ACTIVE AND PASSIVE COUPLING IN WDM PHOTONIC RING AND BUS NETWORKS

BY
ADRIAN GRAH, B. ENG., M. ENG.
JUNE 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

ACTIVE AND PASSIVE COUPLING IN WDM PHOTONIC RING AND BUS NETWORKS

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

McMaster University
Hamilton, Ontario

TITLE:	Active and Passive Coupling in WDM Photonic Ring and Bus Networks
AUTHOR:	Adrian Grah, B. Eng., M. Eng.(McMaster University)
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:	xvii,190

Abstract

Wavelength division multiplexing technologies are proving to be an effective method for bring high bandwidth to data communication networks. Such networks utilize some form of wavelength agility at stations in order to access a set of wavelength channels. Since each channel is represented by a unique wavelengh, transmissions on one channel are independent from other channels. Unfortunately, this multi-channel nature acts to complicate system operation and often results in increased optical hardware requirements.

Recent advances in photonic amplification have re-motivated the use of ring and bus topologies in single-hop wavelength-division-multiplexed (WDM) networks [Muk92]. Most often stations on a ring or bus WDM network utilize multiple transmitters and receivers which are coupled to the fiber via passive devices. Compared to networks based upon passive optical stars, ring and bus networks offer much simplified station synchronization requirements. However, when minimal station hardware is desired, network operation can be complicated.

In this thesis, four different WDM bus/ring networks based upon passive coupling technology are considered. All four have user station hardware designs with various reductions in the number of tapping points, number of transmitters and receivers and their tunability requirements. With each reduction in hardware, protocol complexities and performance reductions are introduced. In all cases, dynamic packet-switched operation is achieved. The designs thus give an indication of the cost/performance tradeoffs which are possible as the amount of hardware is reduced at the user stations.

Common to each of the designs is the use of a headend controller which attaches to each WDM channel. Media access is achieved through information provided by the

controller. In the more hardware intensive designs, media access is achieved through a mini-slot contention mechanism [GT94]. The remaining designs rely on a hybrid opto-electronic request/allocation protocol motivated by DCCN [JT93]. Analytical models are presented which are validated through discrete-event simulations for each design.

Although this passive tap arrangement is very simple, synchronizing to upstream transmissions is much more difficult compared to existing single-channel busses and rings. This affects the types of protocols which are practical in passive designs. In this thesis, we also investigate protocols for multichannel photonic bus/ring LANs which use active wavelength-selective taps rather than passive ones. An example of such an implementation might use couplers based on acousto-optic devices [Che90]. There are a number of advantages in such a design. Since stations actively tap the wavelength channel that they are accessing, synchronization is identical to existing single-channel rings. This results in much simpler media access protocols than in the passive case. In addition, spatial re-use is possible in the ring designs.

In LANs with active taps, when stations re-tune their couplers at an inappropriate time, transit packets may be destroyed. This action is referred to as a "re-tuning collision". The media access protocols must take special actions in order to prevent this from happening. This places restrictions on the physical parameters over which system performance is efficient. Capacity and delay models are also derived, and comparisons are made with the conventional passive designs developed in this thesis.

Acknowledgments

To Dr. Terry Todd, Mark Janoska, Nima Ahmadvand, Hugh Pasika and my parents, Rose and Ivan Grah: So long, and thanks for all the vacations!

Lastly, I owe many thanks to Tanya DiTommaso. *Me big, you small.*

Contents

List of Tables

List of Figures

List of Acronyms

ACTA	Adaptive-Cycle Tunable-Access
ASIC	Application Specific Integrated Circuit
BTMSR	Blind Transmitter Mini-Slot Reservation
BTMSC	Blind Transmitter Mini-Slot Contention
BWB	BandWidth Balancing
CC	Channel Controller
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DCCN	Distributed Channel Controller Network
DNA	Digital Network Architecture
DoD	Department of Defence
DQDB	Distributed Queue Dual Bus
EOS	End Of SIG
EQEB	Estimated Queue Expanded Bus
FDDI	Fiber Distributed Data Interface
FBT	Fused Biconical Taper
Gbps	Giga-bits per second
IEEE	Institute of Electrical and Electronic Engineers
IMP	Interface Message Processors
IP	Internet Protocol
ISO	International Organization for Standardization

LAN	Local Area Network
LLC	Logical Link Control
MAC	Media Access Control
MAN	Metropolitan Area Network
MAC	Media Access Control
OSI	Open Systems Interconnection
PBS	Polarization Beam-Splitter
SIG	Station Insertion Gap
SNA	System Network Architecture
SONET	Synchronous Optical NETwork
SOS	Start Of SIG
TCP	Transaction Control Protocol
TDM	Time Division Multiplexing
TDMA	Time Division Multple Access
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing

Chapter 1

Introduction

Future Local Area Networks (LANs) employing optical Wavelength Division Multiplexing (WDM) techniques may prove to be the answer to exploiting the enormous bandwidth of optical fiber [Bra90]. This chapter serves as an introduction to this thesis which proposes several WDM network designs based on ring and bus topologies. The introduction begins with a short overview of the material and states the motivations for the conducted research. The scope of the thesis is then outlined and the major contributions of the research are stated. The chapter concludes with an outline of the remaining chapters in the thesis.

