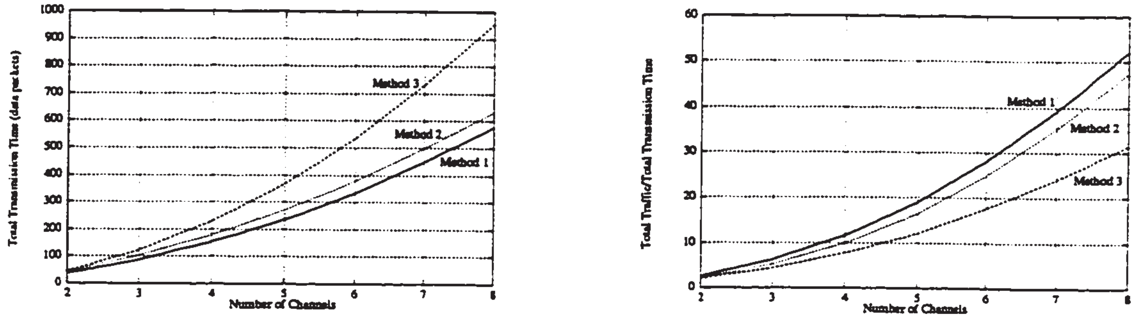
Method 1: Modified Greedy Algorithm for minimum transmission duration time.
Method 2: Modified Greedy Algorithm for minimum number of switchings.
Method 3: Fixed number of switchings.

Figure 4.6: Comparison of the different TSAs. (a) Number of switching modes. (b) Optimality of the number of switching modes. (c) Traffic transmission time. (d) Optimality of the traffic transmission time.

greedy algorithm for minimum number of switchings discussed in Section 4.5.2. Finally, Method 3 shows the results of the fixed number of switching method explained in Section 4.5.3.

In the simulations, WDM cross connect networks are assumed to have the same number of stations in each LON as the number of available channels, i.e. $n = N_s$. Consequently, total number of stations in the network is $N = nN_s = n^2 = N_s^2$.

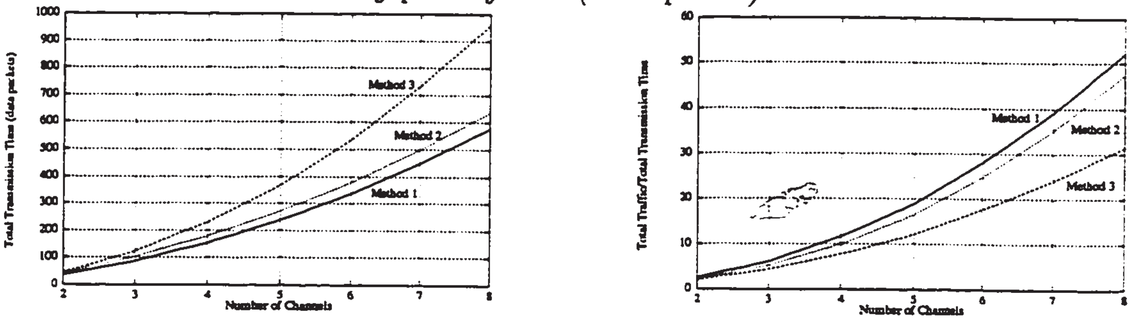
As shown in Figure 4.6(a) and (b), Method 3 is the best in terms of the number of switching modes (N_m), since it always gives the number of switching modes very close to the lower bound. Method 1, on the other hand, is the best in terms of the traffic transmission time (T_D), since it achieves the closest transmission time to the lower



(a) Total transmission time
T_{total}

(b) Performance Efficiency
$Total\ Traffic/T_{total}$

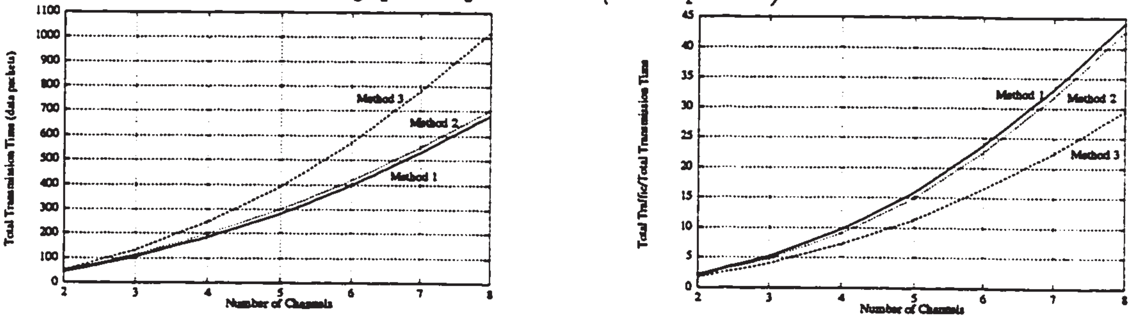
Switching penalty = 0 (data packet).



(c) Total transmission time
T_{total}

(d) Performance Efficiency
$Total\ Traffic/T_{total}$

Switching penalty = 0.01 (data packet).



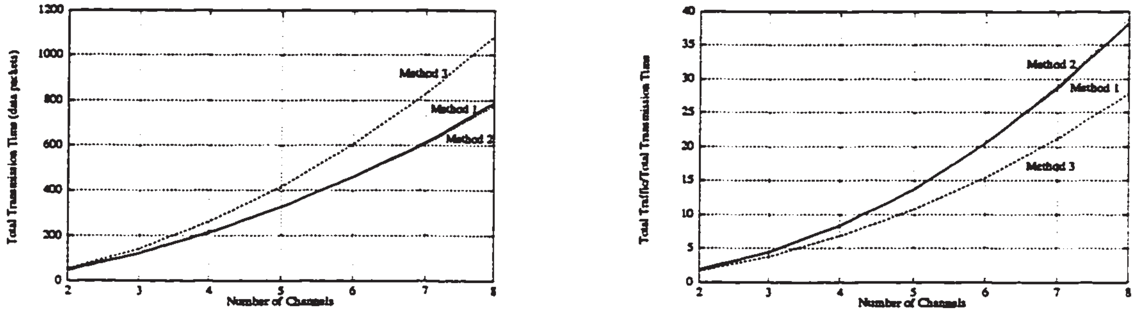
(e) Total transmission time
T_{total}

(f) Performance Efficiency
$Total\ Traffic/T_{total}$

Switching penalty = 1 (data packet).

Method 1: Modified Greedy Algorithm for minimum transmission duration time.
Method 2: Modified Greedy Algorithm for minimum number of switchings.
Method 3: Fixed number of switchings.

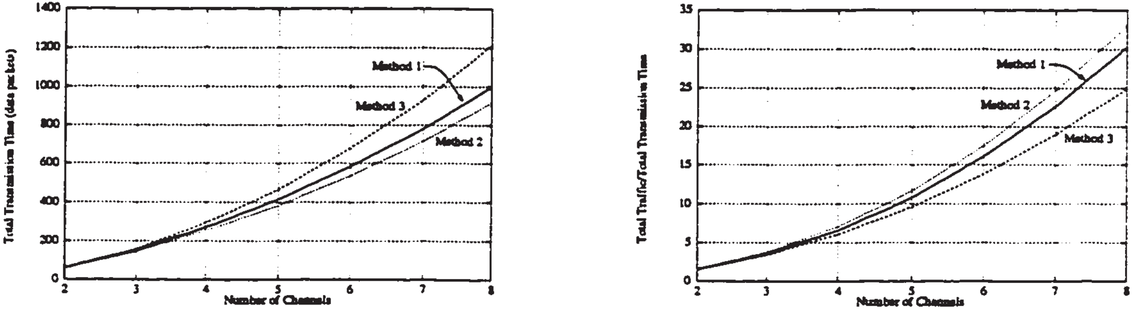
Figure 4.7: Total transmission time for different TSAs for switching penalties of 0, 0.01, and 1 data packet.



(a) Total transmission time T_{total}

(b) Performance Efficiency $Total\ Traffic/T_{total}$

Switching penalty = 2 (data packets).



(c) Total transmission time T_{total}

(d) Performance Efficiency $Total\ Traffic/T_{total}$

Switching penalty = 4 (data packets).

Method 1: Modified Greedy Algorithm for minimum transmission duration time.
Method 2: Modified Greedy Algorithm for minimum number of switchings.
Method 3: Fixed number of switchings.

Figure 4.8: Total transmission time for different TSAs for switching penalties of 2 and 4 data packets.

bound, as shown in Figure 4.6(c) and (d). Method 2 has a performance between Method 1 and 3 in terms of both the traffic transmission time and the number of switching modes. In Figure 4.6(b), the number of switching modes is normalized to the total number of stations N (i.e. N_m/N). Figure 4.6(d) shows the traffic transmission time normalized to the corresponding lower bound, i.e. T_D/T_s. The simulations also show that Methods 2 and 3 are much faster than Method 1.

In Figures 4.7 and 4.8, the efficiencies of the three methods are compared in terms of the total transmission time (T_{total}) where switching penalties are also included. The results are shown for τ = 0, 0.01, 1, 2, 4 (data packets). In right hand part of all these graphs the ratio of average total traffic to the total transmission time is shown. It is obvious that the bigger ratios represent the more efficient methods. We can easily observe that the efficiency of the first method is the best in most of the cases, except for large switching penalties where method 2 is better.

4.7 Distributed Tuning

The tuning operation in a WDM cross connect network is done in a distributed manner at the stations, unlike SS/TDMA which performs the tuning at the switching module located on the satellite. In all the scheduling methods that we have described so far, we assumed that all the stations perform the retuning before the start of each mode of transmission and reception simultaneously, i.e. at the time of mode switching. This kind of synchronized retuning might be the best solution, if there is a central switch as in SS/TDMA. In the case of distributed switching, the question might be asked "Is it possible to optimize the scheduling by pipelining the tuning times and data transmissions?" Through some examples we can easily show that the answer is yes, although in general this optimization is very difficult and in many cases it does not appear to be worthwhile. Particularly, if the transmitter and receiver tuning times are much shorter than the duration of a data time slot, this scheme hardly provides any improvement in the overall transmission time.

In the following example we show the tuning times as the numbers in parentheses. In each case, the corresponding transmitter and/or receiver must be retuned to another channel for the next mode of data communication.

Example 5: The following traffic matrix is given.

$$\mathbf{D} = \begin{array}{|c|c||c|c|} \hline 0 & 6 & 0 & 0 \\ \hline 0 & 0 & 1 & 2 \\ \hline\hline 2 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 2 \\ \hline \end{array}$$

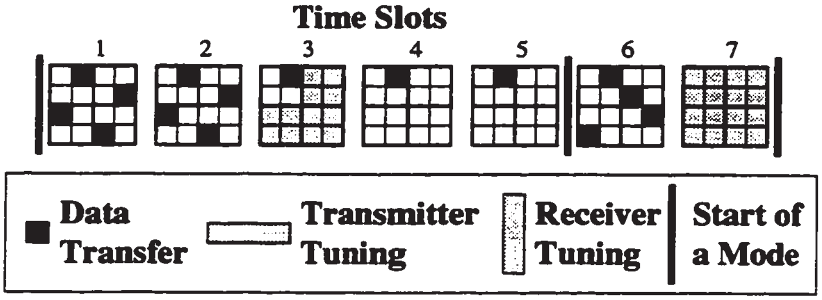


Figure 4.9: Distributed Tuning: pipelining of the tuning times with data transfers corresponding to Example 5.

Using Method 1, it can be decomposed as follows.

$$\mathbf{D} = \begin{array}{|c|c||c|c|}\hline & 5 & & \\ \hline & & & 2 \\ \hline\hline 2 & & & \\ \hline & & 2 & \\ \hline\end{array} + \begin{array}{|c|c||c|c|}\hline & 1 & & \\ \hline & & 1 & \\ \hline\hline & & & 1 \\ \hline 1 & & & \\ \hline\end{array}$$

If we consider the tuning time to be 1 time slot long ($\tau = 1$), then we can write the above decomposition as

$$\mathbf{D} = \begin{array}{|c|c||c|c|}\hline & 5+(0) & & \\ \hline & & & 2+(1) \\ \hline\hline 2+(1) & & & \\ \hline & & 2+(1) & \\ \hline\end{array} + \begin{array}{|c|c||c|c|}\hline & 1+(1) & & \\ \hline & & 1+(1) & \\ \hline\hline & & & 1+(1) \\ \hline 1+(1) & & & \\ \hline\end{array}.$$

In this case, the overall transmission duration is the sum of the maximum elements of the switching modes, which in this case is $T_{total} = 5 + (0) + 1 + (1) = 7$. If we do not consider distributed tuning the overall transmission duration can be found using Equation 4.16 as $T_{total} = (1)(2)+(5+1) = 8$, since $\tau = 1$, $N_m = 2$ and $T_D = 5+1 = 6$. This enhancement is possible since the dominant transmitter or receiver does not need retuning and we can pipeline the tuning of the other transceivers during the data transfers. This process is shown in Figure 4.9 for each time slot. We should note that this enhancement method is sensitive to the ordering of the switching modes. For example, if we change the order of the modes in Example 5, no enhancement is achievable.

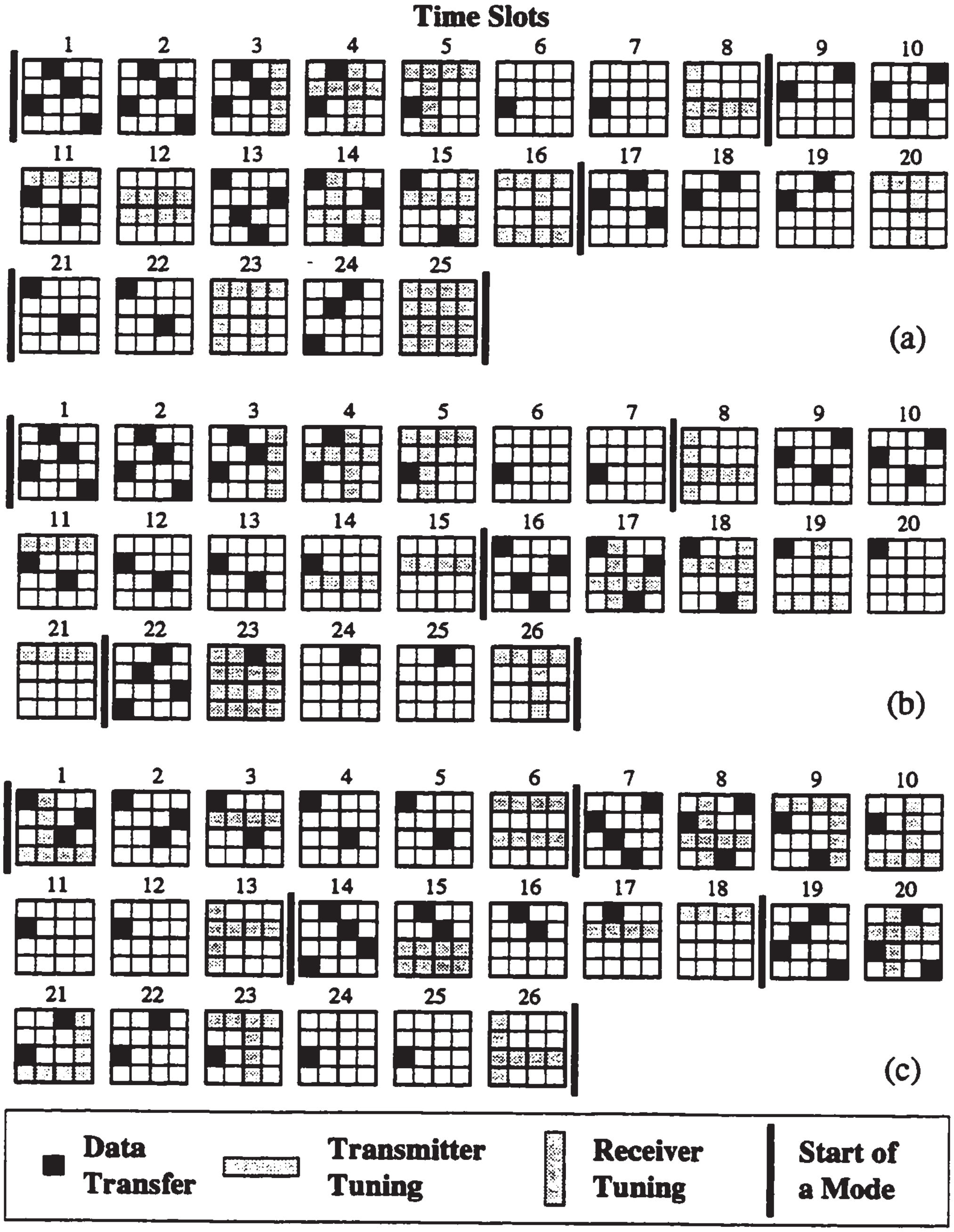


Figure 4.10: Distributed Tuning. (a) Method 1, Example 2. (b) Method 2, Example 3. (c) Method 3, Example 4.

The case shown in the above example is not a common one. In most of the cases considered this distributed scheduling mechanism is not very efficient. Figure 4.10 shows the distributed tuning for all the proposed heuristic methods. The traffic matrices and the corresponding decompositions are the ones previously shown in Examples 2, 3, and 4. In the figure, the corresponding TSAs are illustrated for each time slot. The tuning times are pipelined whenever it is possible. Using Equation 4.16, we can easily see that no improvement in these cases have been achieved by application of pipelined tuning. More investigation and mathematical elaboration are needed to obtain efficient distributed scheduling with pipelined tuning times.

4.8 Random Local Media Access Control Protocols

Random access methods are superior to pre-scheduled ones in terms of their ability to accommodate bursty non-uniform traffic with less delay and less complexity. However, they are not that efficient in the case of uniformly distributed heavy traffic patterns. Random access methods usually have less synchronization requirements. Furthermore, different local media access control (L-MAC) protocols can be applied to different LONs in a WDM cross connect network.

In the architecture shown in Figure 4.11, a control channel λ_C is introduced to coordinate the access of the different stations inside each LON. As shown, the control channel is reused and is local to each LON. The control channel is separated by a wavelength selective coupler (WSC) which filters out the control channel and then resends it to the same LON through the combiner in the receiving side. Each station has one fixed transmitter and one fixed receiver for access to the control channel λ_C.

The connection table for the WDM cross-connect is the same as the one shown in Table 4.2. Again each station is also equipped with one tunable transmitter and one tunable receiver for access to the data channels. The tuning time table of the receivers are assumed to be fixed according to the ITDMA method. We continue our description with the case of $n = 3$ and $N_s = 3$, just to make it more clear. The

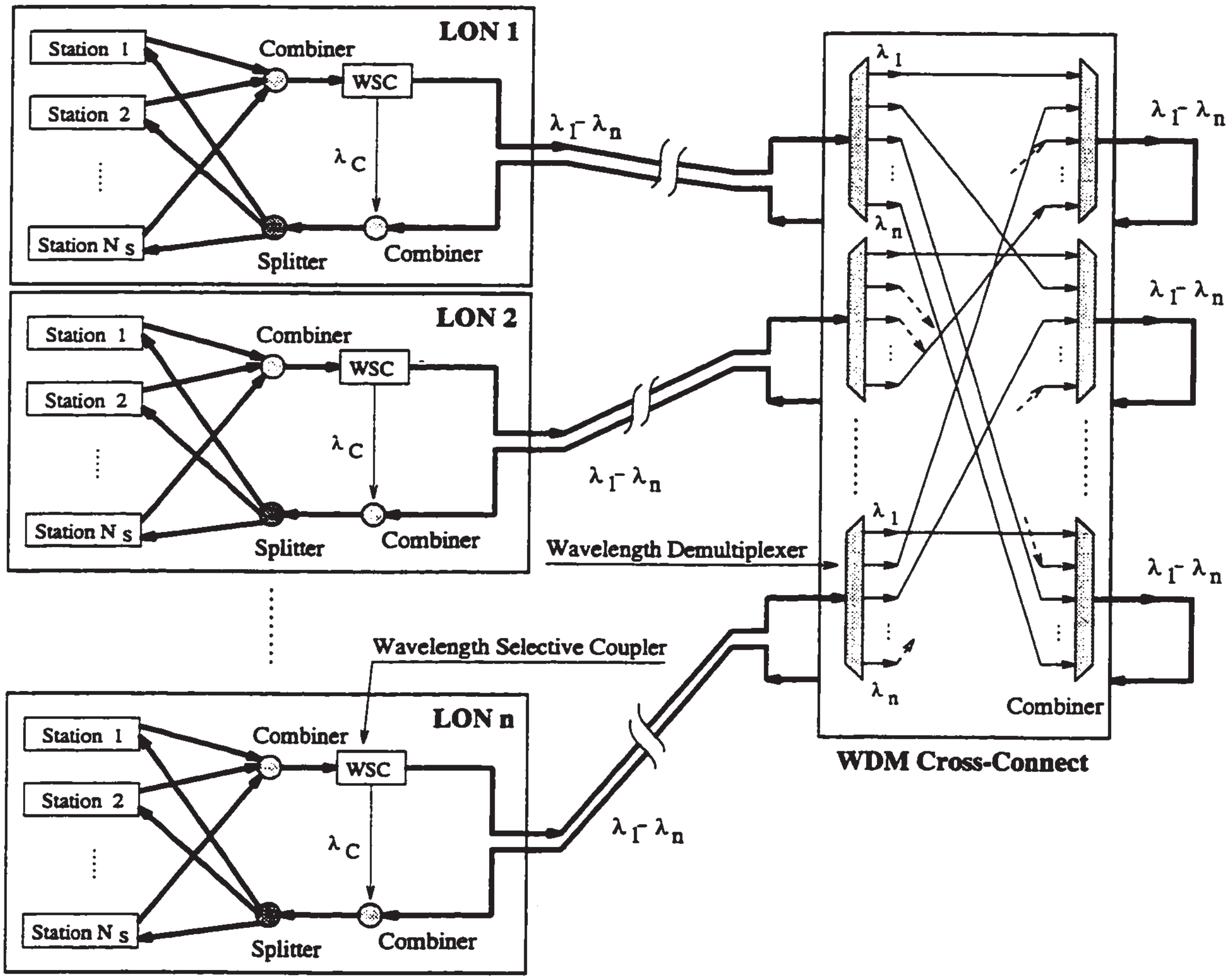


Figure 4.11: Dynamic bandwidth allocation in WDM cross connect. Control Channel is reused.

receiver tuning time table for a frame is assumed to be fixed as

Station	*Time Slot*		
	1	2	3
1	λ_1	λ_2	λ_3
2	λ_3	λ_1	λ_2
3	λ_2	λ_3	λ_1

and the same for all the LONs. The total frame duration of this tuning table is

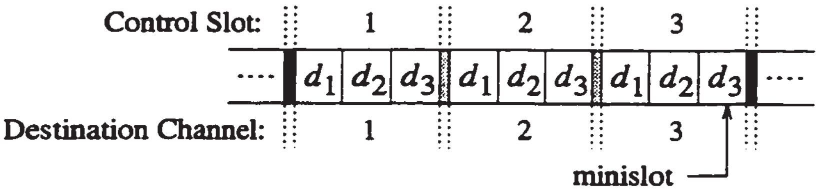


Table 4.4: Control Frame for the first method of random L-MAC.

one third ($1/n^{th}$) of the tuning table for the same system using WR-ITDMA.

In the following, we present two methods of L-MAC for the single-hop WDM cross connect network.

4.8.1 First Method

In this method each control frame is divided in $n^2 = 9$ minislots as shown in Table 4.4. There are $n = 3$ control slots in each control frame corresponding to 3 destination channels (LONs). Each control slot is divided into 3 minislots, one for each destination station in each destination LON ($N_s = 3$); d_i corresponds to the i^{th} destination station.

When station i in LON k wants to send a data packet to station j in LON connected by channel m, it must make a reservation in the j^{th} minislot of control slot m. For example, if station 2 of LON 1 wants to send a data packet to station 3 of LON 3, it must make a successful reservation in minislot 3 of destination channel 3 (control slot 3), since LON 1 (and consequently station 2 of LON 1) is connected to LON 3 through channel 3. If two or more stations from the same LON compete for the same destination station, a collision occurs in that control minislot. Since the control channel is fed back to the LON, all the stations can check for successful reservations and send the corresponding data packets in the proper data slot according to the known fixed tuning time in the receiver side. Those collided reservations have to be attempted again after a random backoff time.

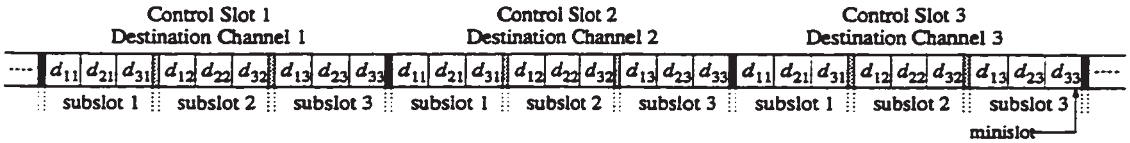


Table 4.5: Control Frame for the second method of random L-MAC.

4.8.2 Second Method

As we observed in the First Method, if there are two or more attempts at making reservations for the same destination station, a collision occurs and the stations have to try again. In the second method we extend the number of control minislots to n times (3 in our example) of the previous case, so that when there is more than one request for the same destination, one of them can get through. In this case, each of the previous minislots is actually expanded to n minislots. The resulting frame is shown in Table 4.5.

In this method, each source station has a specific minislot on each control slot (corresponding to each channel) to any destination station. As shown in Table 4.5, minislot d_{ij} of each control slot (destination channel) is dedicated to the reservations of source station j for its transmissions to station i in the corresponding destination LON. We call a group of minislots dedicated to each of the source stations a "subslot". Source station k which has a data packet to be sent to station j in the LON connected by channel i, must make a reservation in its own dedicated control minislot which is minislot j in subslot k of control slot i. In our example, if station 2 of LON 1 wants to make reservation for transmission to station 3 of LON 3, it must make the reservation in minislot 3 (destination station 3) of subslot 2 (source station 2) of control slot 3 (destination LON 3).

It is apparent that reservations will not collide in this scheme. But contention still exists. In order to resolve this, each station that has a data packet to send includes a random number in its reservation minislot between 1 and a known maximum number; otherwise it just enters 0. When the control channel is fed back to the LON, all the stations can check all the reservations. If there are two or more reservations for the same destination station, the one which has the largest random number wins the competition and it will send its data packet in the corresponding data slot. The other stations compete again in the next control frame.

4.9 Station-to-Station Channel Capacity

Although WR-ITDMA is able to achieve the maximum network capacity in uniform traffic loading, its station-to-station channel capacity is very low. Station-to-Station channel capacity (denoted by $\widetilde{C}$) considers the maximum channel bandwidth that one station can obtain while transmitting to one other station. In the fixed scheduling method of WR-ITDMA this channel capacity is

$$\widetilde{C}_{WR-ITMDA} = \frac{1}{nN_s}, \tag{4.20}$$

where n is the number of available channels and N_s is the number of stations in each LON.

The random access methods described previously try to allocate the channel bandwidths according to traffic demands of the stations. In the first method described in Section 4.8.1, the station-to-station capacity is limited due to collisions in the control slots. For a large number of stations generating packets under Poisson arrival rates, the maximum uniform load station-to-station channel capacity can be written as

$$\widetilde{C}_{first} = \frac{1}{ne} \cdot \frac{T_{DF}}{\max(T_{DF}, T_{CF})} . \tag{4.21}$$

This is because of the fact that the slotted ALOHA protocol used can achieve e^{-1} times the maximum channel capacity for the traffic model mentioned above [SHP91, Tan89, BG92]. In the random local access methods described above, we assume that the receiver tuning time table is fixed. This gives us the maximum available channel capacity as $1/n$. T_{DF} and T_{CF} are the data and control frame durations, respectively.

In the second random local access method explained in Section 4.8.2, contention is resolved without collisions. In this case, the station-to-station capacity is

$$\widetilde{C}_{second} = \frac{1}{n} \cdot \frac{T_{DF}}{\max(T_{DF}, T_{CF})} . \tag{4.22}$$

It can be seen that the first scheme can achieve some improvement over WR-ITDMA, while its station-to-station capacity is dependent on the collision level occuring on the control channel. The second method, however, does not suffer from collisions on the control channel and can achieve the maximum available station-to-station channel capacity at the cost of more minislots per control frame than that of the first method.

In the first random method the number of minislots per control slot is N_s, while in the second method this number is nN_s. Therefore, the first method is less vulnerable to the control channel being a potential bottleneck. We should note however, that in the proposed random local MAC protocols, the reservations are being done on a frame by frame basis. Each data frame is n data time slots. As a result, it is very unlikely that the control frame will become bigger than the data frame, so that the control channel becomes a bottleneck. In both station-to-station capacities shown in Equations 4.21 and 4.22, $T_{CF} \ll T_{DF}$ and $\max(T_{DF}, T_{CF}) = T_{DF}$. Consequently, the term $\frac{T_{DF}}{\max(T_{DF},T_{CF})} = 1$ and has no effect on the station-to-station capacities.

In dynamic TSA for the WDM cross connect networks, station-to-station channel capacity can be 1. This means that the entire channel bandwidth may be assigned for a connection. This however, is achieved by a more complicated MAC protocol than those used in random local access controls.

4.10 Summary

Single-hop architectures which perform media access control across a spatial wavelength cross connect are presented in this chapter. The designs reuse the available wavelength channels so that they can achieve n times the capacity of a single star network, where n is the number of available wavelengths. This large capacity cannot be fully utilized without proper MAC protocols. Dynamic time slot assignment techniques are introduced which dynamically allocate the available bandwidth on different channels according to the traffic demands of the stations.

New scheduling schemes are developed based on extensions to SS/TDMA methods used in satellite communications. It is shown, however, that the bandwidth allocation for WDM cross connect networks is a much more complicated problem than the one for SS/TDMA. Thus, direct application of the scheduling methods is not possible. A number of theorems are presented which indicate the characteristics of the new scheduling problem and define its bounds. Since the problem of the total transmission time optimization is found to be computationally intractable, heuristic methods are proposed to achieve optimum scheduling based on the results obtained from the mathematical model. Total transmission time is a function of two parameters, data transmission duration time and number of switchings. Using the proposed methods, the total transmission time can be optimized by exclusively fixing either the data transmission duration or the number of switchings to the minimum value, and then minimizing the other one. It is shown that the more optimized scheduling method is slower than the others, but more efficient.

Finally, random local medium access control protocols are considered. These protocols are suitable for situations where faster and localized bandwidth allocation is required.

Chapter 5

Dual-Hop WDM Cross Connect Networks

5.1 Prelude

In single-hop networks using a passive optical star coupler, the total throughput is upper bounded by the number of available WDM channels. It is apparent that the number of available wavelengths may initially be less than desired. As a result, spatial reuse may be required in order to obtain a LAN which will support a reasonable number of stations. In [JT94a, KFG94, DBAP93], hierarchical single-hop architectures have been considered which permit sets of wavelengths to be reused across a number of local subnetworks. This is accomplished by wavelength-routing a set of local wavelengths from each local optical network (LON) directly back into each LON for use locally. The remaining wavelengths from all LONs are coupled together by a passive star before returning them to the LONs. As a result, this set of remote wavelengths carries all inter-LON traffic. This scheme can take advantage of the traffic locality which is often associated with large LANs and MANs. Unfortunately, when traffic flows are more uniform, the performance improvements may be very poor.

In [Mat93], a design called MANDALA is considered where media access is performed across a wavelength-routed spatial cross connect. In this design, each LON is connected directly to the cross connect which provides wavelength-routed paths from each LON to all others. As a result, the allocation of bandwidth is not hierarchical and much improved capacity performance is possible even for uniform traffic. In MANDALA, global coordination is achieved using a control channel implemented using additional optical hardware. This is used to schedule all packet transmissions through the cross connect. In such a system, however, the control channel may easily become congested as traffic loading increases. In addition, dynamically tunable station transmitters and receivers are required to operate the system. This drastically complicates the operation of the system compared with that of single star or hierarchical star networks.

In the previous chapter we discussed single-hop WDM cross connect networks and their properties. Using wavelength routing and wavelength reuse we can achieve a much higher network capacity with respect to similar architectures. In this chapter, we propose the dual-hop networks to simplify system operation in WDM cross connect networks. These networks consist of two stages. In the first, the wavelength agility of the sources is used to route packet transmissions from a given local optical network (LON) to a destination LON. As in the single-hop cross connect architecture, a wavelength cross connect is used for this purpose. When packets arrive at a destination LON, they are buffered and subsequently transmitted onto the desired destination wavelength. Input access is determined using a local media access control (L-MAC) protocol. We first consider the use of a form of interleaved TDMA (ITDMA) for this purpose and then we will present a two level protocol to dynamically adapt system operation to different traffic loading conditions.

There are a number of advantages to the dual-hop design. Since buffering is introduced into the LONs, slot synchronization is only required between stations associated with the same LON. This simplifies the station design since those stations will be in close physical proximity. Also in such a system, it is difficult to dynamically allocate destination slots since unlike the single-star case, channel feedback is not available. In the proposed network this problem is alleviated because dynamic contention is resolved remotely, after the buffering stage. In addition, the dual-hop nature of the network eliminates the need for dynamic station receiver agility and the ensuing protocol control complexity. Finally, the design is a hybrid opto-electronic one which exploits the increasing availability of inexpensive commercial ATM (Asynchronous Transfer Mode [SAH94, Onv94]) switch/buffering components.

In future optical networks, certain electronic functions such as routing and buffering may be achieved optically. The advantage of this is that hardware component counts and data bussing may be reduced. Although dual-hop networks would normally use electronic ATM buffering subsystems, we also show that they may be implemented using almost all-optical buffering. Extensions of the systems first proposed in [Haa93] are used to build buffered optical concentrators from optical delay lines. A technique is introduced which permits a significant reduction in the LON buffering requirements through wavelength sharing. Results are presented which consider the reduction in buffering required under different loading conditions at the inputs to the LONs. A simple delay model is also introduced which can be used to approximate the overall mean delay performance of the system. Finally, we propose two level dynamic bandwidth allocation protocols to adapt the network to different traffic patterns.

5.2 Dual-Hop Architecture

The physical architecture of the dual-hop networks considered is shown in Figure 5.1. As in Figure 4.3, the network is unfolded showing station transmitters on the left and receivers on the right. Also, as in WR-ITDMA, stations can wavelength route packets to each of the LONs using the wavelength agility of their transmitters. In both single-hop WDM cross connect networks and MANDALA, when a packet is wavelength routed to its destination LON, the receiving station must be tuned to the correct wavelength in the proper time slot. In a dual-hop network, this interaction is decoupled by placing a buffering module at each LON. There are a number of advantages to considering this design. From a protocol standpoint, the system is simplified since stations engage in media access only with stations in the same LON. Slot synchronization and other functions are eased. Also, stations can be implemented using a single non-agile receiver and a single tunable transmitter, rather than requiring multiple tunable devices per station. Finally, in the buffering stage, commercial ATM

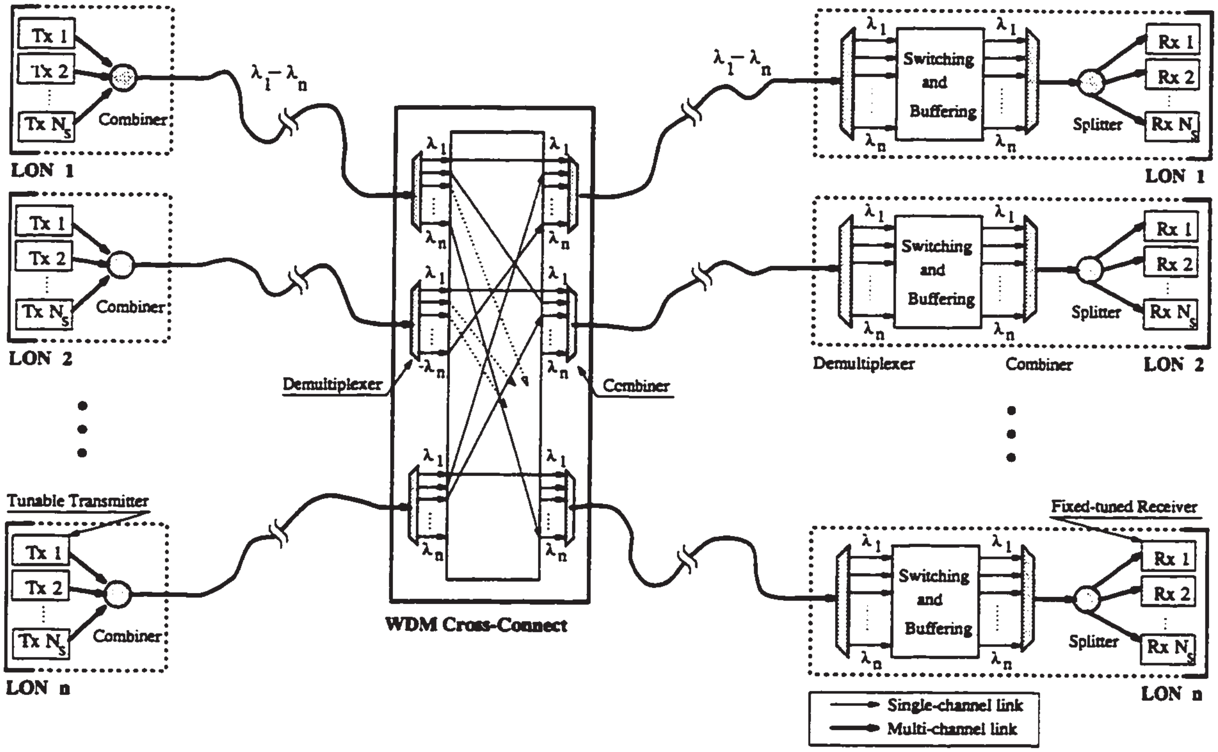


Figure 5.1: Dual-Hop Cross Connect Network Architecture

buffer/switch modules can be used. This permits a simple design based on off-the-shelf electronic switching components.

An output port at each of the LON switching modules is associated with each WDM channel. The switch thus moves packets from the correct input port to the port associated with the destination's receive wavelength. Since the switch sources all second-hop transmissions, there is no media access contention at the outputs. Concurrent transmissions are at different wavelengths and thus may be coupled into a single fiber and delivered to the destination LON as shown. This means that the stations and their buffer module can be physically separate, permitting a fair bit of flexibility in layout. Again, the buffer module would typically be located in a wiring closet in the vicinity of the stations. At first, we consider the use of ITDMA for local media access, and later we use dynamic reservation-based protocols to adapt the system to different loading conditions. Other protocols such as the one proposed in [JT94b] could also be used. The stations each have a single receiver which is assigned to a HOME channel wavelength when the station is initialized. For convenience, we will assume that stations are placed on HOME channels using a round-robin assignment starting with Station 1. It is important to note that the ITDMA frame cycle length is $1/n$ times that of the WR-ITDMA case. This advantage results from the buffering introduced, since a station can reach any destination wavelength once it routes a packet to the proper LON.

It is interesting to briefly compare the requirements of a single electronic switch implementation of this system. If for example, the cross connect used only 10 wavelengths, each running at 1 Gbps, then the maximum network capacity would be 100 Gbps. This would require a switch with 100 input and output ports, each operating at 1 Gbps to achieve full connectivity. Obviously this single switch could be implemented using a multistage interconnection of smaller ones. However, it is clear that the use of wavelength agility and a 10 × 10 passive cross connect results in a very simple overall design using fewer components. Only 10 of 10 × 10 electronic switches would be needed. Commercial feasibility in the near future will depend on the production of stable multi-wavelength sources. It is also clear from the design that adding additional buffering stages cannot improve the uniform capacity of the system or increase wavelength reuse since the upper bound of n^2 is already achieved.

5.3 Dual-Hop Architecture without Wavelength Conversion

In the architecture shown in Figure 5.2, there is no need for wavelength conversion. When a station in any LON wants to send a packet to any station in any LON whose home channel (HOME) is λ_i, it sends that packet on λ_i using its L-MAC. In the central section where a WDM cross connect routes the incoming packets to different LONs, each packet header will be checked and switched to the proper destined LON. For example if a station in LON 1 wants to send a packet to a station in LON 3 whose home channel is channel 4, it will send it on λ_4 and the switching control circuit in WDM cross connect will route that packet to the buffer module corresponding to channel 4 (or λ_4) of LON 3. It can be seen there is no need for wavelength conversion.

If we want to reduce the hardware complexity of the WDM cross connect, the dynamic switching function can be removed. A LON address filtering section, can then be added to the next stage of the cross connect of Figure 5.1 to filter the packets destined to a specific LON, as shown in Figure 5.3. After wavelength demultiplexing of the optical signal received in each input port of the cross connect, each wavelength channel is distributed to all the output ports, where a LON address filter selects and routes only those packets destined to that LON. Here, the LON address filter is a simple header check for LON address only, which is a fixed static filtering. This can be done in the optical domain by using optical correlator, as shown in Figure 5.4.

Comparing Figures 5.2 and 5.3 we find that by removing the dynamic switching section and adding fixed static LON address filtering the hardware is simplified, but the optical signal that is routed to the output port has only $1/n^{th}$ of the power of the signal in the input port because of the splitting.

5.4 Network Capacity

In the following discussion, we briefly consider the maximum traffic capacities of the architectures discussed above, and compare them with several others. This capacity is defined to be the maximum possible system throughput considering the contributions of all channels and may be considered an upper bound on that attainable in practice. The results will be normalized to the transmission rate of a single WDM channel.