1.1 Overview

Designs employing WDM divide the available bandwidth on optical fiber into a set of parallel channels, each running at electronic data rates. Currently, channels running in the low Gbps are practical. In addition to making fiber's bandwidth more accessible to the electronic domain, WDM introduces concurrency into the system. This increases the total network capacity compared to conventional single-channel local and metropolitan area networks (LANs and MANs).

In the past, the study of multichannel photonic LANs and MANs has been based on the use of passive optical star couplers. In these designs, each station is given a separate inbound and outbound fiber to the star. The star distributes optical energy

impinging on the inputs across all outputs (subject to various losses), thus creating a broadcast-and-select network. This design is much better than many others in terms of optical link budget. However, the user stations synchronization requirements and media access protocols are often very complex and stringent when minimal station hardware is desired. For a survey of these networks the reader is referred to [Muk92].

In [Muk92] a number of issues specific to single-hop WDM networks are addressed. A primary concern raised is that of hardware and media access protocol complexity at the user stations. For a cost effective design, it is essential to reduce the hardware required at the user station interface. A single transmitter and receiver per station, with one of the devices being tunable, is the absolute minimum transmission hardware. Current proposals often require multiple tunable transmitters and receivers [Che93, CM92, CSR90, SD93].

Although multi-channel ring and bus networks have many advantages from a protocol standpoint, they have been impractical in the past because of their poor optical power budgets [Gre93]. Recently, more attention has been focussed on such designs [SD93, IK92, BM93]. Advances in both broadband traveling-wave and doped-fiber amplifiers [Sen92, DSP87] as well as wavelength selective couplers [Che90] have motivated this current interest. It is clear that these devices will permit practical photonic bus and ring LANs and MANs in the future.

1.2 Motivation

In this thesis, several multi-channel network architectures based upon either passive or active station coupling are proposed. The motivation behind the project is to examine the hardware requirements of the user stations and study what constraints are intrinsic to such designs. Each design requires the aid of an intelligent headend which attaches to each of the WDM channels. However, the amount of processing required at the headend varies in inverse proportion to the hardware complexity of the user stations.

For networks employing a passive coupled station design, two protocols are introduced which may be applied to both folded bus and ring topologies. The first

relies on the headend to provide bandwidth partitioning in a way such that stations may compete for slot access. The second protocol has the headend provide much more functionality in that it allocates specific slots to stations on a per request basis. In each case, the goal of the design is to reduce the user station hardware requirements while studying the impact such reductions have on media access protocols and performance. Capacity and delay results are presented and compared for the two systems.

A ring or bus network can also be designed using stations which are coupled to the network using active wavelength-selective couplers. Such designs are able to regenerate optical signals at various points in the network. Additionally, an attractive feature of these devices is that they are "fail-safe". That is, in the event of power loss to the device, optical energy will continue to flow through the device subject only to the insertion losses of the device. Lastly, since a specific wavelength can be fully and directly accessed by a station, many protocol options are available that are not in passive designs. It is these device characteristics which make the investigation of their use in WDM networks interesting. Through the study of such systems this thesis will demonstrate the constraints of active couplers in station design and compare this to designs which use passive couplers.

1.3 Scope Of Thesis

According to the network classifications in the OSI model, this thesis describes and studies new Media Access Control (MAC) level protocols as applied to ring and bus topologies using passive and active wavelength-selective couplers. As a result, issues pertaining to network design using the optical fiber media and optical coupling devices of varying capabilities are addressed. Of particular interest are the ramifications for the use of active couplers versus passive couplers on media access.

The presented protocols are based on a broadcast and select topology, in which a large number of transmitters and receivers are assigned to both transmit and receive from a single channel. The protocols' performance objectives are to achieve high

capacity and low queueing delays even though hardware requirements have been reduced with respect to traditional optical WDM networks. Also, problems particular to the station coupling technology employed must be handled in appropriate and efficient manners. Performance will be demonstrated through analytical and simulation models. Although many aspects in the operation of these networks are necessarily explained, they are secondary to the main purpose. For example, even though the physical layer of the networks will be explained, it is done so only in sufficient detail for a general understanding of the network's operation. Clearly many interesting physical issues will need to be omitted for the tractability of the problem at hand. Similarly, the network layers above the MAC sub-layer will receive only passing and necessary attention. Ample references are given to relevant, but secondary material in order to avoid lengthy expositions. Thus it is the thrust of this thesis to provide a study of media access control design and performance as dictated by a particular station design and network topology.

1.4 Contributions

For the purposes of this thesis a thorough development of MAC protocols for optical WDM ring and bus networks using current research and commercial grade devices is conducted. The work is an exploration of the limits that can be achieved in terms of protocol and station hardware simplicity. For the first time, this thesis introduces the use of active wavelength-selective couplers into the design of network stations as a means of providing access to optical signals in an optical packet switched network. Through techiniques developed here, it is shown that many traditional single channel protocols can be exported to the WDM domain. Also, it is found that networks based upon passive coupling still have many advantages over the active coupled networks and novel media access protocols are developed to support a reduced amount of station hardware. Other important contributions to the field of WDM photonic networks are:

- Reduced hardware requirements of traditional passively coupled designs through the use of novel media access methods. A minimum number of optical components used in a single user station would be one transmitter and receiver, with

one of those devices being tunable. Designs which approach this logical minimum are proposed and studied.

- Introduced several protocols suited to high speed WDM networks with reduced station hardware. That is, the proposed protocols are scalable to high bit rates and increased system diameter with minimal effect on performance. Often, as with the case of earlier generation protocols such as the token ring, scaling the network in speed or size results in an increased amount of bandwidth been utilized by system overhead. The proposed networks scale well into the metropolitan area at bit rates into the gigabits per second range.

- Enabled the possibility of wavelength channel reuse into many designs. This can effectively double a systems capacity under uniform loading conditions.

- Simplified the station hardware design and associated protocols to remove potential bottlenecks in the system operation. Since complex algorithms need not be performed at each station, the effect of the electonic processing bottleneck is minimized. Furthermore, since each channel operates at only a fraction of the total system bandwidth, the electronic interfaces can be much more easily implemented.

- Simplified synchronization requirements of optical WDM networks when considering an active station interface. This is a result of selected wavelength signals being regenerated at each station.

- Extended the concept of a channel controller architecture to the headend of a Ring and Bus.

1.5 Organization Of Thesis

This thesis attempts to provide a narrative and summary for the results generated over several years of research in the area of optical WDM ring and bus networks. The work is arranged so that a casual reader will get an overview of the proposed

networks. The interested reader will obtain more in-depth coverage of the simulation and analytical models by detouring into the appendices. The chapters cover the following topics:

- Chapter 2 gives a general introduction to networking and the related photonic componentry which provide the prime motivation for this research. This chapter simply highlights key concepts, although exhaustive references are provided for the interested reader.

- Chapter 3 provides a literature study of current related material. Ring and bus networks of the first, second and third generations of networks are summarized with specific focus on design, media access and performance.

- Chapter 4 begins the original content of this thesis and covers the physical aspects of passively coupled network architectures. This includes station and headend designs and their requirements. Then described are the procedures which stations must follow to access bandwidth on the network. Also, the operation of the headend is detailed including some simple scheduling algorithms. Lastly, performance indices such as capacity and mean delay for the proposed designs and protocols are examined. This is done through analytical and simulation approaches. This is followed by an analysis of the optical power budgets. Results are compared between the architectures as appropriate. This chapter should provide enough information to recreate a simulation model of any of the protocols discussed.

- Chapter 5 covers the physical aspects of the active wavelength-selective coupled network architectures. This includes station and headend designs and their requirements. Options for folded bus, dual bus, and ring topologies are included. Following this, the chapter deals with the related active design protocols. Described are the procedures which stations must follow to access bandwidth on the network. Also, the issue of how to most efficiently provide and calculate the required station insertion gap is addressed. Several options are examined and performance indices are studied. This is followed by an analysis of the optical

power budgets. Results are compared between the architectures as appropriate. This chapter should also provide enough information to recreate a simulation model of any of the protocols discussed.

- Chapter 6 provides conclusions and suggests future work related to optical ring and bus topologies.

Chapter 2

Background

In this chapter a number of topics fundamental to communications network and media access protocol design are introduced. Of particular interest are those issues which are related to the optical device technology upon which this thesis rests. Where appropriate, only brief summaries of material are provided in order to avoid lengthy expositions of previously published material. Ample references are provided to the text books, research journals, and conference proceedings which comprise the basis for this chapter.

2.1 Communication Networks

With the recent explosion of the Internet, the concept of a communication network has become ubiquitous. Certainly one would be hard pressed to find a person in western society who does not have some interaction with a network of any sort in their daily lives. Formally, a communication network can be considered as a system that is used to connect a set of stations or nodes together in order that they may exchange data. Many excellent introductory books are available on this subject and readers interested in a more thorough treatment of this material are directed toward the references [SHP91, Sta87, Tan89]

Historically, communication networks originated in the 1950s as a means for the military to interconnect sensors, weapon systems and command-and-control centers

[SHP91]. Networks have now evolved to find applications in a huge range of areas. Banking and finance, reservation systems, grocery and retail store checkout systems, offices, and factories are samples of the areas into which communication network technology has invaded.