In a network using a single star coupler, with n wavelengths, the maximum capacity is given by

$$\widehat{C}_{STAR} = n. \tag{5.1}$$

In a two-level hierarchical design such as the EDCCN system described in [JT94a], there are n_L local and n_R remote wavelengths. The remote wavelengths are shared globally, but the local ones are spatially reused in each LON. Assuming that the number of LONs is given by L, the maximum capacity is therefore

$$\widehat{C}_H = n_R + Ln_L \tag{5.2}$$

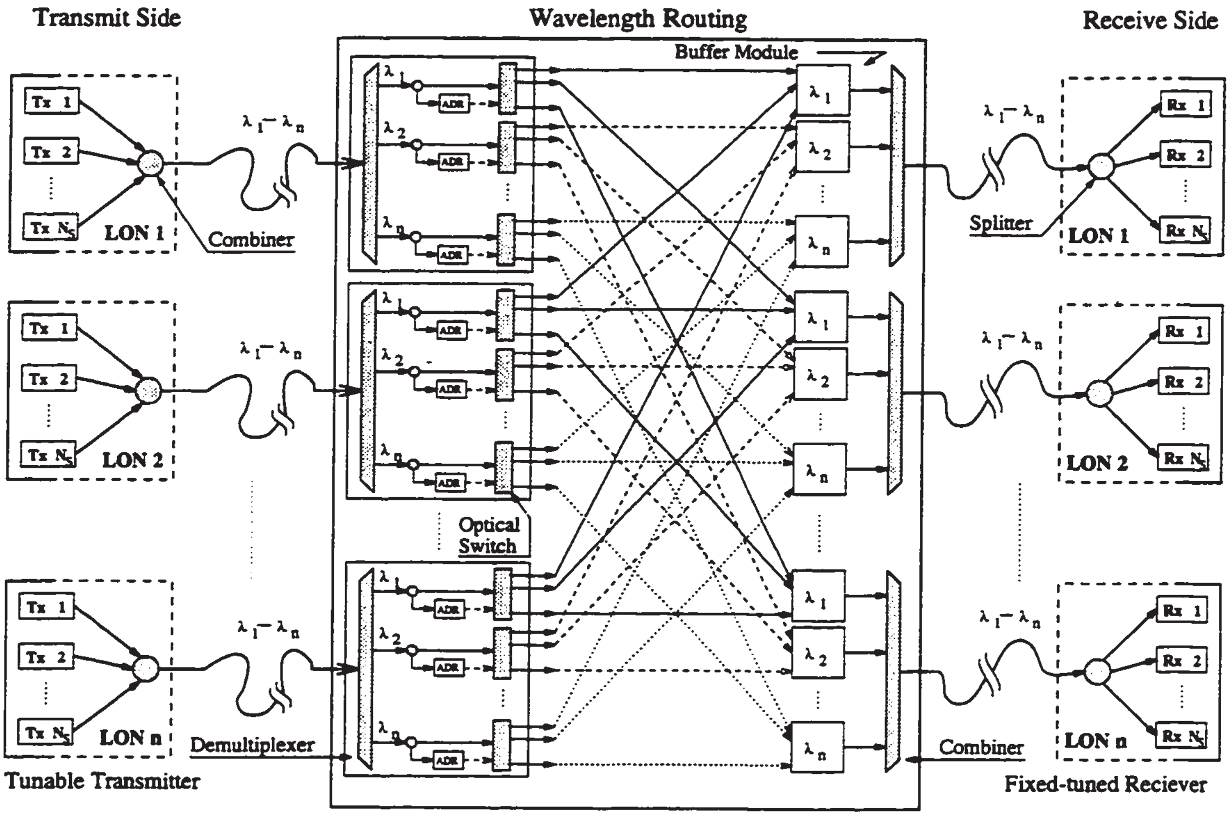


Figure 5.2: Dual-hop architecture with no wavelength conversion.

$$= \hat{C}_{STAR} + (L-1)n_L,$$

since $n_R + n_L = n$. In this case, the capacity may be increased above that of a single star system simply by increasing the number of LONs without having to add additional wavelengths. However, in order to achieve this level of throughput, system traffic loading must have the required degree of locality [JT94a].

The system described in [Mat93] (MANDALA) uses a WDM cross connect. If we assume a large station population, then it can be shown that the maximum capacity is given by

$$\hat{C}_{MANDALA} = \frac{n(n-1)}{2e} \cdot \frac{T_{DF}}{\max(T_{DF}, T_{CF})}, \tag{5.3}$$

where T_{DF} and T_{CF} are the data and control frame durations, respectively. Unfortunately, the number of request minislots per control frame grows with the square of the number of channels which means that very few channels can be accommodated

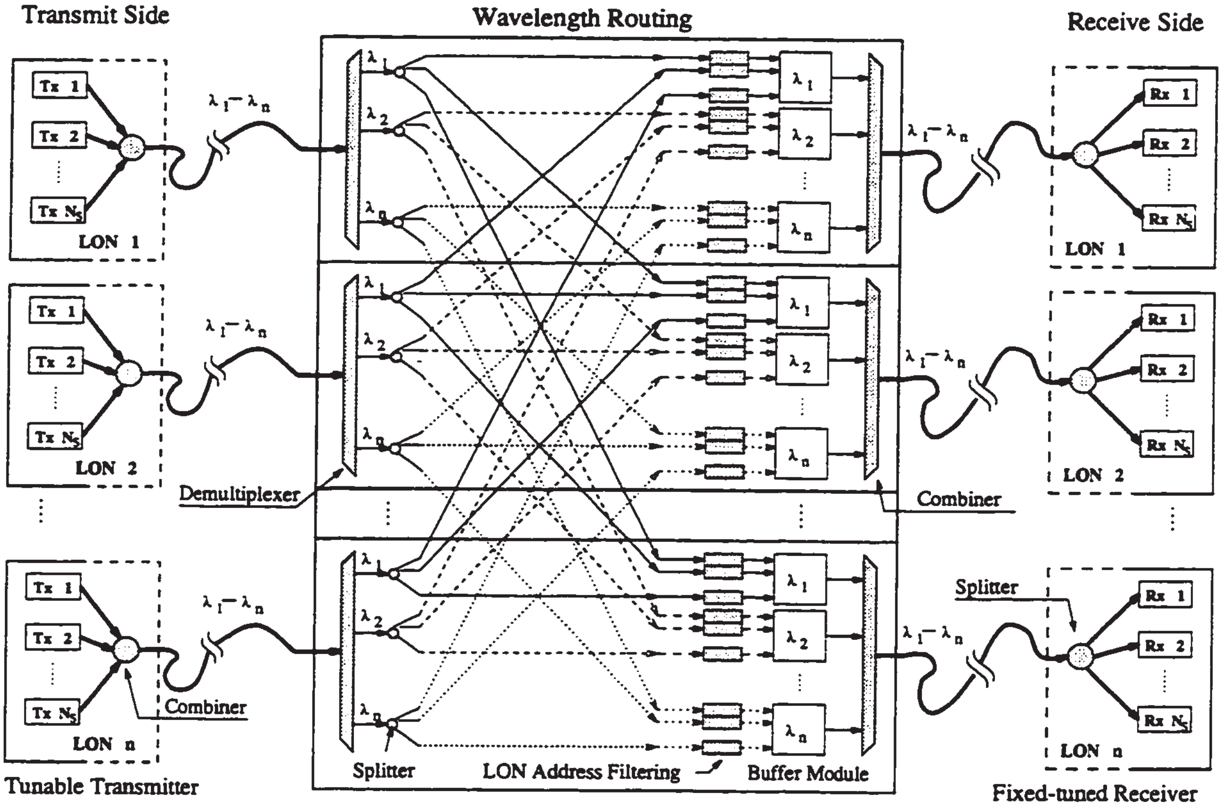


Figure 5.3: Dual-hop architecture with no wavelength conversion. Optical switches are removed and static LON address filtering is added.

before the control channel becomes a bottleneck. When this occurs, it prevents full data channel utilization.

In the dual-hop design, assuming that the number of LONs is equal to the number of wavelengths, the maximum system capacity is given by

$$\begin{aligned} \hat{C}_{DH} &= n^2, \\ &= n\hat{C}_{STAR}, \\ &= \frac{2e\max(T_{DF},T_{CF})}{T_{DF}}\hat{C}_{MANDALA} + n. \end{aligned} \tag{5.4}$$

This is the same as that which can be attained by WR-ITDMA under uniform load conditions. The dual-hop design achieves this without tunable station receivers and permits more dynamic allocation of capacity even when ITDMA is used on the LON uplinks. In addition, compared with MANDALA, the capacity may be considerably

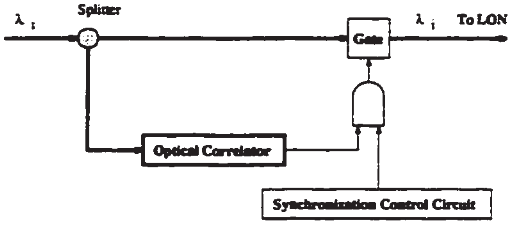


Figure 5.4: Optical LON address filtering.

higher, especially if the second term in the max function from Equation 5.4 dominates. As an example, when $n = 6$, the dual-hop capacity can be as high as 36. However, in MANDALA, assuming that data slot lengths are 10 times that of the minislots, the resulting capacity is less than 4, which is only about 11% of this value. The gains in capacity compared with hierarchical designs are dependent upon the degree of traffic locality present. If the station traffic flow is uniform for example, then a hierarchical network provides very little advantage over that of even a single star system.

5.5 LON Buffering

A salient feature of dual-hop networks is that packets are buffered at the LONs prior to delivery. A brief discussion of LON buffering using conventional electronic shared memory switches is first given. This is followed by a detailed discussion which shows that dual hop networks can also be implemented using the optical buffering techniques, such as the one used in [Haa93, T^+94, HAB94, TG87, G^+94].

5.5.1 Electronic Buffering

In this section, we will briefly consider two basic design alternatives and their suitability to the dual-hop architecture. The intent is that in a practical implementation, simple off-the-shelf ATM switch components would be used. Accordingly, we will consider $N \times M$ switching subsystems as our basic building blocks in the design of the LONs. For purposes of discussion, we will assume that these components are based on a simple shared-memory design as discussed in [GHJ94]. Thus an $N \times M$ subsystem

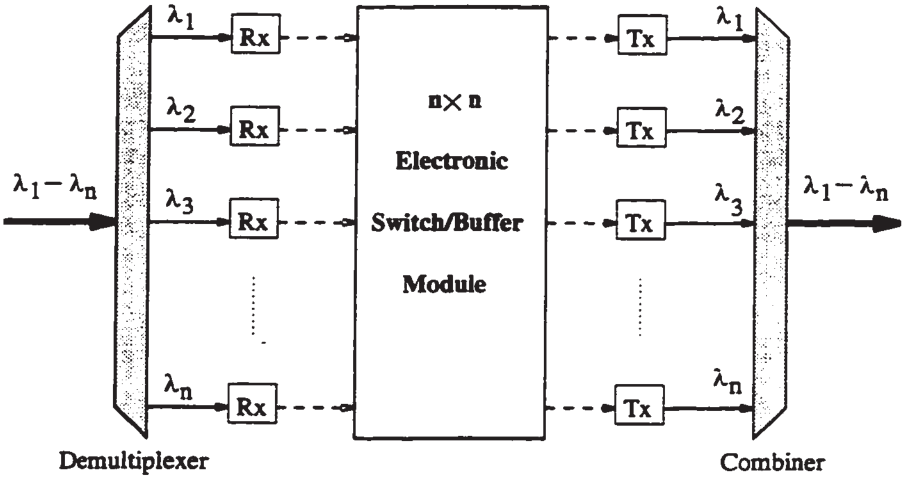


Figure 5.5: Switching and Buffering stage at each LON. An $n \times n$ electronic packet switch/buffer is used.

would be strictly non-blocking and has an internal speedup that is proportional to $N + M$.

In Figure 5.5, a dual-hop design is shown which uses one $n \times n$ switching subsystem at each LON. This implies that the switch has the required internal speedup to accommodate all $2n$ ports. If this is not the case (due to a high transmission rate or large value of n), further optical processing can be used. In Figure 5.6, each input is first sorted according to desired output wavelength. This is done by translating the wavelength of each incoming packet to its destination wavelength. This may be accomplished electronically by simply reading the packet header as it arrives and retransmitting it at the desired wavelength. This can be done using a single fast tunable laser or a transmitter array. In the future, this step may be performed optically via direct wavelength conversion [Bra90, Gre93]; also refer to Section 3.3.5. Following this, the packets are wavelength-routed to the appropriate buffering block. The result of this action yields n outputs, one for each destination wavelength. Each one is then applied to the buffer block which performs an $n \times 1$ concentration function. In Figure 5.6, a single memory block is used for each of the n inputs. Although this design requires n memory blocks in each LON, the internal speedup of each is now only proportional to $n + 1$. Note that in both cases, output queuing is used thus avoiding head of line blocking.

In the two buffering designs considered, we can clearly trade off the internal memory speedup required at the expense of a larger number of buffering modules. However, in addition to this, the depth of the buffers for a given level of loss performance may be reduced in the shared buffer case. Simulation results illustrating this tradeoff are shown in Figure 5.7. In the experiments, an exact simulation was used assuming independent Poisson packet arrival processes at each user station. Arrivals consist of single-slot packets, uniformly destined across all stations and are buffered in infinite queues at each station. As discussed above, ITDMA is used on the LON uplinks. The small boxes in the figure show the input link utilization at the buffers. The results compare the loss performance for the two buffering schemes when the total amount of buffering in each case is equal. This gives some measure of the price to be paid in terms of additional loss in going to a smaller buffer speedup. For example, it can be seen that when the link utilizations exceed 90% in an 8 wavelength system, a loss of roughly 10^{-3} requires at least twice the total buffering when separate buffers are used. The figures clearly show the advantage of the shared buffer design in this case, especially when the number of LONs is large. However it should be noted that different numerical advantages will result depending upon the exact details of the traffic model. In particular, under focused load situations, the single buffer design may result in much higher relative performance.

In Figure 5.8, simulation packet loss results are shown for a more bursty traffic model. As before, message arrivals at the stations are taken to be Poisson and the entire system is simulated. However, in this case the number of packets per message is taken to be geometrically distributed, with the mean indicated as the average burst length in the figures. As before, results are presented for both the shared and multiple buffer designs. The value of buffer sharing is clearly more important here compared with that of Figure 5.7. It can be seen from the graphs that in the non-shared buffer case, increases in buffer size reduces the loss rate of the system in a similar way for both 4 and 8 wavelength cases. For example, in Figure 5.7(a), increasing the buffer size from 4 per channel to 7 at a link utilization of 0.7 reduces the loss rate to almost 10% of that with 4 buffers. The same effect can be seen in Figure 5.7(c)

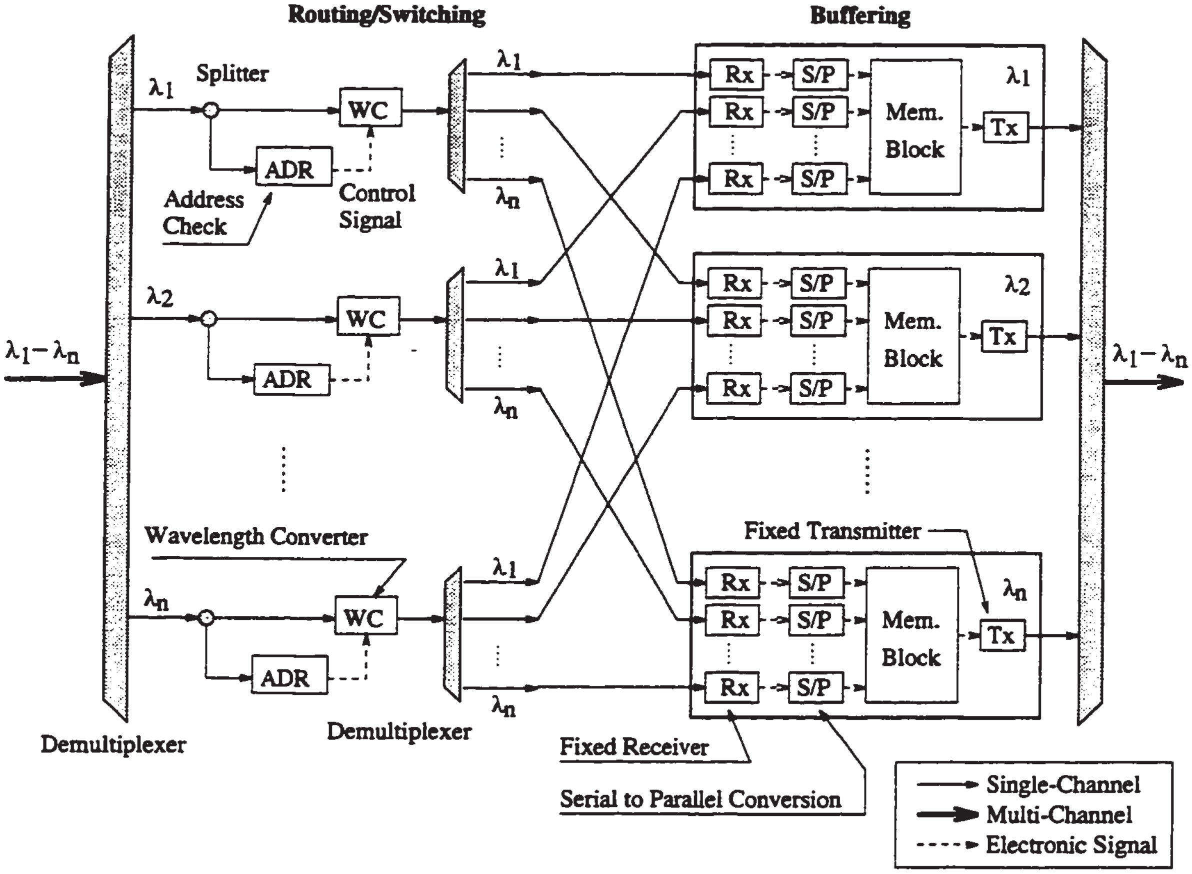
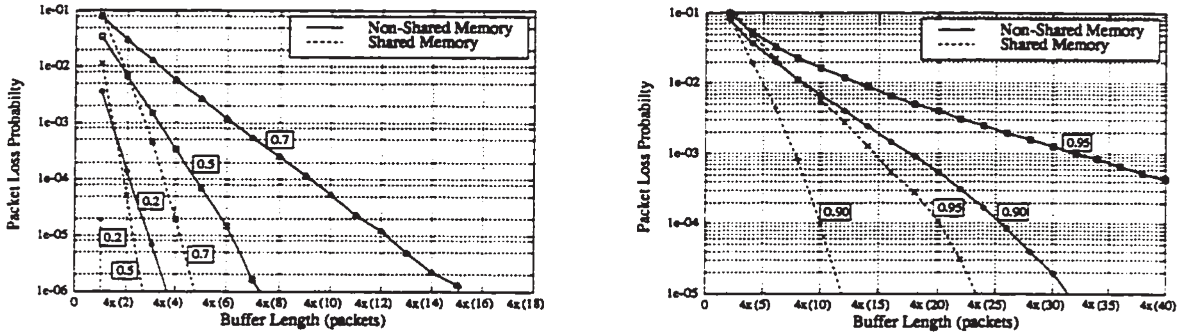


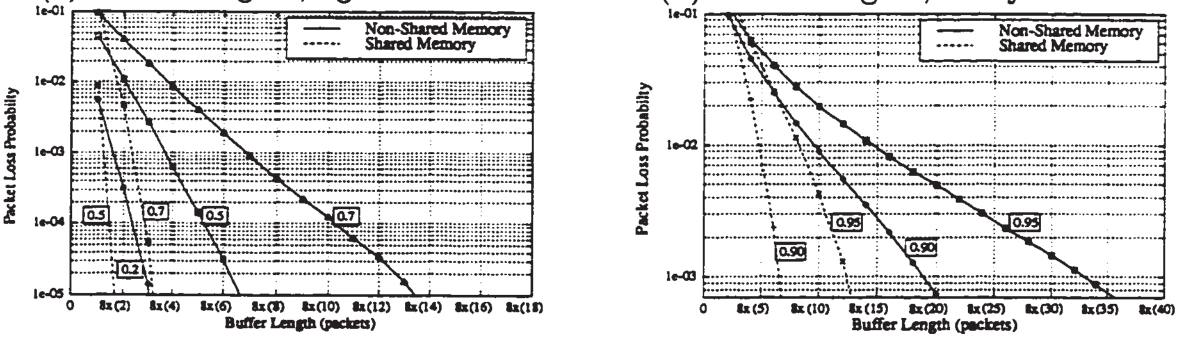
Figure 5.6: Switching and Buffering stage at each LON. Optical processing is used to sort packets destined to different channels.

for 8 wavelengths. This also holds for the other curves and is true for the non-shared buffer case under bursty traffic, as can be seen from Figure 5.8. This trend is not the same in the shared buffer system. In this case, an increase in buffer size decreases the loss rate to a larger degree in the system with a higher number of channels. This can be seen by comparing the shared buffer performance in graphs (a) and (c) or (b) and (d) of Figure 5.7. The slope of the curves is clearly higher in the 8 wavelength system compared with that of the 4 wavelength case. This is also true for the shared buffer under bursty traffic.