Those who look to install a network of any sort, do so with varying motivations. Some installations are geared to share information like a large database with several or even several hundreds of users. Other networks are installed to provide e-mail services and possibly a "paper-less office" environment. Still other networks are used to share resources like large disks, printers and modems.

Whatever the original motives have been, the effect of networking has been to create a distributed working and living environment. Offices and branch offices can now be in full contact with one another even though they span cities, provinces or countries. Banking can be done at just about any of the "banking machines" that can be found world wide. We can mail order for goods using a credit card number which is again processed by another large network. It is also now conceivable to work at home and be connected to resources at work via a network. The proliferation of the Internet will only re-enforce our dependence on the computer network.

2.2 Basic Network Principles

A goal in the development of networks is to share a common communication facility among multiple users. Each user will be referred to as a *user station* or just *station*, although the literature also refers to these as *hosts*, *end systems*, or *nodes*. User stations are interconnected by a communication subnet whose responsibility it is to transport messages from station to station. A subnet simply refers to a particular portion of a larger network that connects a set of interested user stations.

Typically, networks are categorized by their geographical coverage. Networks spanning distances of not more than a few kilometers are classed as Local Area Networks (LANs). Metropolitan Area Networks (MANs) deal with distances of up to a few dozen kilometers. The last category is the Wide Area Network (WAN) which can extend to world-wide coverage [Tan89].

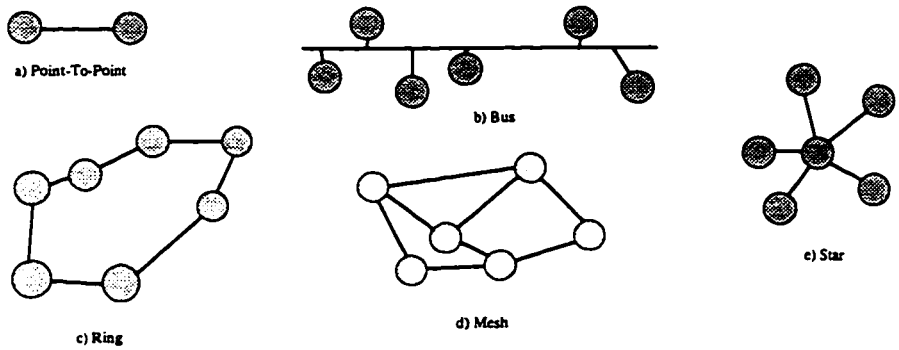


Figure 2.1: Common Network Link Topologies. a) Point-to-Point. b) Bus. c) Ring. d) General Mesh. e) Star.

A further criterion for classifying networks is the network's topology or geometry in reference to the particular pattern used in connecting stations to one another. The most popular network topologies are point-to-point, bus, ring, mesh, and star which are illustrated in Figure 2.1.

Point-to-point topologies are the most simple, connecting a pair of stations to one another as illustrated in Figure 2.1a. Such designs require that a single transmitter and receiver be connected to the link at both stations. Each transmission from source to destination passes over this dedicated link. The bus topology, Figure 2.1b, places stations along a common bus via taps where data flow may be bidirectional or unidirectional depending on the medium and technology used. Each station requires at least a transmitter/receiver pair which is connected to the shared medium in some way. The result is that, in total, multiple transmitters and receivers will be connected to the same medium. Data travels from a source station to its destination in a single hop without buffering or routing by intermediate stations. The ring topology as shown in Figure 2.1c, connects stations in series along a path which loops back onto itself. The ring topology typically employs unidirectional media such that data flows in one direction around the ring passing through any intermediate stations. Each communication link has only a single transmitter and a single receiver connected to it. Figure 2.1d shows a mesh topology which employs multiple links to create a web like network between stations. Each link may be either bidirectional or unidirectional. A special case of the mesh is one in which each station is "fully connected" to all

other stations. Each link will have a single transmitter and receiver connected, thus each station may be required to maintain multiple receivers and transmitters. Data flowing from source to destination may need to pass through intermediate stations where routing may occur. Routing simply refers to the selection of a particular link such that data will eventually reach its destination. Mesh topologies offer good redundancy and fault tolerance. However, for large networks, fully connected systems are impractical as they require $\frac{N(N-1)}{2}$ bidirectional links, where N is the number of attached stations. Lastly star topologies as that shown in Figure 2.1e, use some central switching point to which all stations are connected via point-to-point links. This switching point may be either an active or passive device. Each transmission must pass through this switching point in order to reach its destination.

The above described network topologies may be further categorized by their use of either switched point-to-point channels or multiaccess broadcast channels. Switching modes of operation compose the historical form of computer communication subnets and are often based upon point-to-point, star, ring, or mesh topologies. Two main techniques of switching have been proposed: circuit switching, and packet switching [Sta87]. The multiaccess broadcast approach encompasses the use of broadcast channels which topologically may be implemented as a bus, ring, or broadcast star.