The packet loss rate for the non-shared buffer option is slightly better for the 4 wavelength system. For example, the loss probability for 5 buffers per channel at 0.5



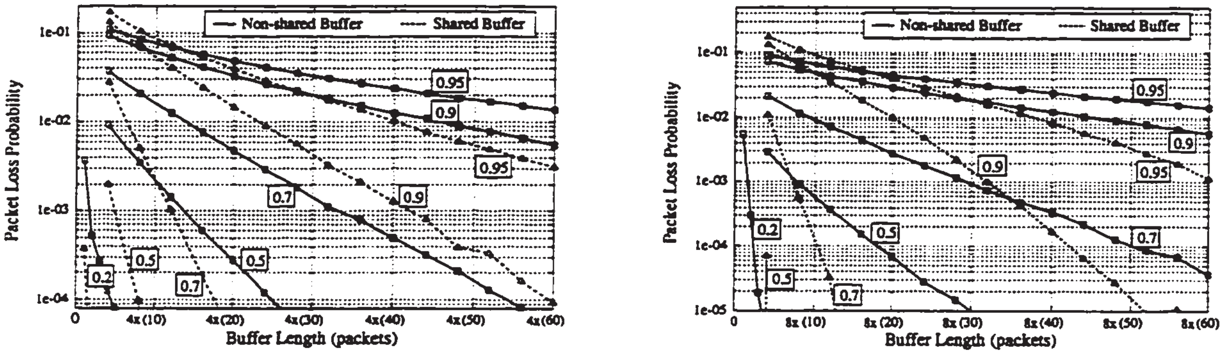
(a) 4 Wavelengths, light load (b) 4 Wavelengths, heavy load



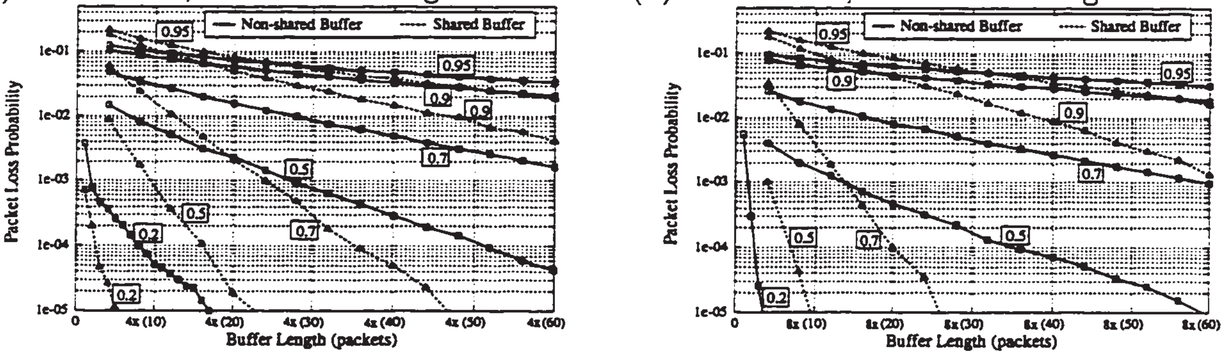
(c) 8 Wavelengths, light load (d) 8 Wavelengths, heavy load

Figure 5.7: Packet loss probabilities for non-shared and shared LON buffers. The small boxes in the figure show the input link utilization at the buffers.

utilization in a 4 wavelength system is slightly less than 10^{-4}, while the corresponding value for 8 wavelengths is slightly greater than this value. The same is true for the non-shared buffer under bursty traffic. It can be seen that in the shared buffer case, the loss rate is lower for the larger number of wavelengths. The performance of the shared buffer is clearly better for a higher number of wavelengths in bursty traffic. As seen in Figure 5.8, the loss for 40 buffers per channel at 0.9 utilization and mean burst length of 4 in an 8 wavelength system is of the order of 10^{-4} which is 10 times less than the corresponding loss rate for a 4 wavelength system. This trend continues when the bursts are longer.



(a) 4 Channels, Mean burst length = 4 (b) 8 Channels, Mean burst length = 4



(c) 4 Channels, Mean burst length = 8 (d) 8 Channels, Mean burst length = 8

Figure 5.8: Packet loss probability for bursty traffic. The small boxes in the figure show the input link utilization at the buffers.

5.5.2 Optical Buffering

In [Haa93, T^+94, HAB94, TG87, G^+94], almost all-optical techniques were discussed for implementing buffering and switching functions. The designs are considered "almost" since dynamic control and configuration of the switching elements is performed under electronic control. In this section we discuss how dual-hop networks could be implemented in this way.

To accommodate optical buffering, and to minimize the electronic control complexity, the final stage for each LON would be based on the design proposed in Figure 5.6. However, in this case, each buffering block shown would be replaced by an optical buffer module. The resulting architecture is shown in Figure 5.9. As before, when packets arrive from different LONs, power is coupled out and the packet header is read.

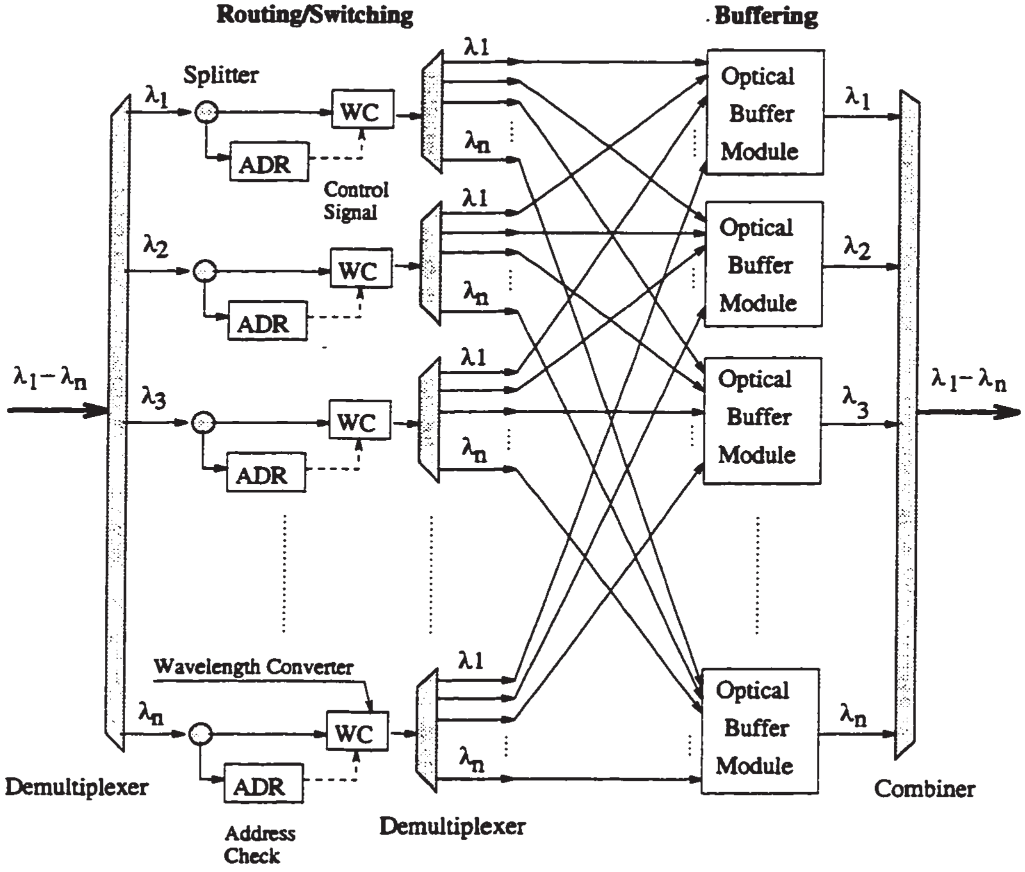


Figure 5.9: Switching and Buffering stage at each LON. Optical buffering is being used in this architecture.

The transmission is then converted to the destination wavelength of the packet. Following this, the packets are wavelength-routed to the appropriate buffering stage. At these points, the packets are buffered optically [Haa93, T+94, HAB94, TG87, G+94]. By using this particular design, the memory modules perform a simple $n \times 1$ concentration function, and the ensuing control functions are vastly simplified compared to that of a general optical buffer module [Haa93].

Figure 5.10(a) shows one method for implementing the optical buffer module. As previously mentioned, this system performs a simple $n \times 1$ concentration function. In the concentrator, m optical delay-lines, which are quantized in single-packet delay times, are used. Each incoming packet to the module will be switched to one of the these lines according to a simple control procedure, described below. We will refer to

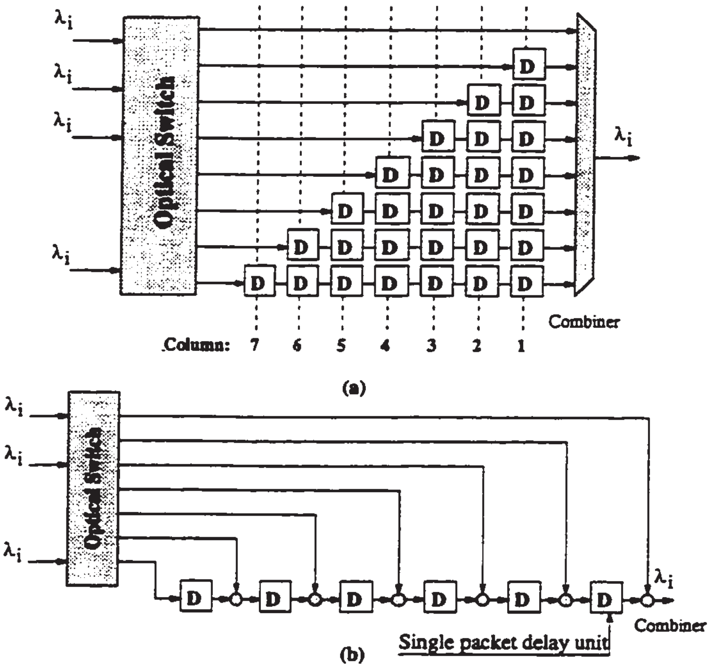


Figure 5.10: (a) Multiple Delay-Line Buffer: Incoming packets are switched to different delay lines. (b) Shared Delay-Line Buffer: Incoming packets are switched to different taps of a shared delay-line.

this design as the multiple delay-line buffer, MDLB. An alternative which uses only one tapped delay-line is shown in Figure 5.10(b). This is referred to as the shared delay-line buffer, SDLB. Here the incoming packets to the module will be switched to different taps according to the control procedure to be discussed shortly. As described above, there is one buffer module per wavelength at each LON. Thus we require n^2 buffer modules in total.

Multichannel Optical Buffers

A novel feature of the design is that it is possible to collapse the n modules per LON into a single one as follows. This is possible by noting that the wavelength in each of the n modules is different. Thus a single set of delay lines may be used to simultaneously buffer packets at different wavelengths. This arrangement is shown in Figure 5.11. The n sets of n inputs are brought into their own switch as before. The

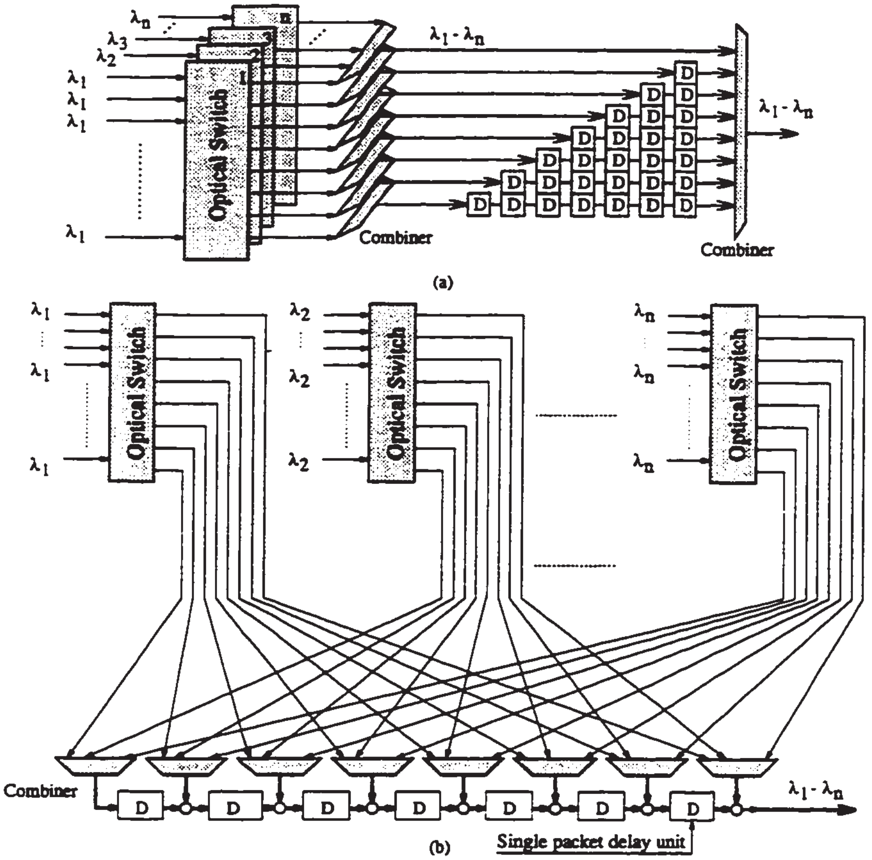


Figure 5.11: Shared Buffer: All wavelengths share a common delay line. (a) Multiple delay-line buffer-module is shared amongst all wavelengths. (b) Shared delay-line buffer-module is shared amongst all wavelengths.

same outputs for each switch are then coupled together and applied to the inputs of the delay lines. This setup provides for n independent and concurrent buffering modules, thus giving an n-fold reduction in total buffering requirements. Figure 5.11(b) shows how this technique can be applied to the shared delay-line buffer case.

Optical Buffer Control

The control procedures for the optical buffer modules should be as simple as possible to avoid processing delays and reduce electronic component counts. They must coordinate the switching of incoming packets to one of the delay lines in MDLB or to one of the taps in SDLB. It will be seen that the design considered results in very simple

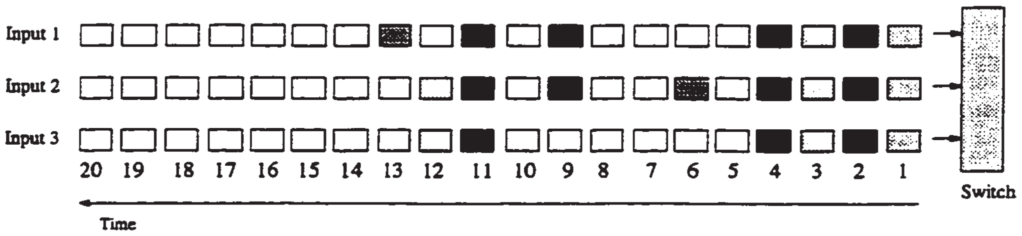


Figure 5.12: An example of optical buffer-module control. Here, there are three input lines to the buffer-module. Packets are randomly distributed on the input ports.

control functions. In order to illustrate this, a simple example will be considered. Incoming packets to a given module can occur on any of the inputs and there may be more than one in any slot. In Figure 5.12, an example for one channel is shown with three input ports and three arrivals per slot in the first four slots. Slot arrivals are shown shaded in the figure. Different shades are used to distinguish arrivals in different time slots. Clearly, a maximum of one packet per slot can be sent out and the remaining packets must be buffered.

In Figure 5.13, the control procedure is shown using the input shown in Figure 5.12 for both the MDLB and SDLB systems. A total of 18 clock cycles are shown in Figure 5.13. In this figure, Option 1 shows the state of the MDLB system, while Option 2 shows SDLB. In MDLB, the next available delay line is defined to be the shortest delay line in the first empty column, starting from the output. The column labeling convention is shown in Figure 5.10(a). In Figure 5.14, an example is given illustrating how the next available delay line is determined. A pointer is used to track this line number and is increased by one for each arrival and decremented by one in each clock cycle. For the SDLB case, the control procedure is even simpler. The next available delay buffer is the first free one starting from the output. Figure 5.14 gives an example of this. In both designs, when two or more packets simultaneously arrive to the buffer module, the switch routes them to the next available buffer slots in accordance with the above definitions. For example, the first 2 cycles in Figure 5.13 show 3 packet arrivals in each slot. In the first, one packet is sent out and the other 2 are buffered. Similarly, in the second slot, a second packet is transmitted and the three new packets are buffered. It can be seen that the procedures used maintain

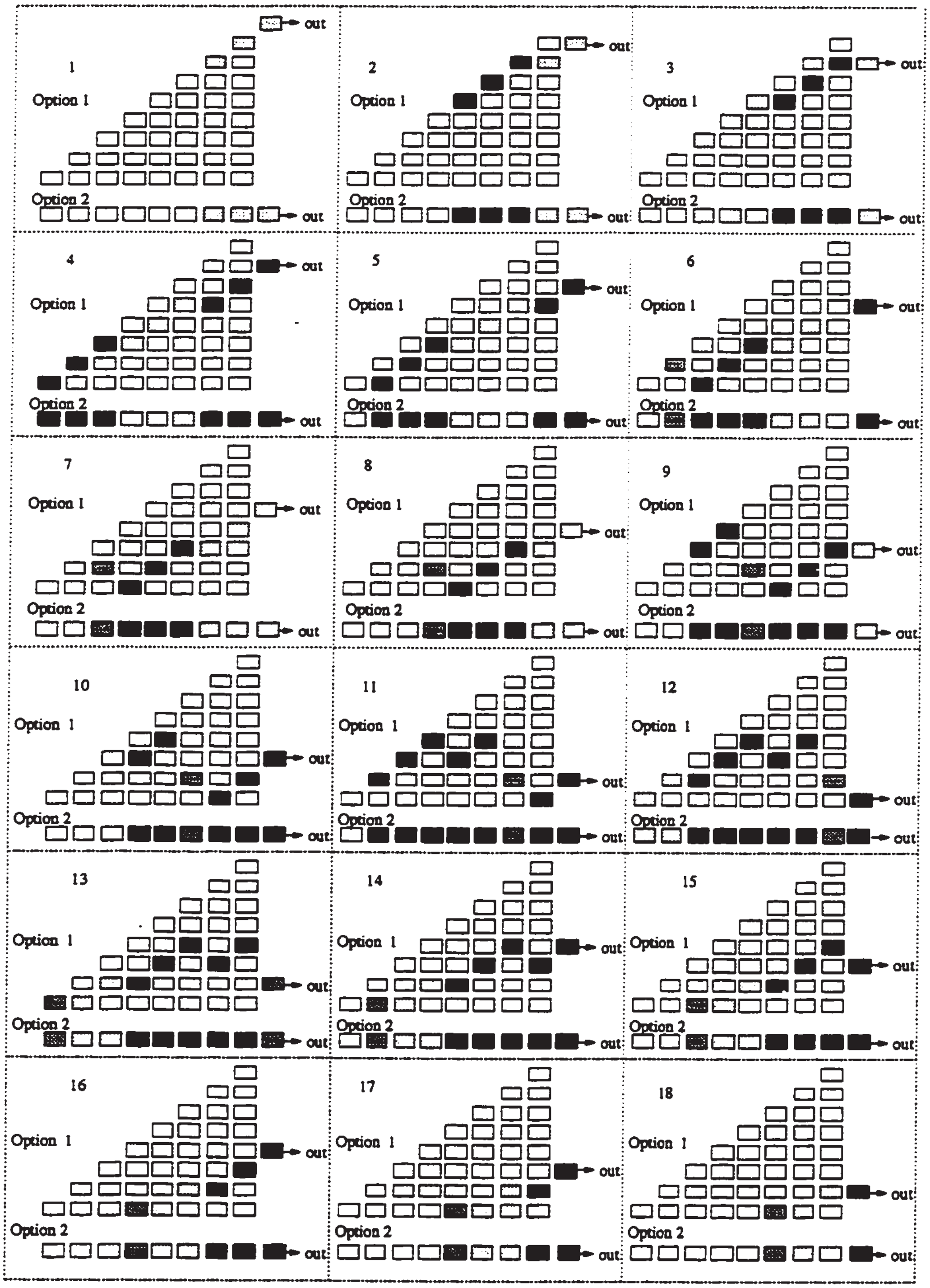


Figure 5.13: An example of optical buffer-module control. Option 1: Multiple Delay-Line Buffer, Option 2: Shared Delay-Line Buffer. The figure shows a total of 18 clock cycles.

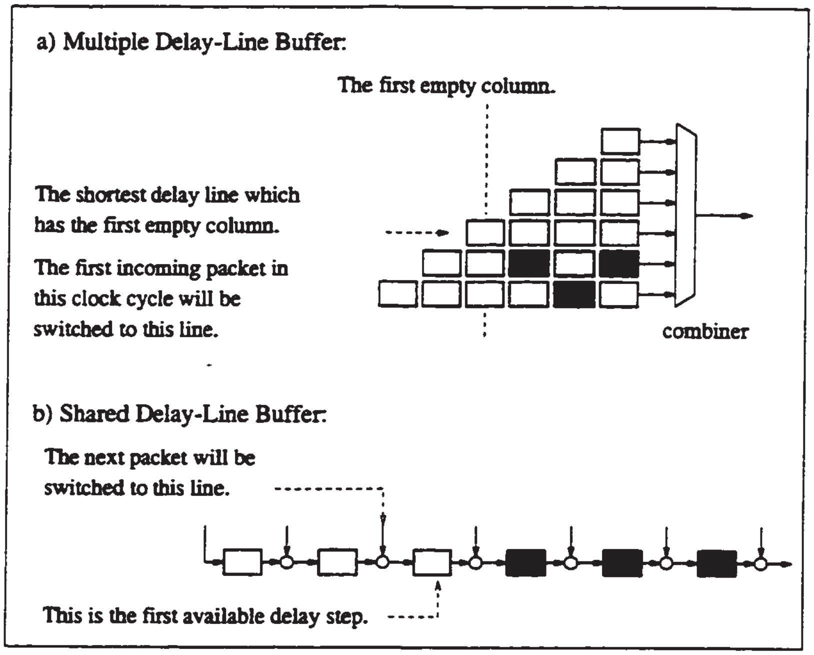


Figure 5.14: Simple buffer control procedures. (a) Multiple Delay-Line Buffer, (b) Shared Delay-Line Buffer.

packet sequencing.

5.6 Network Performance

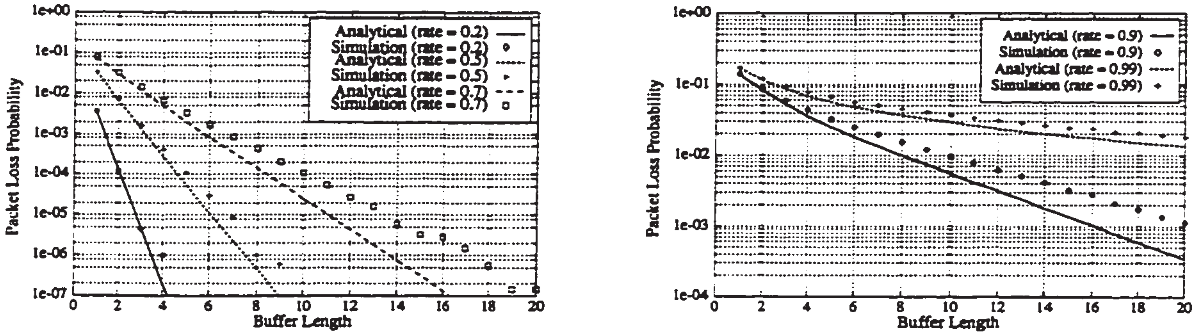
In this section we consider both buffer loss and mean delay performance of the network. In both cases, analytic models are used to approximate the system under certain conditions. The intent is to provide simple closed form results which can be used in the design of these networks. The models are validated using simulation results.

Considering the issue of packet loss, we will assume that packets arrive according to a Poisson arrival process at each station and are uniformly destined to different channel outputs. It is also assumed that there is ample buffering at each station and that the source of loss is in the shared LON buffers. Since ITDMA is being used on the uplink channels, the packet arrival process at each LON buffer is very complex. In the analytic model, we will consider the arrivals on each buffer input link to be memoryless and independent. In this case, each buffer may be modeled by a simple finite-state Markov chain. The details are very straightforward and are shown in Appendix A. In Figure 5.15 results are presented showing the packet loss performance of the system. Simulated loss results are compared to the simple analytic model discussed in Appendix A. The simulations are exact in that they model the complete system including the effects of station buffering and the ITDMA protocol. It is important to note that the buffering requirements indicated in the graphs are the total for the entire LON when the wavelength-sharing scheme from Figure 5.11 is used. In this case, a throughput n times the non-sharing case can be obtained for the same loss value. The results also show that as the number of channels is increased, the model does a much better job of predicting the actual packet loss. This is encouraging since this type of parameter value is of much more interest and is also more difficult to simulate. Deviations from the model are caused by correlation introduced by station buffering and through the media access protocol. Errors caused by these effects would be expected to be reduced for larger system sizes. It is also noted that these results should be considered optimistic, since the loss may be higher under focused or bursty traffic models.

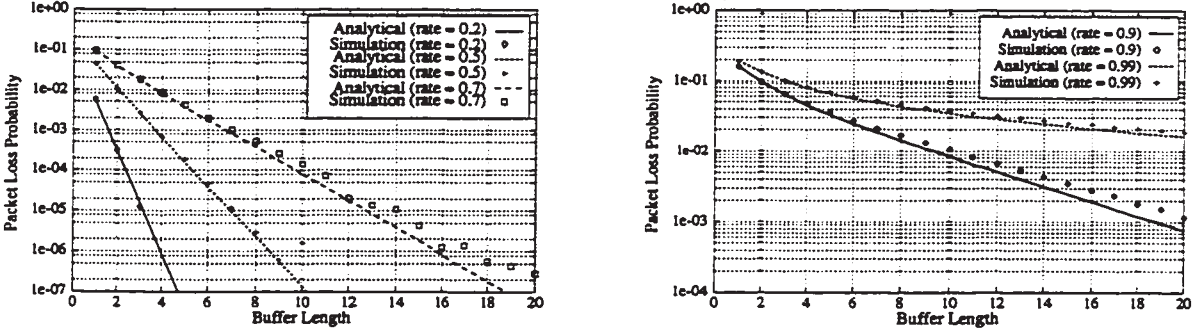
It should be noted that in case of packet loss, higher layer protocols are responsible for recovery techniques in order to insure system integrity. These schemes include retransmissions of lost packets (or data frames) and/or error corrections. These procedures may also be used when data packets are somehow corrupted in the optical path.

A precise mean delay model of the entire system is also very complex, due to coupling between the input and output sections. If we assume Poisson packet arrivals to the stations in each LON, the exact output distribution from the ITDMA stage can not easily be found. In order to obtain an approximate model of the system, we will assume that the ITDMA and buffering sections are independent and that packet arrivals to the buffering section follow a memoryless arrival process. The details of this model are also discussed in Appendix A.

In Figure 5.16 this approximation is shown in comparison with exact simulation results. Included in the graphs is the delay of a single passive star network using



(a) 4 Channels, light load (b) 4 Channels, heavy load



(c) 8 Channels, light load (d) 8 Channels, heavy load

Figure 5.15: Comparison of Packet Loss Probabilities found by analytical model and Simulation.

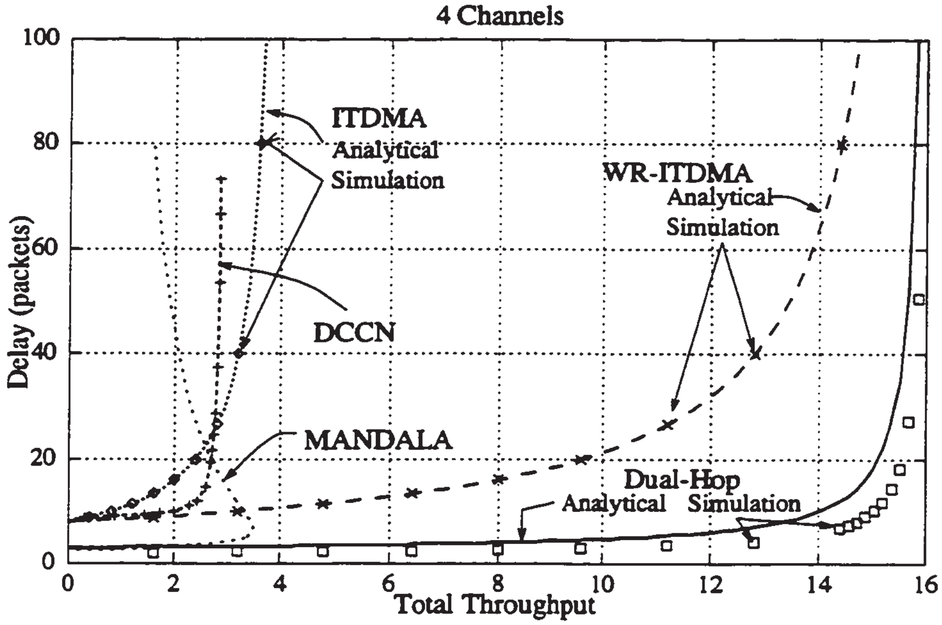
ITDMA and a single star reservation-based system referred to as Distributed Channel Controller Network (DCCN) [JT94b]. In DCCN, each wavelength is assigned a channel controller that implements a movable-boundary TDMA scheme between the channels. Each controller makes contiguous slot block allocations on its assigned channel, and communicates these to the other channel controllers. Stations on a given channel then contend for the appropriate slot blocks by sending requests to their channel controller during a request subframe. Slot allocations are then returned by the channel controllers during an allocation subframe.

In the figure, We have also included the mean delay performance for the MANDALA network, from [Mat93], and for WR-ITDMA. Results are shown for both 4 and 8 wavelength systems. The performance advantages of the dual-hop network are readily apparent. The total delay in the 8 wavelength dual-hop system is almost twice that of the 4 wavelength system under light load conditions. This trend continues up to roughly half of the maximum throughput, since under light load, the delay in the first stage (D_{local}) is the dominant factor. The average delay for ITDMA under light uniform loading is approximately half of the frame duration. We can also see this effect through Equation A.6, which shows that for light traffic the second term in the equation is the dominant one. For heavier loads, because of queuing delay in the stations, the first term of Equation A.6 will obviously come into play. The same behavior can be found in the curves for WR-ITDMA, but in this case the frame length is larger by a factor of n. When comparing the dual-hop results to WR-ITDMA under heavy loads, the improvements can be attributed to the fact that much more dynamic scheduling is achieved. Note that in WR-ITDMA the frames are defined so that a station can reach any destination channel within a single frame. Thus a frame length of n^2 slots is required to obtain full network connectivity. In the dual-hop case, each station can reach any destination channel with a frame of n slots. As a result, under light load the delay for WR-ITDMA is almost n times that of the dual-hop system. At higher arrival rates, the buffering delay, D_{buf}, becomes comparable to the first stage delay. This however, is not dependent on the number of channels, as can be seen in Equation A.8.

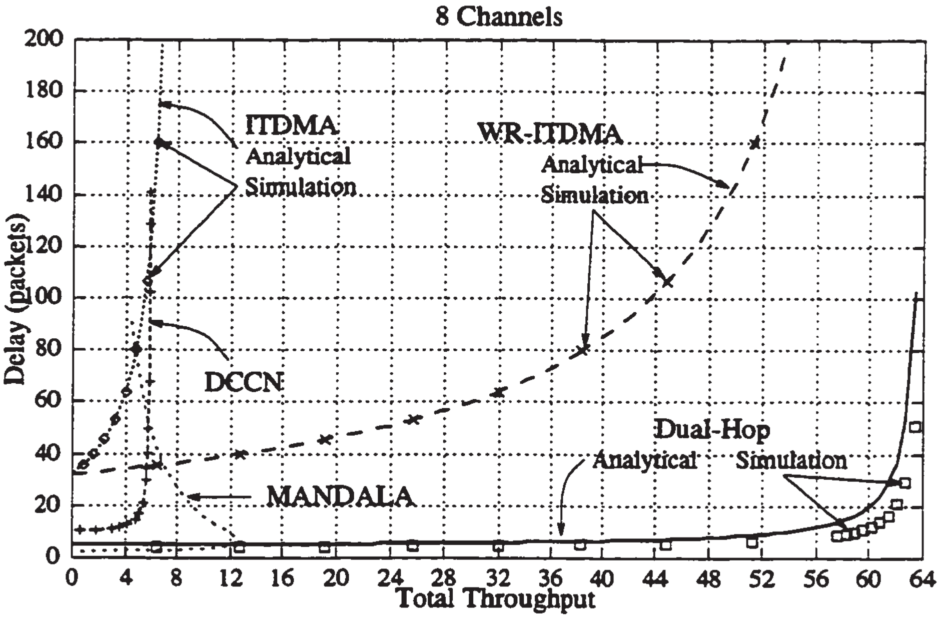
In both graphs it can be seen that the performance of MANDALA is severely limited by the bottleneck imposed by its control channel. The graphs also show that the approximate delay model is a reasonable one, although it tends to slightly overestimate the mean delay. This is attributed to the fact that the model does not take into account the true output process from the first stage of the network. Again, it can be seen that the analytic model appears to perform better as the number of wavelengths is increased.

In the previous section, performance comparisons were made assuming a traffic model where flows are uniform across all stations. In this section, we present simulation results of mean delay under a non-uniform traffic model. In this case, when a packet arrives at a station, with probability γ is destined uniformly across those stations in the same LON. Otherwise it's destination is chosen randomly across those stations outside its LON.

Figures 5.17 and 5.18 show the comparisons for a 4-wavelength system. In Figure



(a) 4 Wavelengths



(b) 8 Wavelengths

Figure 5.16: Delay vs Total Throughput.

5.17, the inter-LON traffic rate is held constant while the local traffic rate is varied. The flows per channel are shown in the boxes adjacent to the curves. Similarly, in Figure 5.18, the local traffic rate is fixed (as shown) while the inter-LON traffic rate is changed. Comparisons are made with WR-ITDMA, MANDALA and with the single star systems discussed earlier. It can be seen that in all cases improved performance is obtained by the dual-hop network. As before, this results from increased dynamic channel allocation. Also, as seen in the uniform traffic case, the control channel bottleneck in MANDALA severely restricts its capacity performance.

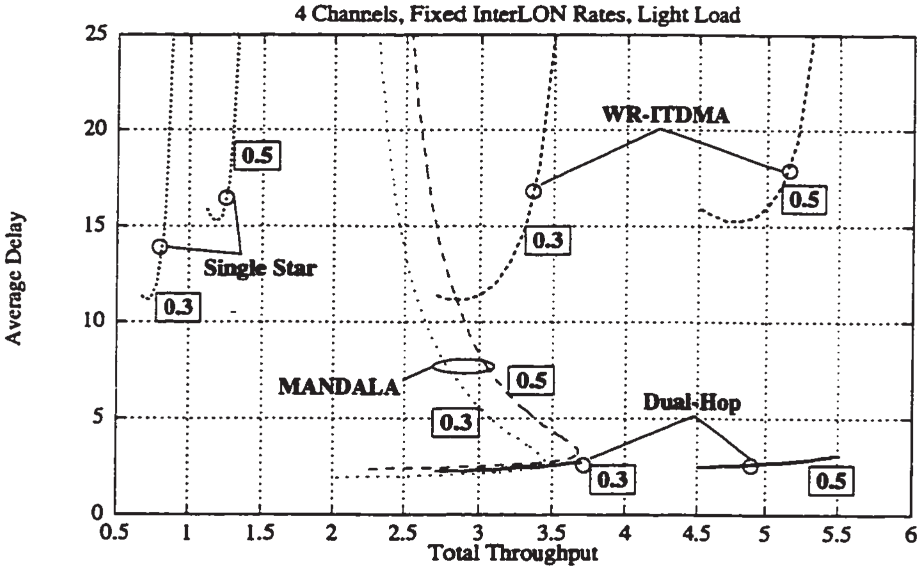
5.7 Dynamic Bandwidth Allocation

A two-level dynamic bandwidth allocation scheme can be incorporated into the dual-hop network. The new architecture is shown in Figure 5.19. As shown, one channel is dedicated as a control channel, λ_C. Each station is equipped with a transmitter and a receiver, both fixed tuned to λ_C. Requests and allocations are communicated on the control channel between stations and control agents described below.

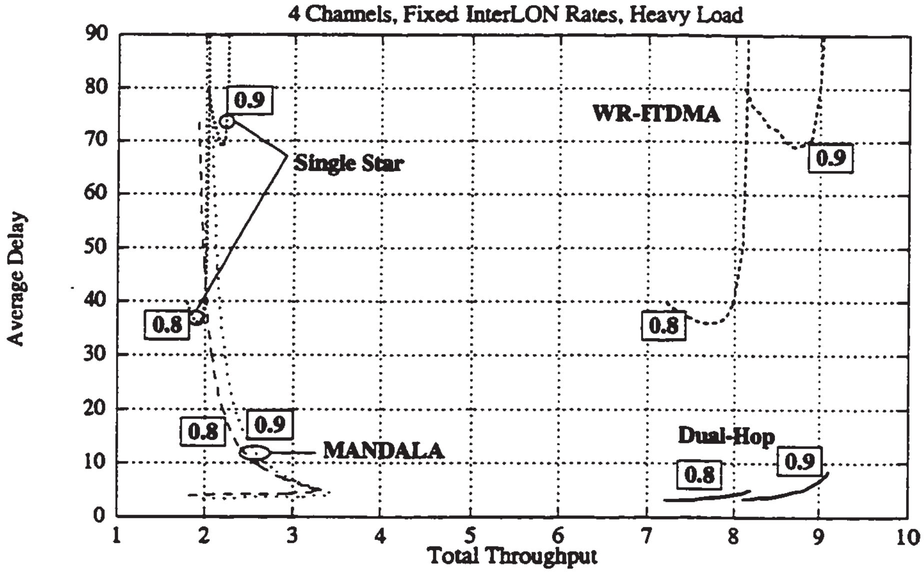
Two levels of network control exist: local controllers (LCs) are responsible for local (Intra-LON) bandwidth allocation, while the master controller (MC) allocates bandwidth for inter-LON communications.

As shown in Figure 5.19, two wavelength selective couplers are added to each LON. The wavelength selective couplers filters out λ_C from the outgoing fiber and routes that channel to the corresponding LC. Through this path, the LC receives the requests from the stations in its LON. Another wavelength selective coupler is used to send the allocations assigned by the LC back to the stations in that LON. An enlarged view of this section is shown in Figure 5.20. In this figure, the receive wavelength selective coupler is denoted by WSC1 and the transmit one is labeled as WSC2. It can be seen that the control channel is reused in all the LONs.

The LCs are also connected to the MC through the control channel, as shown in Figure 5.20. Each LC sends its requests to the MC through WSC1 and receives allocations performed by the MC through WSC2. Therefore, the LCs are serving two networks, one to communicate to their LONs and the other to communicate to the

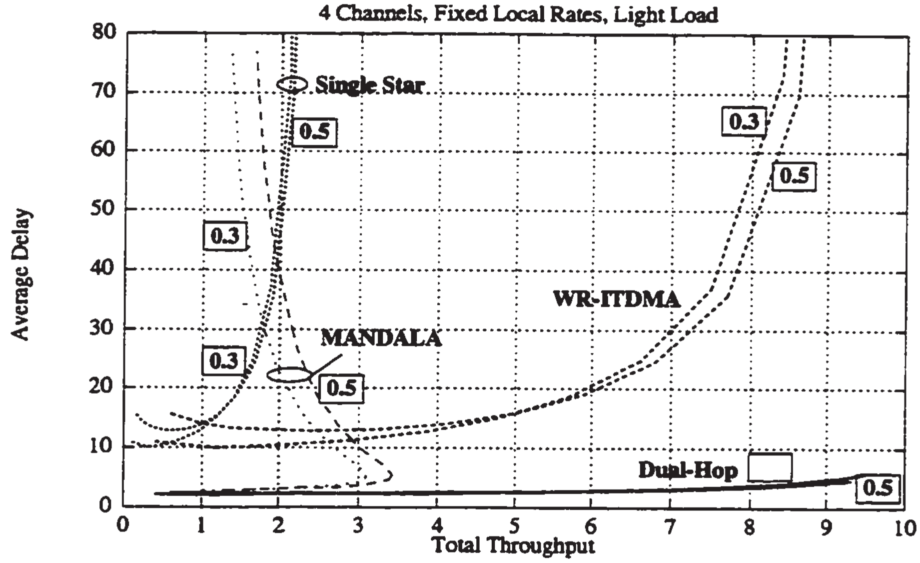


(a) 4 Channels, Light Inter-LON Loading

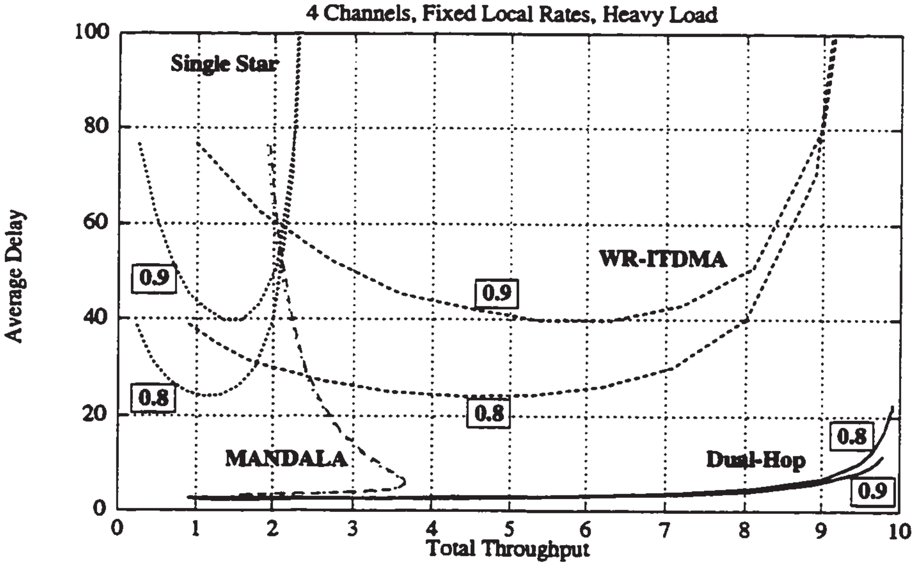


(b) 4 Channels, Heavy Inter-LON Loading

Figure 5.17: Non-Uniform Traffic Comparisons, Fixed Inter-LON Traffic. The inter-LON flow rates per channel are shown in the boxes adjacent to the curves.