2.2.1 Circuit Switching

A simple illustration of the circuit switching technique is a telephone call which provides a connection-oriented service. This approach creates a single fixed end-to-end path between the source and destination stations through possibly many switching elements which comprise the subnet. Each of these switching elements creates a fixed data path such that the line remains intact and is used exclusively by the involved stations until the connection is terminated. To use such a network requires a set up phase in which the circuit switches are set between the source and destination, after which communication may occur at the line's full bandwidth. This technique is useful for many types of communication such as voice, but for computer data traffic which tends to be bursty in nature, circuit switching is often inefficient since a connection

may remain idle for large amounts of time. To increase network utilization and reduce costs, a method whereby many users can share network lines on demand, such as with packet switching, was developed.

2.2.2 Packet Switching

In traditional packet switching a subnet is composed of a number of point-to-point links connecting pairs of *packet switching nodes* or *IMPs* (Interface Message Processors). Each IMP is a specialized switching element which forwards data arriving on an incoming line to an appropriate outgoing line. Every message generated by a station is divided into smaller fixed length units called packets, containing data and some amount of overhead information. Packet overhead is composed of the information required to allow the network IMPs to determine the packet's destination, provide error control, and facilitate the receiving station to reconstruct the message generated by the source station. This "connectionless" service can be thought to mimic the postal system where every packet (letter) carries a full destination address. Each packet is then multiplexed onto the network thus sharing the medium with many other stations. If a packet is destined between two IMPs that are not directly connected, it must be routed through other IMPs. Packets arriving at intermediate IMPs must be buffered until the required output line is free, at which time it may be forwarded. Each time a packet is stored or retransmitted at an IMP on its way to its destination is referred to as a hop. Since links are now shared among many competing users, these networks which are called store-and-forward, packet switched, or multihop, tend to be much more fully utilized than the circuit switched counterpart [Sta87]. A problem with this approach is that delays associated with queueing at IMPs can seriously increase the delay a packet experiences on its way to the destination during times of network congestion.

2.2.3 Multiaccess Broadcasting

Multiaccess broadcast networks share a single communication channel among multiple user stations. These networks are usually single hop and as such do not require

intermediate buffering or equipment for routing packets to any given destination. Once a packet is placed on the network, it finds its way to the destination without the aid of intermediate stations. This eliminates many problems associated with multi-hop networks since the queuing delays introduced at intermediate stations are removed.

This is accomplished by means of an address field contained within each packet which uniquely specifies the destination station. Also, since packets are received by all stations there exists the possibility of packet broadcasting and multicasting in which packets are addressed to a subset of destination stations. When a packet is received by a station, the address field is examined, and if the packet was not intended for that user station, it is simply ignored. In order to resolve transmission conflicts that can occur when two or more user stations wish to broadcast simultaneously, some arbitration scheme is required such that the network will have only one master at a time. This strategy is typically implemented through a distributed algorithm which coordinates and arbitrates stations' rights to transmit a packet on the media. Such media access protocols can typically be categorized as being contention based or conflict-free.

For a contention based protocol, all user stations are permitted to transmit without regard to there being a packet already on the medium. For a successful transmission to occur, a packet must avoid "collisions" with packets transmitted by other user stations. This requires first that the transmitting station has not started its transmission while another transmission is already in progress or about to start. Secondly, successful transmission requires that no other stations begin transmitting while the user station is in the process of transmission. If either of these conditions are not met it must be assumed that none of the packets will be received correctly and the transmission must be attempted again. Popular examples of contention based protocols are known as ALOHA, Slotted ALOHA, CSMA and CSMA/CD [SHP91] which will be discussed in some detail in the next chapter.

In a network using a conflict-free protocol, transmission is regulated by the algorithms so that collisions will not occur. This may consist of the protocol allocating

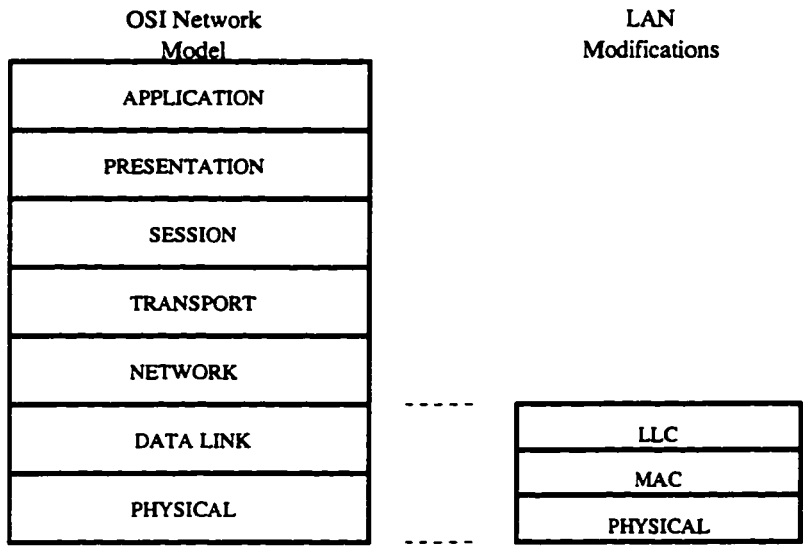


Figure 2.2: Network reference models: OSI and LAN modification.

a specific time for transmission to a user station in either a static or dynamic manner. Time Division Multiplexing (TDMA) [Tan89] is a static form of a conflict-free protocol whereas the IEEE 802.5 Token Ring [Tan89] would be a dynamic version.