(a) 4 Channels, Light Local Loading



(b) 4 Channels, Heavy Local Loading

Figure 5.18: Non-Uniform Traffic Comparisons, Fixed Local Traffic. The local flow rates per channel are shown in the boxes adjacent to the curves.

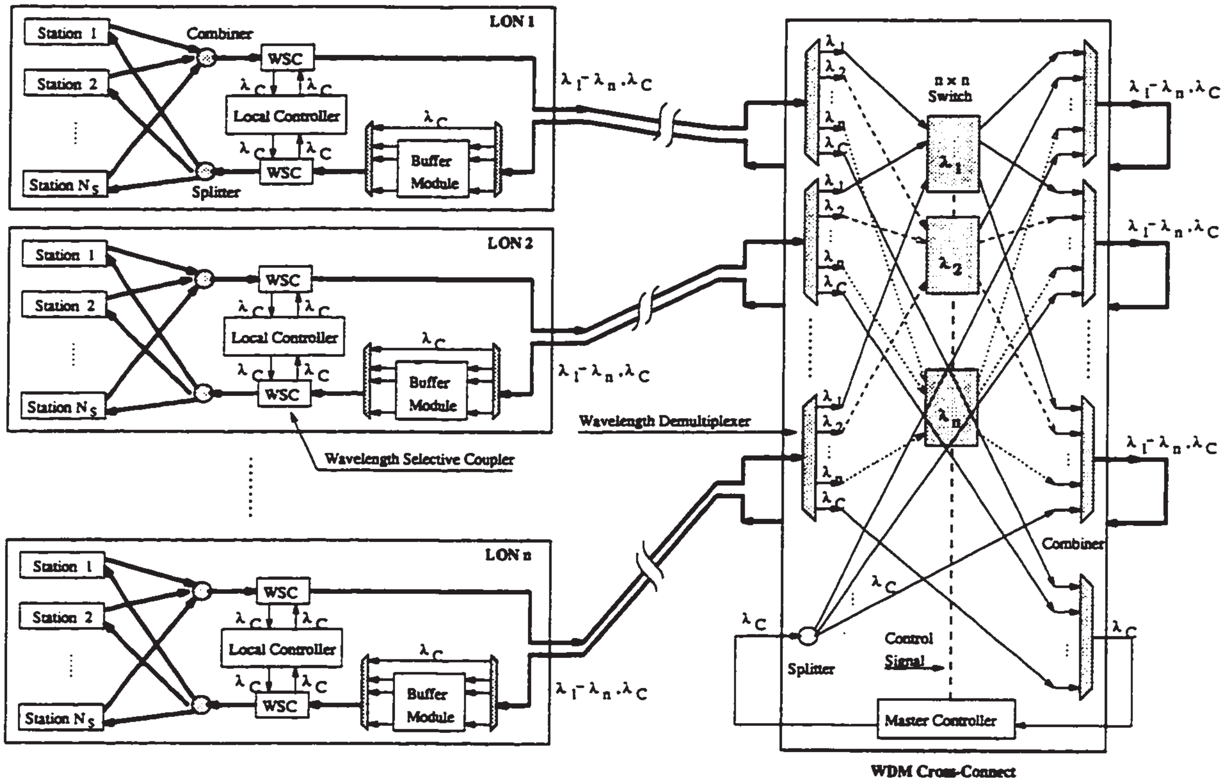


Figure 5.19: Dynamic bandwidth allocation in dual-hop system using two level control.

MC. The corresponding paths are also shown in Figure 5.20. The stations in each LON send their requests to the LC associated with that LON. The LC then allocates the available bandwidth to different communications based on the traffic demands of the stations in its LON. This task is performed by assignment of the time slots in the data frames.

5.7.1 Local Control

As mentioned above, the control channel is spatially reused in each LON. In each LON, the stations send their bandwidth requirements in a pre-assigned time slot in a control frame on the control channel to the LC of that LON. Each station simply reports the number of data slots that it needs in the next data frame for different destination channels (LONs). Table 5.1 shows the request frame structure. As shown,

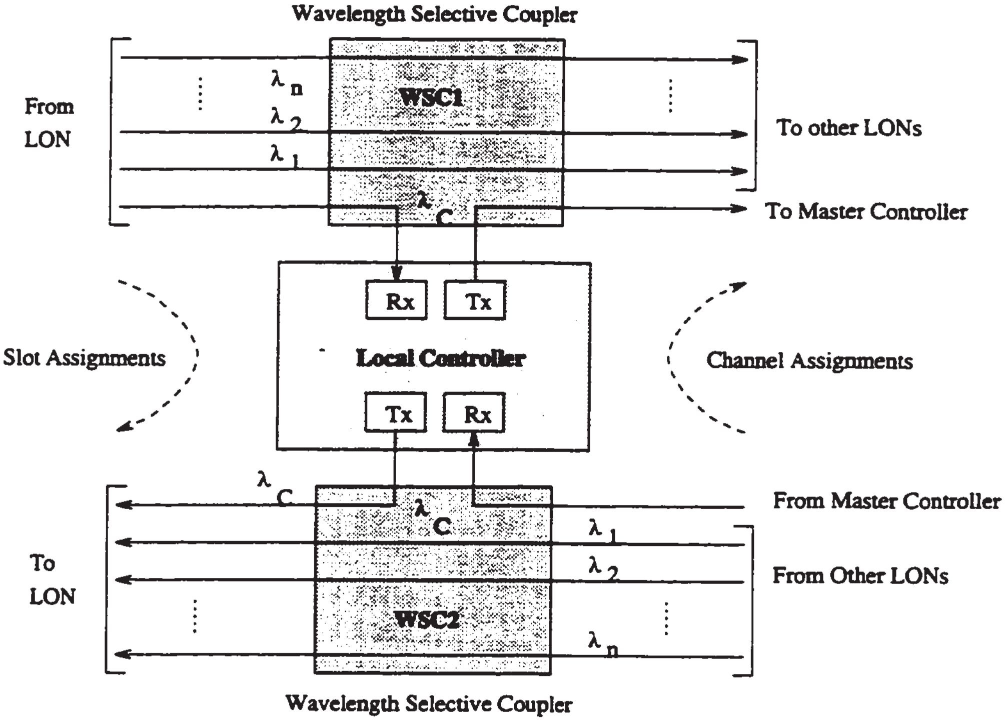


Figure 5.20: Local Controller (LC) design.

a TDMA scheduling is used for access of the stations to the control channel. A group of n control slots are dedicated to each station, where n is the number of available channels. In this table, ν_{ij} is the number of data slots that station i needs for its packets destined to LON j $(i, j = 1, \cdots, n)$. Recall that in the dual-hop architecture the destination channel corresponds to the destination LON.

The LC sends the time slot assignments to the stations of its LON using the allocation frame shown in Table 5.2. Each frame is divided into n groups of allocation slots. Each group corresponds to a destination channel. It is assumed that each data frame has p data slots. In the allocation frame, s_{ij} shows the station that is assigned to the j^{th} time slot of the next data frame on channel i, where $i = 1, \cdots, n$, $j = 1, \cdots, p$, and $1 \leq s_{ij} \leq N_s$. D_k shows the destination LON connected by channel k $(k = 1, \cdots, n$, and $1 \leq D_k \leq n)$. These channel assignments are done by the MC and the LC reports them to the stations in its LON.

Station 1				Station 2					Station N_s			
ν_{11}	ν_{12}	$\cdots$	ν_{1n}	ν_{21}	ν_{22}	$\cdots$	ν_{2n}	$\cdots$	ν_{N_s1}	ν_{N_s2}	$\cdots$	ν_{N_sn}

Table 5.1: Request frame from stations to local controller (LC) on λ_C.

Channel 1					Channel 2						Channel n				
D_1	s_{11}	s_{12}	$\cdots$	s_{1p}	D_2	s_{21}	s_{22}	$\cdots$	s_{2p}	$\cdots$	D_n	s_{n1}	s_{n2}	$\cdots$	s_{np}

Table 5.2: Allocation frame from local controller (LC) to stations on λ_C.

5.7.2 Master Control

Each LC sends its estimated channel bandwidth requirements for inter-LON communications to the MC by accessing λ_C through WSC1 as shown in Figure 5.20. The request and report frame shown in Table 5.3 is used for this purpose. As shown, each LC also reports the packet loss level in its LON to the MC. In the request and report frame, κ_{ij} denotes the estimate of the LC of LON i for the number of channels (wavelengths) needed for the transmissions from LON i destined to LON j, and $Drop_i$ indicates the recent packet loss level in LON i. The packet loss indication can be used in the global channel allocation, and also as a feedback from the destination LONs. This measure might be used in the local traffic scheduling to avoid catastrophic packet losses.

The MC assigns channels to the different inter-LON links based on the requests from all the LONs, and sends the allocation table to the LCs on λ_C. The allocation frame used is shown in Table 5.4. The MC changes the inter-LON channel allocations by controlling the $n \times n$ switches in the WDM cross connect. The MC also relays the packet loss reports received from each LC to all of the LCs, so that they can use this data in their media access control to reduce packet loss in a specific LON by reducing the rate of data transmissions destined to that LON. In Table 5.4, $Drop_i$ reports the recent packet loss level in LON i, and l_{ij} indicates the destination LON connected to LON i through channel j.

The optical switching modules in the central WDM cross connect can be removed if the architecture shown in Figure 5.21 is used. In this architecture, there is no need for dynamic switching control in the WDM cross connect and all the packets are filtered before entering the outbound links. The dynamic allocation protocol

Local Controller 1					*Local Controller 2*						*Local Controller n*				
κ_{11}	κ_{12}	$\cdots$	κ_{1n}	$Drop_1$	κ_{21}	κ_{22}	$\cdots$	κ_{2n}	$Drop_2$	$\cdots$	κ_{n1}	κ_{n2}	$\cdots$	κ_{nn}	$Drop_n$

Table 5.3: Request and report frame from Local Controllers (LC) to Master Controller (MC) on λ_C.

Packet Loss Report				*LON 1*				*LON 2*					*LON n*			
$Drop_1$	$Drop_2$	$\cdots$	$Drop_n$	l_{11}	l_{12}	$\cdots$	l_{1n}	l_{21}	l_{22}	$\cdots$	l_{2n}	$\cdots$	l_{n1}	l_{n2}	$\cdots$	l_{nn}

Table 5.4: Channel Allocation frame from Master controller (MC) to Local Controller (LC) on λ_C.

performed by the MC and the LCs ensure collisionless connections.

5.7.3 Allocation Protocol

The dynamic time slot assignment (TSA) used by the LCs is described in this section. We will show that the proposed TSA technique is simple and cannot be a potential bottleneck.

When LC receives requests, it performs the TSA on different channels based on the received station demands. The resulting TSA is then sent to the stations. In order to simplify the descriptions, we continue with an example.

Assume 4 channels are available and there are 4 LONs. In our example, we assume that there is one station per channel, hence 4 stations per LON. For this example, the resulting frame is 4 × 4, in which rows correspond to channels and columns to time slots. The element in row i and column j of a frame denotes the station that is assigned to time slot j on channel i. The LC assigns the time slots available on different channels to the stations in a way that each station must at most access one channel in each time slot. This is necessary since each station has only one tunable transmitter. Since the TSA scheme is the same for all LONs, we consider one of the LCs, say at of LON 1. We also assume that channel i ($i = 1, 2, 3, 4$) connects LON 1 to LON i.

Consider the following request frame.

3	1	0	0	0	0	1	2	0	3	1	0	4	0	0	0

Using the request format frame shown in Table 5.1, the demand table can be easily

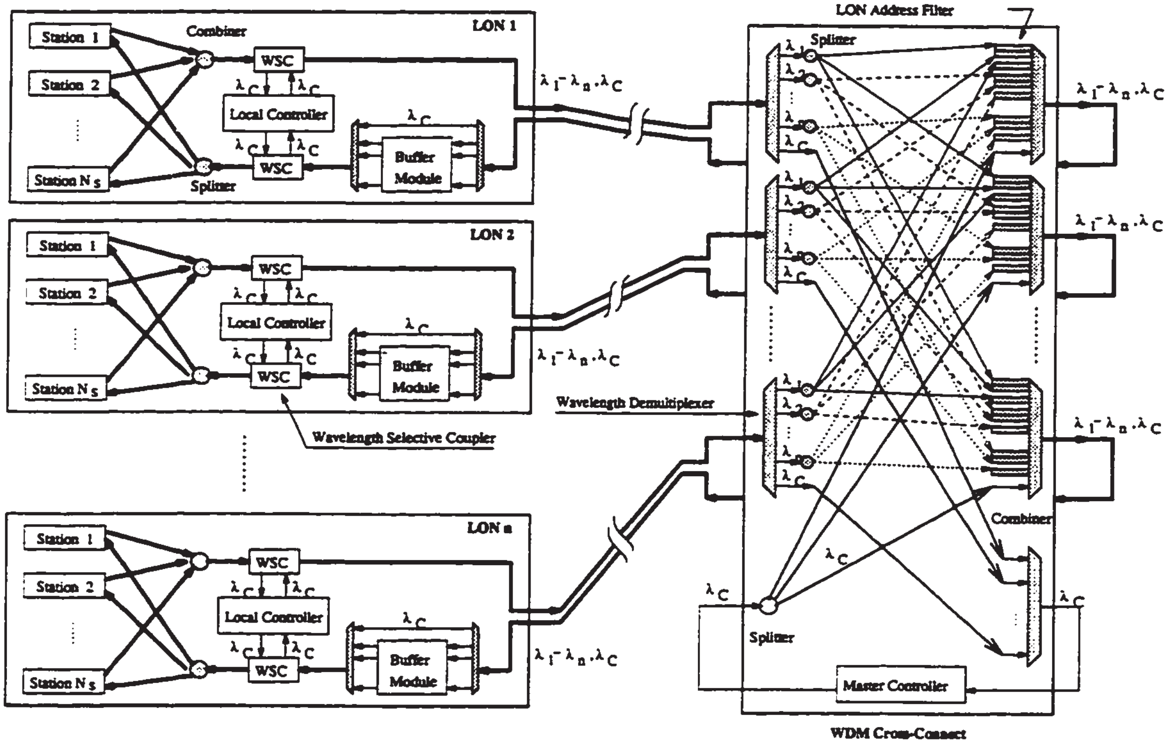


Figure 5.21: Dynamic bandwidth allocation in dual-hop system using two level control. The LON address filtering replaces the optical switches.

formed.

Demand Table (initial)

Destination LON:	1	2	3	4
Station: 1	3	1	0	0
2	0	0	1	2
3	0	3	1	0
4	4	0	0	0

In this table, the number 3 in the first row shows that station 1 requested 3 slots on the channel connected to LON 1, 1 slot for LON 2, and no time slots on the channels connected to LONs 3 and 4. Rows 2, 3, and 4 are showing the demands of stations 2, 3, and 4, respectively.

For the TSA, we start by checking each entry of this table. In order to obtain a fair assignment, we assign one time slot in each scan of the table as follows. Scanning the table for non-zero entries, we assign one time slot on the requested channel to the requesting station and decrement the corresponding entry in the demand table by one. In this procedure, only one channel can be assigned to each station in each time slot, since each station is equipped with a single transmitter. Processing the above demand table we find these results.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4			⟶ LON 1
2	3	1			⟶ LON 2
3	2	3			⟶ LON 3
4		2			⟶ LON 4

Demand Table (after step 1)

Destination LON:	1	2	3	4
Station: 1	2	0	0	0
2	0	0	0	1
3	0	2	0	0
4	3	0	0	0

We continue this procedure until all the time slots are assigned, all the demands are satisfied, or no more assignments can be done, whichever occurs sooner. If we carry on with the TSA for this example the result of the second step (scan) is as follows.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4	1	4	⟶ LON 1
2	3	1	3		⟶ LON 2
3	2	3			⟶ LON 3
4		2	2		⟶ LON 4

Demand Table (after step 2)

Destination LON:	1	2	3	4
Station: 1	1	0	0	0
2	0	0	0	0
3	0	1	0	0
4	2	0	0	0

The third step gives the following frame.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4	1	4	⟶ LON 1
2	3	1	3	3	⟶ LON 2
3	2	3			⟶ LON 3
4		2	2		⟶ LON 4

Demand Table (after step 3)

Destination LON:	1	2	3	4
Station: 1	1	0	0	0
2	0	0	0	0
3	0	0	0	0
4	2	0	0	0

No more assignments on demand can be done after this step. The unassigned time slots can be allocated to those stations that have not been assigned to other channels in the corresponding time slot. The final frame is shown below.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4	1	4	⟶ LON 1
2	3	1	3	3	⟶ LON 2
3	2	3	4	1	⟶ LON 3
4	4	2	2	2	⟶ LON 4

The LC sends the final frame to the stations in its LON. It uses the allocation frame format shown in Table 5.2 for this purpose. For the above example, the allocation frame is as follows.

1	1	4	1	4	2	3	1	3	3	3	2	3	4	1	4	4	2	2	2

The numbers in the dotted boxes show the destination LON of each channel. The unfulfilled demands have to be repeated by stations in the subsequent frames.

The LC sends its estimates of the number of channels that its LON needs for inter-LON communications to the MC. The MC allocates the connection channels for inter-LON traffic based on the demands from all the LCs. We will not discuss the protocols used by the MC for bandwidth allocation, since this allocation strategy can be based on well known techniques. It is worthwhile to mention however, that this bandwidth allocation can be performed on a long time basis based on the statistical characteristics of the network traffic.

As another example, we consider the case that two channels connect LON 1 to

LON 2 (say channel 2 and 4), and one channel connects LON 1 to LON 3 (say channel 3) and the other one (channel 1) links LON 1 to itself. This means that the traffic going from LON 1 to LON 2 is high and there is no traffic from LON 1 destined to LON 4. The request frame below is an example of this.

1	3	1	0	0	2	1	0	0	3	1	0	1	1	0	0

The demand table can be easily obtained from the request frame.

Demand Table (initial)

Destination LON:	1	2	3	4
Station: 1	1	3	1	0
2	0	2	1	0
3	0	3	1	0
4	1	1	0	0

The TSA frame after the first step of the TSA step is shown below.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4			⟶ LON 1
2	2	1	3	4	⟶ LON 2
3	3	2	1		⟶ LON 3
4					⟶ LON 2

Demand Table (after step 1)

Destination LON:	1	2	3	4
Station: 1	0	2	0	0
2	0	1	0	0
3	0	2	0	0
4	0	0	0	0

The result of the second (last) step is as follows.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4			⟶ LON 1
2	2	1	3	4	⟶ LON 2
3	3	2	1		⟶ LON 3
4		3	2	1	⟶ LON 2

Demand Table (after step 2)

Destination LON:	1	2	3	4
Station: 1	0	1	0	0
2	0	0	0	0
3	0	1	0	0
4	0	0	0	0

Finally, the TSA results in the following frame.

TSA Frame

Time Slot:	1	2	3	4	
Channel: 1	1	4	4	2	⟶ LON 1
2	2	1	3	4	⟶ LON 2
3	3	2	1	3	⟶ LON 3
4	4	3	2	1	⟶ LON 2

The LC sends the TSA frame to the stations in an allocation frame as follows.

1	1	4	4	2	2	2	1	3	4	3	3	2	1	3	2	4	3	2	1

Again the numbers in the dotted boxes show the destination LON of each channel. As shown, there are two channels destined to LON 2 and there is no channel connecting LON 1 to LON 4.

5.8 Station-to-station and LON-to-LON Capacities

Station-to-Station channel capacity ($\tilde{C}$) defines the maximum channel bandwidth that can be allocated to the communication between two stations in the network. Generally this capacity depends on the network architecture and the medium access control protocol used.

As seen in Section 5.8, WR-ITDMA can provide $1/nN_s$ for the station-to-station capacity in a WDM cross connect network, where there are n wavelength channels and N_s stations per LON (Equation 4.20). The dual-hop architecture decouples the local access from the global access control. Thus, even if we use ITDMA for the L-MAC, the frame length is $1/n$ times that for the general case of WR-ITDMA. Consequently, the station-to-station capacity in this case is

$$\tilde{C}_{dual-hop} = \frac{1}{N_s}. \tag{5.5}$$

Dynamic bandwidth allocation techniques introduced in the previous sections allow us to ramp up this capacity to 1. The Local Control part of the dynamic bandwidth allocation enables us to provide each station with the full bandwidth of a channel to communicate with another station in any LON. Thus,

$$\tilde{C}_{Local\ Control} = \frac{T_{DF}}{\max(T_{DF}, T_{CF})}, \tag{5.6}$$

where T_{DF} and T_{CF} are the data and control frame durations, respectively. The reservation operation is performed on a frame basis, where each frame duration is N_s data packets. It can easily be seen that it is very unlikely that the control frame duration will be larger than the duration of a data frame, since $T_{CF} \ll T_{DF}$. Therefore, $\frac{T_{DF}}{\max(T_{DF}, T_{CF})} = 1$. Consequently, we can write

$$\tilde{C}_{Local\ Control} = 1. \tag{5.7}$$

Master Control, on the other hand, assists in allocating bandwidth to LON-to-LON communications. In the single-hop and dual-hop WDM cross connect architectures, the LON-to-LON capacity is limited to the number of channels connecting two LONs. In the two level dynamic control techniques described for the dual-hop architecture, it is possible to assign all the available data channels for communications between two LONs. As a result, the LON-to-LON channel capacity can be as high as n.

Table 5.5 summarizes all the station-to-station, LON-to-LON, and total capacities for the proposed architectures and protocols. We assume that there are n channels and N_s stations per LON. The control frame duration T_{CF} is also assumed to be much less than the duration of a data frame T_{DF} ($T_{CF} \ll T_{DF}$ and $\frac{T_{DF}}{\max(T_{DF}, T_{CF})} = 1$).

WDM Cross Connect		Capacity		
Architecture	Access Method	Station-to-Station	LON-to-LON	Total
Single-Hop	WR-ITDMA	$1/nN_s$	1	$\min(n^2, N_s^2)$
Single-Hop	dynamic TSA	1	1	n^2
Single-Hop	First Local Random	$1/(ne)$	$1/e$	n^2/e
Single-Hop	Second Local Random	$1/n$	1	n^2
Dual-Hop	ITDMA for L-MAC	$1/N_s$	1	n^2
Dual-Hop	Dynamic Local	1	1	n^2
Dual-Hop	Dynamic Local & Master	1	n	n^2

Table 5.5: Station-to-Station, LON-to-LON and total capacities for single-hop and dual-hop WDM cross connect architectures.

5.9 System Scalability

The dual-hop architecture is considered to be *scalable*. Figure 5.22 shows a *triple-hop* architecture, where an extension to the dual-hop cross connect network is illustrated. As shown in the figure, n ($n = 4$ in the figure) of the LONs are amalgamated as a LON-cluster, and two levels of wavelength routing and shared buffering are used.

The same connection table such as the one shown in Table 4.2 can be used in all the WDM cross connects. Each station in each LON uses its transmitter agility to select one of the channels with respect to the LON-cluster of the destination station. The data packet sent is then wavelength routed to the proper LON-cluster switching section. In this switching stage, the wavelength of the received signal is converted to the wavelength corresponding to the destination LON. Finally, in the LON switching section, the wavelength of each incoming signal is translated to the home channel of the destination station. In the system with 4 wavelengths ($n = 4$) shown in Figure 5.22, if any station in LON 2 of LON-cluster 1 wants to communicate to station 2 of LON 3 of LON-cluster 1, it must send its data packet on λ_n (λ_4 in the example). This is because of the fact that all the stations in LON 2 of LON-cluster 1 are connected to LON-cluster 1 through λ_n (λ_4 in the example). In the first level of wavelength routing, this packet is wavelength converted and switched to λ_3 to be sent to LON 3 of this LON-cluster. This packet will be converted to λ_2 in the second switching level to be received by station 2 in LON 3 whose fixed-tuned receiver is tuned to HOME channel λ_2.

It is obvious that very high capacities on the order of n^3 are obtainable. However,

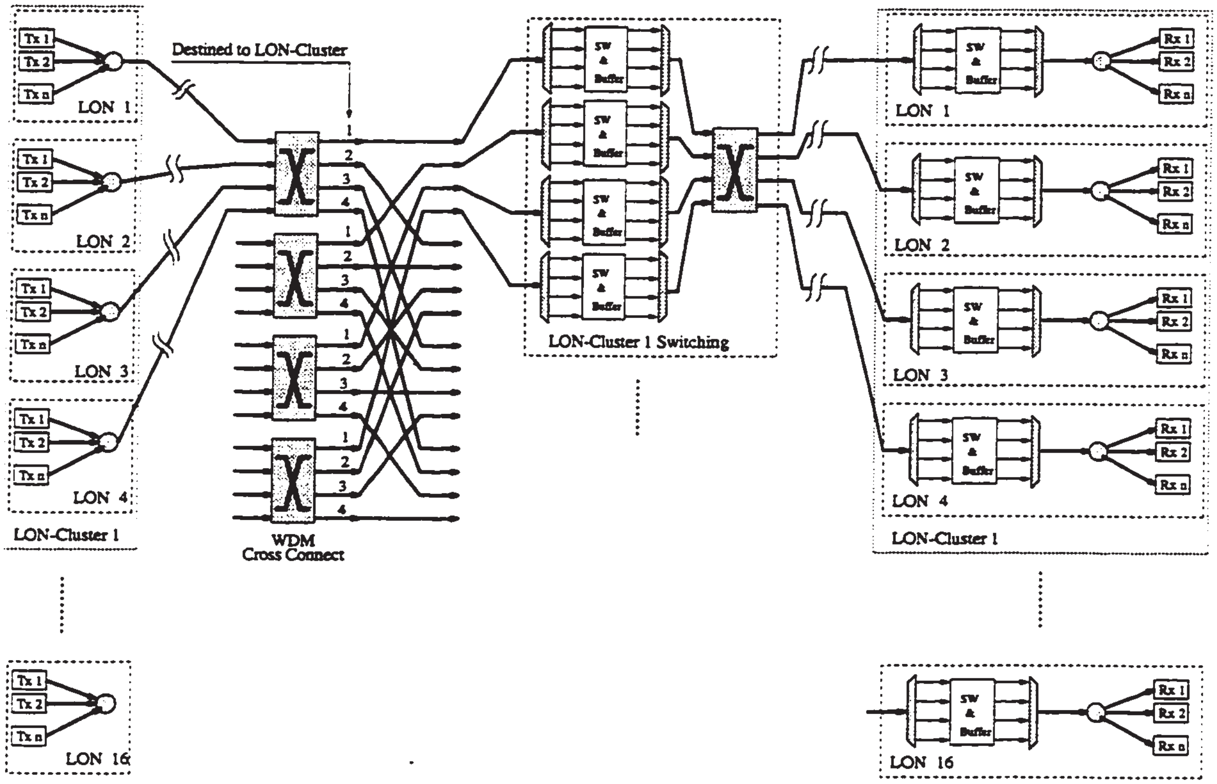


Figure 5.22: Triple-Hop Architecture. Two levels of wavelength routing and switching are used.

we should note that the station-to-station channel capacity is $1/n^{th}$ that of the dual-hop case, and higher levels of packet loss would be likely. Our intent in this section is to show that architectural scaling of the proposed architectures is possible to design a WDM cross connect network for the case where the number of stations is very large and only very few wavelength channels are available.

Architecture	Figures	Transmitter		Receiver	
		Tunable	Fixed	Tunable	Fixed
Single-Hop	4.3	nN_s	0	nN_s	0
Dynamic Single-Hop	4.11	nN_s	nN_s	nN_s	nN_s
Dual-Hop with WC	5.1	nN_s	0	0	nN_s
Dual-Hop without WC (1)	5.2	nN_s	0	0	nN_s
Dual-Hop without WC (2)	5.3	nN_s	0	0	nN_s
Dynamic Dual-Hop	5.19	nN_s	$nN_s + 2n$	0	$2nN_s + 2n$

Table 5.6: Component Counts for Proposed Designs: Part (a).

Architecture	DeMux	Combiner	Splitter	WSC	Buffer Module(*)
Single-Hop	n	$2n$	n	0	0
Dynamic Single-Hop	n	$3n$	n	n	0
Dual-Hop with WC	$2n$	$3n$	n	0	n
Dual-Hop without WC (1)	n	$2n$	n	0	n
Dual-Hop without WC (2)	n	$2n$	$n^2 + n$	0	n
Dynamic Dual-Hop	$2n$	$3n + 1$	$n + 1$	$2n$	n
(*) Refer to Tables 5.9 and 5.10.					

Table 5.7: Component Counts for Proposed Designs: Part (b).

Architecture	Optical Switch	Others
Single-Hop	0	
Dynamic Single-Hop	0	
Dual-Hop with WC	0	
Dual-Hop without WC (1)	n^2 of $(1 \times n)$	
Dual-Hop without WC (2)	0	n^2 Address Check Circuits, n^3 LON Address Filter
Dynamic Dual-Hop	n of $(n \times n)$	Master and Local Controllers

Table 5.8: Component Counts for Proposed Designs: Part (c).

5.10 Component Counts

In this section, we present the component counts for the different architectures proposed in this chapter. Before making any conclusions, we note that even though the count for a specific component may be less in one design than in another, it does not necessarily mean that the former design is superior. The reduction of some hardware components may result in some level of complexity in the system operation that can not be quantized. The reliability of the architecture can be viewed as a function of the component counts of that system. However, the reliability of the entire system and its operation is obviously not dependent on the component counts alone.

The component counts for the proposed architectures are shown in Tables 5.6,

Buffer Module	Figures	Fixed Tx	Fixed Rx	WC	DeMux	Combiner
Shared Electronic	5.5	n	n	0	0	0
Non-Shared Electronic	5.6	n	n^2	n^2	n^2	0
Optical Shared MDLB	5.11(a)	0	0	n^2	n^2	m
Optical Shared SDLB	5.11(b)	0	0	n^2	0	m

Table 5.9: Component Counts for Buffer Modules: Part (a).

Buffer Module	Optical Switch	Memory Blocks
Shared Electronic	0	1 $n \times n$ ATM Switch/Buffer
Non-Shared Electronic	0	n^2 $n \times 1$ Memory Block (Concentrator)
Optical Shared MDLB	m	n Shared MDLB
Optical Shared SDLB	m	n Shared SDLB

Table 5.10: Component Counts for Buffer Modules: Part (b).

5.7, and 5.8. Buffer modules are used in the dual-hop architectures. The component counts for different buffer modules are shown in Tables 5.9 and 5.10.

In the tables, it is assumed that there are n wavelength channels and N_s stations per LON. m shows the maximum length of the delay line buffer for the shared delay line buffer (SDLB), as well as the number of delay lines in the multiple delay line buffer (MDLB). DeMux refers to the wavelength demultiplexer. WSC stands for wavelength selective coupler. Tx and Rx are short for transmitter and receiver, respectively. WC stands for wavelength converter (together with its control circuit). Finally, dual-hop without WC (wavelength conversion) (1) refers to the dual-hop architecture without wavelength conversion shown in Figure 5.2, and dual-hop without WC (2) represents the architecture shown in Figure 5.3, as also indicated in Table 5.6.

As seen in the tables, the dual-hop architectures have added some components to the single-hop designs. This is the price paid to simplify the single-hop system operation. It is apparent that we can not easily quantize the effect of this simplification and its consequences on the dynamic adaptability of the resulting design. However it can be conjectured that the reliability of the single-hop architecture is higher than that of the dual-hop option.

5.11 Summary

The dual-hop architectures presented in this chapter introduce cluster shared buffering in WDM cross connect networks. The basic single-hop WDM cross connect architecture presented in Chapter 4 is used as the basic building block. LON shared buffering techniques are added to simplify network control and station hardware design. These schemes also eliminate the need for station receiver tunability, relax synchronization requirements, decouple local control, and achieve a very high throughput while maintain very low access delays.

Different architectural variations of the dual-hop architectures are presented and compared in terms of the ease of operation, packet loss rates, and component counts. Optical processing techniques are used to expedite the buffering and switching operations. In the almost all-optical buffering module design, it is shown that optical delay line buffers can replace the electronic buffers. Application of a wavelength sharing technique results in a drastic decrease in buffering requirements.

Finally, a dynamic two-level bandwidth allocation architecture and its corresponding control protocols are proposed. In this approach, the bandwidth allocation is performed based on collected traffic demands of the stations. The local controllers are responsible for the LON access control, while the master controller performs the inter-LON channel bandwidth allocation.

Experiments are performed which favorably compare the dual-hop system performance with other architectures.

As the final note, we should mention that the field of optical networks is a unique area that can be advanced by combining both systems and device knowledge. The interested reader is referred to the ample references provided in the thesis for more information regarding the optical device technology and its state-of-the-art.

Chapter 6

Conclusions

In the future, optical components will permit the construction of LANs using parallel channel architectures based on wavelength division multiplexing (WDM). However, the number of usable wavelengths may be less than desired. As a result, designs which permit wavelength reuse may be required. One excellent option in this case is to perform media access across a spatial wavelength cross connect. In such a system, the total capacity may be as much as n times that of a single passive star network, where n is the number of available wavelengths. Unfortunately, previous designs tend to be complex since stations have to coordinate the operation of both tunable transmitters and receivers.

Wavelength cross connect networks have been considered in this thesis. In the proposed designs, the network is clustered into local optical networks (LONs). Each LON includes a set of stations which are located in close physical proximity. The WDM cross connect is then used to interconnect these LONs.

The networks considered are classified into two categories: single-hop and dual-hop. A number of medium access protocols and several architectural variations of these systems were proposed. In each case, the resulting design was compared to other appropriate architectures to highlight the advantages of the proposed systems.

In single-hop WDM cross connect networks, stations use the wavelength agility of their transmitters to wavelength-route their transmissions to the destination LON. At the receiving side, a tunable receiver separates the proper channel and receives data packets destined to itself. Extensions to the techniques used in satellite communication systems were introduced. These schemes, apply dynamic time slot assignments (TSAs) based on collected station traffic demands (i.e. the network traffic matrix). New results which show the characteristics and bounds on the performance of the scheduling techniques were developed. Since fully optimized scheduling for single-hop WDM cross connect networks are shown to be computationally intractable, three heuristic methods were proposed. Using computer simulation, the performance of these three methods were compared in terms of various performance measures. The simulations show that the proposed methods are efficient and fast. Finally, random local medium access control protocols were introduced that can be used where faster operation of the system is required.

In the dual-hop cross connect networks, cluster shared buffering techniques are incorporated into the network architecture. In these networks, as in the single-hop case, tunable light sources are used at the transmitting side. Again stations wavelength-route their transmissions to a destination LON. At the output section of each LON, packets are buffered and subsequently transmitted onto the desired destination wavelength. There are a number of advantages to this design including the fact that receiver tunability is not needed. In addition, this type of network can exploit the availability of inexpensive electronic ATM buffering components. Simulation and analytical results show significant improvements in network throughput compared to other architectures proposed in the literature. A two level dynamically controlled architecture was proposed to enable the dual-hop architecture to adapt to different network traffic conditions.

In the thesis, it was shown that the dual-hop designs can accommodate the use of almost all optical buffering techniques. A technique is proposed which reduces the buffering requirements at the LONs using wavelength sharing. Results are presented that consider the loss performance of the network for different loading conditions. A simple delay model was presented for estimating network performance under uniform loading.

Some of the design advantages over other architectures proposed in the literature are summarized in the next section. In the last part of this chapter some of the extensions and fields of research related to this thesis are listed as future work.

6.1 Design Advantages

Wavelength division multiplexing, wavelength reuse, wavelength routing, network clustering and shared buffering are used in the architectures proposed in this thesis to achieve improvements in network capacity. The above techniques together with two stage contention resolution using local media access control (L-MAC) and cluster shared buffering, enables us to

- simplify the station hardware design,
- simplify the media access control,
- reduce access delay,
- increase the total throughput,
- simplify synchronization requirements,
- have better resource sharing,
- apply transparent and flexible local medium access control protocols,
- decouple the cluster control from global network control,
- use two-level dynamic bandwidth allocation,
- and use commercial ATM switch/buffer components.

The introduction of multi-channel shared optical delay-line buffering helps to

- drastically reduce the buffering requirements,
- reduce the component counts,
- remove potential bottlenecks in the optical path,

- use wavelength channel distinction as an effective routing method,
- and realize all-optical wavelength routing architectures.

In the single-hop WDM cross connect networks, the optical path designs are considered to be as simple as possible, while all the control and routing is being done at the edges of the network. Efficient scheduling techniques are introduced for this class of wavelength cross connected networks. In the dual-hop architectures, cluster shared buffering is proposed which enables more optical processing and consequently very simple control protocols.

6.2 Future Work

This thesis motivates some new areas of research. A number of these are identified and discussed.

- **Local Medium Access Control Protocols:** Dual-hop architectures decouple local (intra-LON) and inter-LON control. Each LON can thus use a different L-MAC from those used in other LONs. These local protocols resolve contention on the transmitting side. It may be possible to define a set of L-MACs which contain different protocols for different loading conditions. The local traffic management agent, then selects one of the known protocols from the set and sends the initialization signal to all the stations in that LON. By applying this method, the LON can swiftly adapt itself to the current traffic pattern.
- **Better Analytical Models:** Simple approximate analytical models are used in this thesis to investigate dual-hop network performance. These models are approximate in terms of neglecting the interaction of the two stages of the network, i.e. local stage and buffering section. More mathematical elaboration would be helpful to find better analytical models which will describe the exact operation of the network.
- **More Scheduling Techniques:** As shown in Chapter 4, the time slot assignment problem for WDM cross connect networks is much more complicated

that for SS/TDMA systems. More research is needed to find some of the exact bounds for the scheduling schemes. It is possible that heuristic methods could be found to achieve these optimum measures.

- **Pipelined Distributed Retuning:** In Section 4.7, we discussed the idea of pipelining the transmitter and receiver tuning times in single-hop WDM cross connect networks. We have observed that if this technique is used together with the proposed time slot assignment (TSA) methods, the improvement in the overall transmission time is negligible. This raises the question that if the pipelining were incorporated into the TSA method from the beginning, whether more efficient scheduling schemes could be found.

Appendix A

Analytical Models

In this loss model we will assume the non-shared electronic buffering case, which is equivalent (in loss performance) to the two optical buffering cases introduced in Section 5.5. Assuming that packet arrivals are independent and memoryless on each LON buffer input, the resulting system is a very simple one and can be modeled by a finite-state discrete-time Markov process [Hay86]. Figure A.1 shows the state diagram for this simple model. We define X_i to be the state of i buffers in use. There are $(m+1)$ states in the Markov chain, with states $X_0, X_1, \ldots, X_m$ corresponding to the use of each delay-line or delay-step and state X_m corresponds to a buffer full state. We define p as the probability of packet arrival on one of the inputs. Since the inputs are i.i.d, the probability of k $(k = 0, 1, \ldots, n)$ simultaneous incoming packets on k input ports is binomial and given by $p_k = \binom{n}{k} p^k (1-p)^{n-k}$.

It is also assumed there are only m buffers (or m delay lines) in each buffer module. Thus if the system is in state m, all buffers are in use, and only one packet can be queued in this frame cycle, assuming that one may also be sent out. Two or more packet arrivals on multiple ports to the buffer-module cause packet loss. Similarly, if three or more packets are received while in state $(m-1)$, only two can be buffered and the remaining will be dropped.