2.3 OSI Model

In order to provide a common framework for the development of network protocols, the Open Systems Interconnection (OSI) model was developed [Tan89]. First proposed in 1978 and finally adopted by the International Organization for Standardization (ISO) in 1983, the model was intended to be applied to Metropolitan Area Networks (MANs). Each layer, built upon its predecessor, offers certain services to higher layers while shielding those layers from the implementation details of the services.

Table 2.1 describes some of the principle functions of the OSI reference model. The ISO extended the OSI model to include Logical Link Control (LLC) and Media Access Control (MAC) levels to better support LAN and WAN networks as proposed by the Institute of Electrical and Electronics Engineers (IEEE) in a standard called

IEEE 802. The MAC layer organizes communications over the link as discussed in Section 2.2.3. It encompasses a set of logical rules which allow stations to access a common communication channel. The LLC layer assembles and disassembles data frames or packets and provides address and error checking information [Tan89].

The OSI model has been widely accepted as the definitive model for protocol development. One reason it has become a standard is that it encapsulates network functions within a confined area and allows for diversity in the overall network. That is, the model's levels are independent of one another, and if the interface between levels is consistent, it is possible to drop new versions of any layer into the network without upsetting the other layers. For example, it is possible to run TCP/IP [Tan89] as the network/transport layers on a physical layer such as coaxial, as is often done with Ethernet. It is equally valid to run NOVELL/SPX [Tan89] as the network/transport layers over the same physical network.

2.4 Transmission Media

So far our discussion of networks has not considered the medium on and through which the communications take place. Indeed for communication to occur, data must be moved from one point to another. The purpose of the physical layer is to provide the transportation for a raw bit stream from one station to another. Commonly used media include magnetic media, twisted pair copper wiring, baseband and broadband coaxial cable, and free space radio frequency signals. However, one medium is emerging as superior: optical fiber.

2.4.1 Fiber Optics

Light has been used as a medium for communication for a long time now. Signal fires and torches were early attempts at exploiting light for communication. In 1870, John Tyndall was the first to demonstrate the principles of guided light transmission [Tyn78]. Soon after, Alexander Graham Bell was the first to use light as a communication medium in any modern sense [Bel80]. However, Bell's invention was limited

Level	Principle Functions
1. PHYSICAL	For transmitting raw, unstructured data over a physical medium. Main design issues deal with mechanical, electrical, optical and procedural interfaces.
2. DATA LINK	Transforms the raw transmission into what appears to be an error free data block; controls the transmission and reception of blocks with necessary synchronization, error and flow control.
3. NETWORK	Controls the operation of the subnet. Issues include how packets are routed, congestion control, addressing and packetizing. Provides upper layers layers with independence from data transmission technologies.
4. TRANSPORT	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.
5. SESSION	For tying terminals to applications during a session; establishes, manages and terminates connections (sessions).
6. PRESENTATION	Concerned with the syntax and semantics of the information transmitted. May provide encryption, compression and/or reformatting.
7. APPLICATION	Contains a variety of protocols which are commonly used such a terminal translations and file naming conventions. Electronic mail, remote job entry, directory lookup, and various other general-purpose facilities are also implemented here.

Table 2.1: Principal functions of the OSI reference model.

in many ways. It could only transmit up to 200m and was limited to line of sight. Optical communication required a more refined technology.

The invention of the laser in 1960 sparked new interest in the realm of optical communication [Mai60] and provided the beginning of the technologies that were to follow. The laser offered a coherent and powerful light source that could easily be modulated at high frequencies.

Bell had used a diaphragm to modulate sunlight which allowed speech to be transmitted up to 200 m. Lasers far exceeded Bell's limitations and made free space optical transmission a viable possibility for short haul communication links. However, free space optical systems still had to contend with its inherent difficulties to become a universally usable technology.

It is the effects of atmosphere that limit the distance free space optical links may span. Phenomena like snow, rain, fog, dust, and atmospheric turbulence, all severely inhibit the performance of free space optical systems. Furthermore, free space optics is restricted to line of sight.

Attempts were made to make optical communication more feasible and less susceptible to atmospheric degradation. Proposals to use a dielectric waveguide or optical fiber drawn from glass were made in 1966 [KH66, Wer66].

The ten years following the initial proposals saw optical attenuation in fiber drop from over 1000 dB km^{-1} to under 5 dB km^{-1}. Fiber approached and bettered the performance of coaxial cable which shows a 5 to 10 dB km^{-1} loss. Furthermore, new splicing techniques for optical fiber were perfected allowing for more complicated optical systems to be constructed than ever before. Fiber optics began to make a large impact on the way communication technologies were approached.

To accommodate the newly emerging fiber technology, much research was aimed at ways of improving the other components that would comprise an optical network. Semiconductor optical sources and detectors that were compatible with optical fiber began to emerge. Initially, these devices had life expectancy of only a few hours. However, by 1977 lifetime had been extended to 7000 hours through the use of gallium arsenide alloys [GPPB77].

Today, lifetime has been extended to a mean of 25 years and beyond for InGaAsP

devices [Mil88]. Further, direct modulation at Gbps rates is also attainable. It is not surprising that these advances in device technology and low fiber losses of about 0.2 dB km^{-1} have made fiber optical communication the fastest growing area in communication ever. Every day hundreds of kilometers of new fiber are being installed in the U.S. alone [Kra91].

Many other advantages to the use of fiber optics exist. Among those advantages are the economics of low optical losses relative to copper wire. The implication being that repeater spacing can be greatly increased over copper, thus reducing the number of repeaters used on a line. Optical signals are also immune to electric and magnetic interference, making a photonic network particularly interesting in environments of extreme magnetic and electric activity such as a factory floor. Conversely, the optical signal in a fiber emits no radiation to the environment which could prove usefull in the proximity of sensitive equipment. When faults in optical fiber do occur, it is possible to locate the faults using devices called reflectometers. Finally, optical fiber is small and lightweight, making it appropriate in environments where weight and size constraints are at a premium such as on board a satellite.

2.4.2 Wavelength Division Multiplexing

Optical fiber offers a tremendous amount of bandwidth, on the order of 25 THz for a single strand, in the low loss, low dispersion wavelength window of 1.55μm [Gre93]. However, due to the large difference in electronic and optical speeds, and immature optical device technology, current optical systems are limited to use only a fraction of fiber's potential bandwidth.

In order to avoid what has been called the electro-optical bottleneck, designers have resorted to wavelength division multiplexing (WDM) techniques. This consists of splitting the available bandwidth of fiber over several lower bit rate wavelength channels. Each of these channels is independent and runs at available device and electronic speeds. Conceptually it is like taking a single fiber and splitting it up into several smaller fibers with less bandwidth as shown in Figure 2.3. Currently, the state-of-the-art is considered to be Giga-bit per second (Gbps) channels.

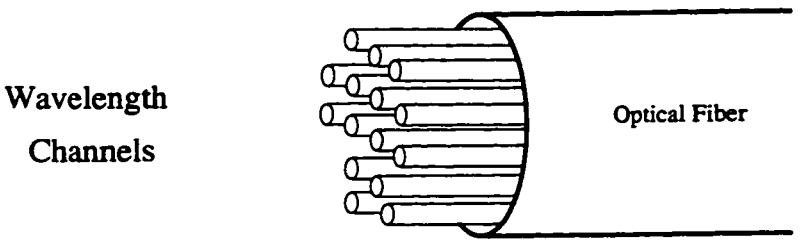


Figure 2.3: Wavelength Division Multiplexing Concept

In addition to making fiber's bandwidth more accessible to the electronic domain, WDM also introduces concurrency into the system. This means independent and simultaneous transmissions can occur on each channel. A contraindication of WDM is the protocol problems it introduces.

2.5 Optical Device Technology

It will now prove beneficial to provide a discussion of the various devices which form the basis of a photonic network. The brief discussion will highlight the characteristics and limitations of the optical building blocks which are commonly used. Some of the devices in the ensuing discussion are broadband devices, yet they still prove useful in a WDM environment. Other devices have been specifically engineered to exploit various physics to generate useful WDM functionality. Many text books covering basic optics to advanced device concepts are available such as [Sen92, Gre93, Gow93, Jon88] for the interested reader. The discussion begins with the most basic piece of a photonic network: the optical fiber.

2.5.1 Fiber

As mentioned earlier in Section 2.4.1 the use of fiber optics holds many advantages over its alternatives. To understand the operation of optical fiber requires more space than can be justified for this simple summary. However, much can be learned by studying fiber with a ray trace approach [Pal92]. Viewed this way, optical fiber operates on the principle of total internal reflection which occurs due to the bending

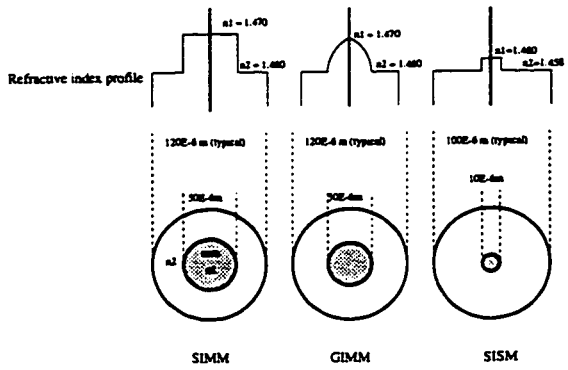


Figure 2.4: Basic Optical Fiber Types

of light at the boundary of two materials with different refractive indices. When a ray of light approaches such a boundary at an angle greater than or equal to an angle called the critical angle, all of the incident light will be reflected at the boundary. This critical angle is a function of the ratio of the two refractive indices at the boundary. The fiber cable itself comes in three basic configurations: Step-Index Multi-Mode (SIMM), Graded-Index Multi-Mode (GIMM), and Step-Index Single-Mode (SISM) [ST91] as illustrated in Figure 2.4. The simplest configuration is the SIMM fiber. In such a configuration, the fiber consists of two distinct refractive indices, the core and the cladding. The refractive indices are shown in Figure 2.4 by the refractive index profile. For GIMM fiber, notice that the refractive index profile increases gradually in the core in contrast to the abrupt increase of the SIMM fiber. The single mode fiber is very similar to the SIMM fiber except that it uses a much smaller core in order to limit the number of modes that may be developed in the fiber.

Simply put, a mode in a fiber can be thought of as a group of light rays bouncing through the waveguide at a given angle of incidence and reflection [Gow93]. Because light will add constructively or destructively, it turns out that only distinct angles of incidence and reflection will be supported in the fiber, while all others will be suppressed. These angles or modes are a function of the fiber's dimensions, the ratio between core and cladding refractive indicies, and the wavelength of light used. Typically, there are hundreds of modes in a multi-mode fiber. Figure 2.5 illustrates a number of modes as developed in SIMM and GIMM fiber. Of course, in single mode

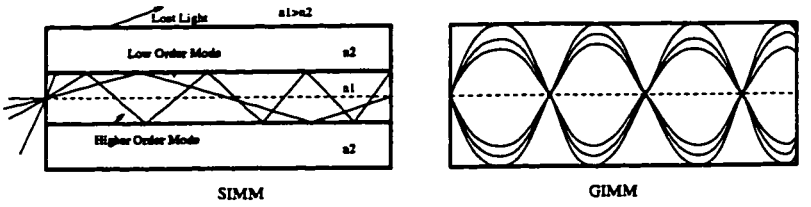


Figure 2.5: Optical Modes developed in Fiber

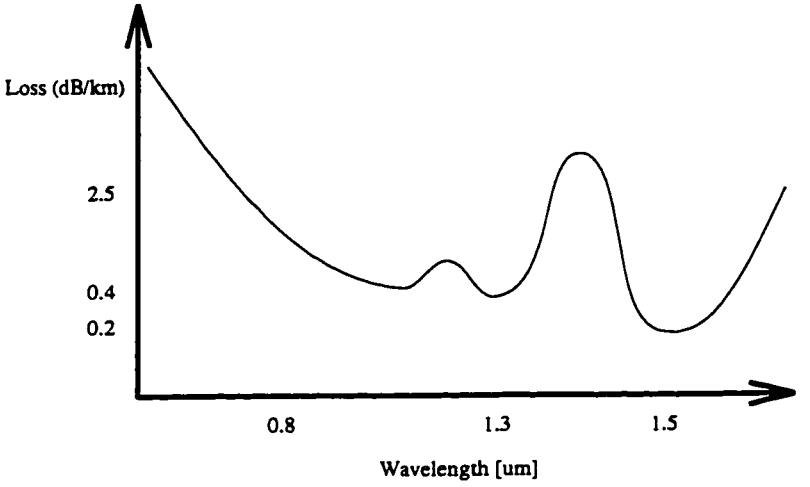


Figure 2.6: Optical Losses in Fiber as a function of Wavelength

fiber, only one such mode will exist.

These modes are of significance as the various modes will traverse the fiber at different speeds resulting in an effect called modal dispersion [Pal92]. This modal dispersion is a limiting factor in the distance that any signal of a given bit rate may be transmitted over the fiber. In order to avoid this degradation, many communication links rely on single mode fiber in which this effect is absent.

A further consideration in the use of optical fiber is the dependence of signal attenuation on the wavelength of light propagating the fiber. Figure 2.6 illustrates the two main low loss windows available in optical fiber. The first window occurs at $1.3\mu m$ and the second at $1.5\mu m$ [Sen92]. Historically, optical communications was initiated by the development of sources at the $0.8\mu m$ wavelength. Advances in source and receiver technology, however, have enabled the use of the longer wavelength, low loss regions of fiber. Furthermore, refinements in fiber construction techniques were