The transition probability matrix $\mathbf{P}$, an $(m+1) \times (m+1)$ matrix, of this Markov

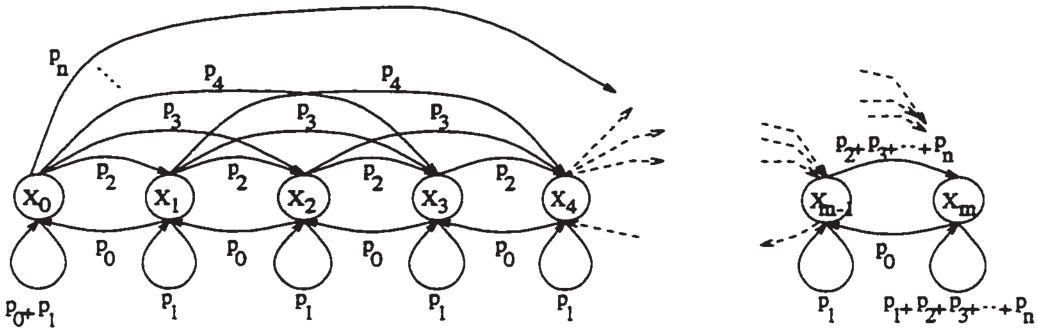


Figure A.1: State diagram of Markov chain corresponding to buffer-modules. p_i is the probability of i $(i = 0, 1, \ldots, n)$ simultaneous packets on input ports.

chain is given by

$$\mathbf{P} = \begin{bmatrix} p_0+p_1 & p_2 & p_3 & \cdots & 0 & 0 \\ p_0 & p_1 & p_2 & \cdots & 0 & 0 \\ 0 & p_0 & p_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{n-1} & p_n \\ 0 & 0 & 0 & \cdots & p_{n-2} & p_{n-1}+p_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_1 & p_2+p_3+\cdots+p_n \\ 0 & 0 & 0 & \cdots & p_0 & p_1+p_2+p_3+\cdots+p_n \end{bmatrix}. \tag{A.1}$$

This system may be solved in the usual way by solving the system of linear equations, $\Pi\mathbf{P} = \Pi = [\pi_0, \pi_1, \ldots, \pi_m]$ [Hay86, Pap91]. Using Π, the average number of dropped packets per slot ($\overline{N}_{dropped}$) can be written as

$$\begin{aligned} \overline{N}_{dropped} &= \pi_m[(n-1)p_n + (n-2)p_{n-1} + \cdots + 3p_4 + 2p_3 + p_2] \\ &\quad + \pi_{m-1}[(n-2)p_n + (n-3)p_{n-1} + \cdots + 2p_4 + p_3] \\ &\quad \cdots \\ &\quad + \pi_{m-(n-1)}[2p_n + p_{n-1}] \\ &\quad + \pi_{m-(n-2)}[p_n]. \end{aligned} \tag{A.2}$$

Note that in the above results, we assume that $\pi_j = 0$ when $j < 0$. The average number of arrivals per slot is simply $\overline{N} = np$, and therefore, packet loss ratio (P_{loss}) is given by

$$P_{loss} = \frac{\overline{N}_{dropped}}{\overline{N}} = \frac{\overline{N}_{dropped}}{np}. \tag{A.3}$$

A much simpler model was also introduced in Section 5.6. In this case, we assume that the ITDMA and buffering sections are independent and that packet arrivals to the buffering section follow a memoryless arrival process. The total delay, D_{total}, can be calculated as the sum of three components, namely local delay (ITDMA), D_{local}, one way propagation delay, D_{prop}, which is fixed, and finally the buffering delay, D_{buf}. Therefore

$$D_{Total} = D_{local} + D_{prop} + D_{buf}. \tag{A.4}$$

Referring to [Hay86], D_{local} is given by

$$D_{local} = \frac{\lambda_0 T_F^2 \overline{X^2}}{2(1-\rho)} - \frac{T_F}{2} + \overline{X} T_F + T, \tag{A.5}$$

in which $\overline{X}$ is the average message length equal to unity in our case. Similarly, $\overline{X^2} = 1$. T_F is the frame length, that is equal to n (number of channels) for n stations per LON. λ_0 is the arrival rate to the queues in each station, $\rho = \lambda_0 T_F \overline{X} = \lambda_0 n$, and finally T is packet transmission time ($T = 1$). Simplifying Equation A.5, we obtain

$$D_{local} = \frac{\lambda_0 n^2}{2(1-\rho)} + \frac{n}{2} + 1. \tag{A.6}$$

For the buffering section (without packet loss) with Poisson arrivals we employ the $M/D/1$ model. Since by assumption, we have n independent Poisson inputs with arrival rate λ_0, the total is also Poisson with arrival rate $\lambda = n\lambda_0$. The queuing delay can be written as ([Hay86])

$$D_{buf} = \frac{\overline{X}(2-\rho)}{2(1-\rho)}. \tag{A.7}$$

Here, the message length is fixed and equal to unity, therefore

$$D_{buf} = \frac{2-\rho}{2(1-\rho)}, \tag{A.8}$$

and $\rho = \lambda\overline{X} = \lambda$.

Bibliography

[AGKV88] E. Arthurs, M. S. Goodman, H. Kobrinski, and M.P. Vecchi. HYPASS: An optoelectronic hybrid packet switching system. *IEEE Journal on Selected Areas in Communications*, 6(9):1500–1510, December 1988.

[AT95] Nima Ahmadvand and Terence D. Todd. Dual-hop LANs using station wavelength routing. *Proceedings of the fourth International Conference on Computer Communications and Networks*, pages 622–629, 1995.

[AT97] Nima Ahmadvand and Terence D. Todd. Dual-hop LANs using a wavelength routing cross connect. *Computer Networks and ISDN Systems*, 1997. Accepted for publication.

[Bal95] V.K. Balakrishnan. *Network Optimization.* Chapman and Hall Mathematics, 1995.

[BCW81] G. Bongiovanni, D. Coppersmith, and C. K. Wong. An optimum time slot assignment algorithm for an SS/TDMA system with variable number of transponders. *IEEE Transactions On Communications*, COM-29(5):721–26, May 1981.

[BG92] Dimitri Bertsekas and Robert Gallager. *Data Networks.* Prentice Hall, second edition, 1992.

[Bra90] Charles A. Brackett. Dense wavelength division multiplexing networks: Principles and applications. *IEEE Journal On Selected Areas In Communications*, 8(6):948–964, August 1990.

[CSW89] F.R.K. Chung, J.A. Salehi, and V.K. Wei. Optical orthogonal codes: Design, analysis, and applications. *IEEE Transactions on Information Theory*, 35(3):595–605, May 1989.

[DBAP93] Patrick W. Dowd, Kalyani K. Bogineni, Khaled A. Aly, and James A. Perreault. Hierarchical scalable photonic architectures for high-performance processor interconnection. *IEEE Transactions on Computers*, 42(9):1105–1120, September 1993.

[DPM+93] T. Durhuss, R.J.S. Pedersen, B. Mikkelsen, K.E. Stubkjaer, M. Oberg, and S. Nilsson. Optical wavelength conversion over 18 nm at 2.5 Gb/s by DBR-laser. *IEEE Photonics Technology Letters*, 5(1):86–88, January 1993.

[DS94] P.W. Dowd and K.M. Sivalingam. A multi-level WDM access protocol for an optically interconnected parallel computer. *IEEE Infocom*, pages 400–408, 1994.

[EKY92] K.Y. Eng, M.J. Karol, and Y. Yeh. A growable packet (ATM) switch architecture: Design principles and appliations. *IEEE Transactions on Communications*, 40(2):423–430, February 1992.

[FV88] Gerard J. Foschini and Giovanni Vannucci. Using spread-spectrum in a high-capacity fiber-optic local network. *Journal of Lightwave Technology*, 6(3):370–379, March 1988.

[G+93] Paul E. Green et al. All-optical packet-switched metropolitan-area network proposal. *Journal of Lightwave Technology*, 11(5/6), May/June 1993.

[G+94] J. M. Gabriagues et al. Design, modeling and implementation of the ATMOS project fibre delay line photonic switching matrix. *Optical and Quantum Electronics*, 26:S497–S516, 1994.

[Gan92] Aura Ganz. Efficient algorithms for SS/TDMA scheduling. *IEEE Transactions On Communications*, 40(8):1367–1374, August 1992.

[GG92] A. Ganz and Y. Gao. A time-wavelength assignment algorithm for a WDM star network. In *Infocom'92*, pages 2144–2150. IEEE, 1992.

[GHJ94] J. Garcia-Haro and A. Jajszczyk. ATM shared-memory switching architectures. *IEEE Network*, pages 18–26, July/August 1994.

[GJ79] Micahel R. Garey and David S. Johnson. *Computers and Interactability, A Guide to the Theory of NP-Completeness.* W.H. Freeman and Company, 1979.

[GJP+91] K.S. Gilhousen, I.M. Jacobs, R. Padovani, A.J. Viterbi, L.A. Weaver, and C.E. Wheatley. On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40(2):303–312, May 1991.

[GJPW90] K.S. Gilhousen, I.M. Jacobs, R. Padovani, and L.A. Weaver. Increased capacity using CDMA for mobile satellite communicaitons. *IEEE Transactions on Selected Areas in Communications*, 8:503–514, May 1990.

[Goo89] Matthew S. Goodman. Multiwavelength networks and new approaches to packet switching. *IEEE Communications Magazine*, pages 27–35, October 1989.

[Gre93] Paul E. Green. *Fiber Optic Networks.* Prentice Hall, 1993.

[GW85] I.S. Gopal and C.K. Wong. Minimizing the number of switching in an SS/TDMA system. *IEEE Transactions on Communications*, COM-33(6):497–501, June 1985.

[GWK+92] B. Glance, J.M. Wiesenfeld, U. Koren, A.H. Gnauck, H.M. Presby, and A. Jourdan. High perforamnce optical wavelength shifter. *Electronics Letters*, 28(18):1714–1715, August 1992.

[Haa93] Z. Haas. The "Staggering Switch": An electronically controlled optical packet switch. *Journal of Lightwave Technology*, 11(5/6):925–936, May/June 1993.

[HAB94] D.K. Hunter, I. Andonovic, and P.E. Barnsley. Demonstration of optical timeslot interchanging through 2 × 2 crosspoints and feedforward delay lines. *Electronics Letters*, 30(11):875–876, May 1994.

[Hal67] Marshal Hall. *Combinatorial Theory*. Blaisdell Publishing Company, 1967.

[Hay86] Jeremiah F. Hayes. *Modeling and Analysis of Computer Communications Networks*. Plenum Press, 1986.

[HKS87] Isam M. I. Habbar, Mohsen Kavehrad, and Carl-Erik W. Sundberg. Protocols for very high-speed optical fiber local area networks using a passive star topology. *Journal Of Lightwave Technology*, 5(12):1782–1794, December 1987.

[HWM94] N. Huang, C. Wu, and G. Ma. A time-wavelength scheduling algorithm for interconnected WDM star networks. *IEEE Infocom*, pages 552–559, 1994.

[Inu78] Thomas Inukai. Comments on analysis of a switch matrix for an SS/TDMA system. *Proceedings Of The IEEE*, 66(12):1669–70, December 1978.

[Inu79] Thomas Inukai. An efficient SS/TDMA time slot assignment algorithm. *IEEE Transactions On Communications*, COM-27(10):1449–1455, October 1979.

[IUMY77] Y. Ito, Y. Urano, T. Muratani, and M. Yamaguchi. Analysis of a switch matrix for an SS/TDMA system. *Proceedings Of The IEEE*, 65(3):411–419, March 1977.

[JM93] Feiling Jia and Biswanath Mukherjee. The receiver collision avoidance (RCA) protocol for a single-hop WDM lightwave network. *Journal of Lightwave Technology*, 11:1053–1065, May/June 1993.

[JM95] Andrzej Jajszczyk and H.T. Mouftah. Tree-type photonic switching networks. *IEEE Network*, pages 10–16, January/February 1995.

[Jon88] William B. Jones. *Optical Fiber Communication Systems*. Holt, Rinehart and Winston, Inc., 1988.

[JT94a] M. Janoska and T. D. Todd. A single-hop wavelength-routed local area network. *IEEE International Conference on Applications of Photonics Technology*, June 1994.

[JT94b] Mark W. Janoska and Terence D. Todd. A simplified optical star network using distributed channel controllers. *Journal of Lightwave Technology*, 12(11):2011–2022, November 1994.

[KFG94] B. Kannan, Shivi Fotedar, and Maria Gerla. A two level optical star WDM metropolitan area network. *IEEE Globecom '94*, pages 563–566, December 1994.

[KK80] L. Kleinrock and F. Kamoun. Optimal clustering structures for hierarchical topological design of large computer networks. *Networks*, 10:221–248, 1980.

[KKYW92] K. Kondo, M. Kuno, S. Yamakoshi, and K. Wakao. A tunable wavelength-conversion laser. *IEEE Journal of Quantum Electronics*, 28(5):1343–1348, May 1992.

[KSG91] Mohsen Kavehrad, Ganti Sudhakar, and Nicolas Georganas. Slotted aloha and reservation aloha protocols for very high-speed optical fiber local area networks using passive star topolgy. *Journal Of Lightwave Technology*, 8(10):1411–1422, October 1991.

[LA93] J. P. Labourdette and A. S. Acampora. Logical clustering for the optimization and analysis of a rearrangeable distributed ATM switch. In *Infocom'93*, pages 464–472. IEEE, 1993.

[Law76] Eugene L. Lawler. *Combinatorial Optimization: Networks and Matroids.* Holt, Rinehart and Winston, 1976.

[LL95] K. Lee and V. O. K. Li. Optimization of a WDM optical packet switch with wavelength converters. *IEEE Infocom '95*, pages 423–430, 1995.

[Mat93] T. Matsumoto. Multiple-access optical network architecture employing a wavelength-and-network-divison technique: MANDALA. *Globecom '93*, pages 463–467, 1993.

[Meh90] Nader Mehravari. Improved multiple access schemes for very high speed optical fiber local area networks using a passive star topology. *Journal of Lightwave Technology*, 8:520–530, April 1990.

[Mid93a] John E. Midwinter. *Photonics in Switching*, volume II. Academic Press, 1993.

[Mid93b] John E. Midwinter. *Photonics in Switching*, volume I. Academic Press, 1993.

[Muk92a] Biswanath Mukherjee. The receiver collision avoidance (RCA) protocol for a single-hop WDM lightwave network. In *Proc., ICC '92*, June 1992.

[Muk92b] Biswanath Mukherjee. WDM-based local lightwave networks, part I: Single-hop systems. *IEEE Network Magazine*, pages 12–27, May 1992.

[Muk92c] Biswanath Mukherjee. WDM-based local lightwave networks, part II: Mulithop systems. *IEEE Network Magazine*, pages 20–32, July 1992.

[Onv94] Raif O. Onvural. *Asynchronous Transfer Mode Networks: Performance Issues.* Artech House Inc., 1994.

[Pal92] Joseph C. Palais. *Fiber Optic Communications.* Prentice Hall, 1992.

[Pap91] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill Inc., 1991.

[PP90] P.R. Prucnal and P.A. Perrier. Optically-processed routing for fast packet switching. *IEEE LCS Magazine*, pages 54–67, May 1990.

[Pr88] C.A. Pomalaza-raez. A note on efficient SS/TDMA assignment algorithm. *IEEE Transactions on Communications*, 36(9):1078–1082, September 1988.

[Ram93] Rajiv Ramsawami. Multiwavelength lightwave networks for computer communication. *IEEE Communications Magazine*, pages 78–88, February 1993.

[SAH94] T.N. Saadawi, M.H. Ammar, and A.E. Hakeem. *Fundamentals of Telecommunication Networks.* John Wiley and Sons, 1994.

[Sen92] John M. Senior. *Optical Fiber Communications Principles and Practice.* Prentice Hall, second edition, 1992.

[SHP91] John D. Spragins, Josdeph L. Hammond, and Krzysztof Pawlikowski. *Telecommunications Protocols And Design.* Addison-Wesley, 1991.

[SKG91] G.N.M. Sudhakar, M. Kavehrad, and N.D. Georganas. Multi-control channel very high-speed optical fiber local area networks and their interconnections using a passive star topology. *IEEE Globecom '91*, pages 624–628, 1991.

[SPE$^+$93] R. Schnabel, W. Pieper, M. Ehrhardt, M. Eiselt, and H.G. Weber. Wavelength convrsion and switching of high speed data signals using semiconductor laser amplifiers. *Electronics Letters*, 29(23):2047–2048, November 1993.

[ST91] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics.* Wiley, 1991.

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

[Sun89] Carl A. Sunshine. *Computer Network Architectures and Protocols.* Plenum Press, second edition, 1989.

[T+94] M. Tsukada et al. Experiments on photonic cell switching with an optical input buffer. *Electronics Letters*, 30(13):1081–2, June 1994.

[Tan89] Andrew S. Tanenbaum. *Computer Networks.* Prentice Hall, second edition, 1989.

[TG87] R.A. Thompson and P.P. Giordano. An experimental photonic time-slot interchanger using optical fibers as reentrant delay-line memories. *Journal of Lightwave Technology*, LT-5(1):154–162, January 1987.

[TKY+94] K. Takahata, K. Kasaya, H. Yasaka, Y. Kondo, and M. Ikeda. A unidirectional output optical frequency conversion device with an asymmetric-k dbr structure. *IEEE Journal of Quantum Electronics*, 30(5):1219–1226, May 1994.

[TNO85] S. Tamura, S. Nakano, and K. Okazaki. Optical code-multiplex transmission by gold sequences. *Journal of Lightwave Technology*, LT-3(1):121–127, February 1985.

[Van89] G. Vannucci. Combining frequency-division and code-division multiplexing in a high-capacity optical network. *IEEE Network*, pages 21–30, March 1989.

[Vit95] Andrew J. Viterbi. *CDMA: Principles of Spread Spectrum Communication.* Addsion-Wesley Publishing Company, 1995.

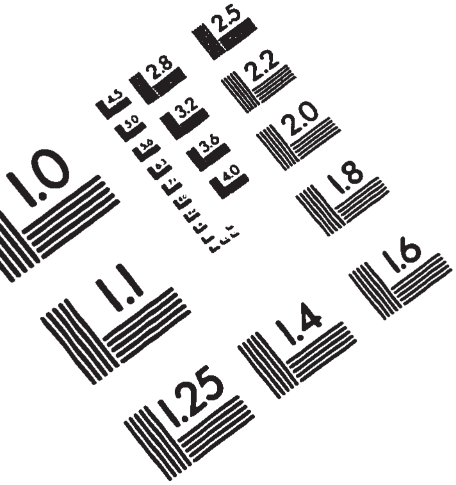
[Wal91] Jean Walrand. *Communication Networks: A First Course.* Asken Associates Inc., 1991.

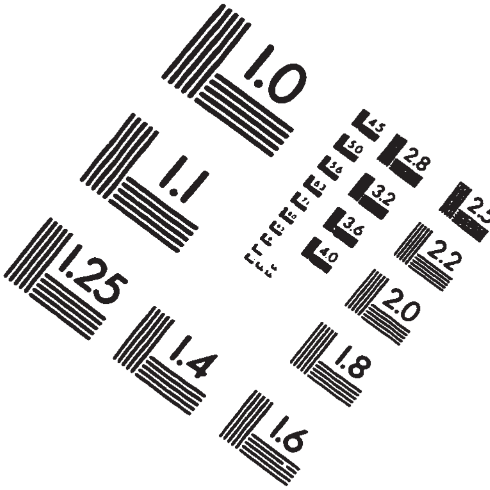
[YIT+94] H. Yasaka, H. Ishii, K. Takahata, K. Oe, Y. Yoshikuni, and H. Tsuchiya. Broad-range wavelength conversion of 10 gbit/s signal using a superstructure grating distributed bragg reflector laser. *Electronics Letters*, 30(2):133–134, January 1994.

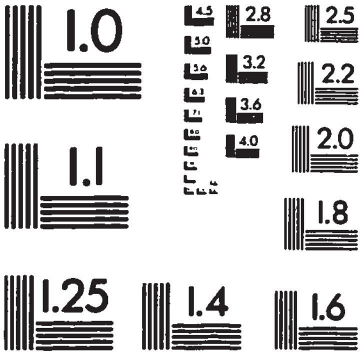
[YIYO95] H. Yasaka, H. Ishii, Y. Yoshikuni, and K. Oe. Repeated wavelength conversion of 10 gbit/s signal using wavelength-tunable semiconductor lasers. *IEEE Photonics Technology Letters*, 7(2):161–163, February 1995.

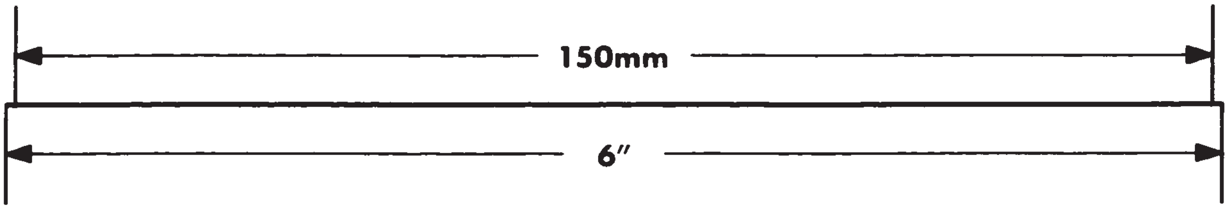
[ZPD+94] J. Zhou, N. Park, J.W. Dawson, K.J. Vahala, M.A. Newkirk, and B. I. Miller. Efficiency of broadband four-wave mixing wavelength conversion using semiconductor traveling-wave amplifiers. *IEEE Photonics Technology Letters*, 6(1):50–52, January 1994.

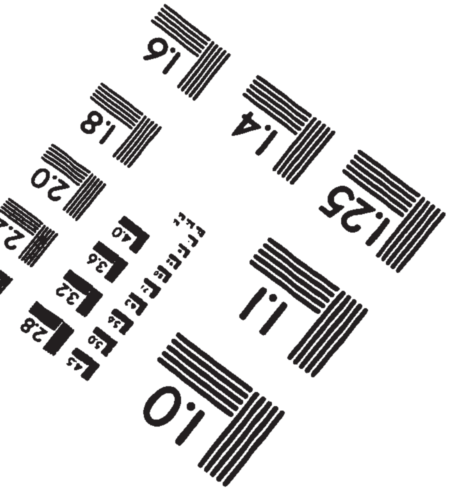
IMAGE EVALUATION TEST TARGET (QA-3)

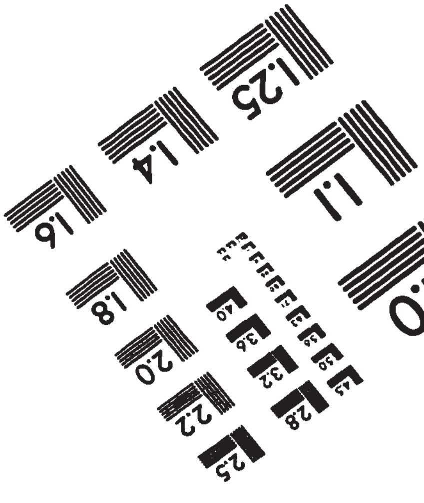












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved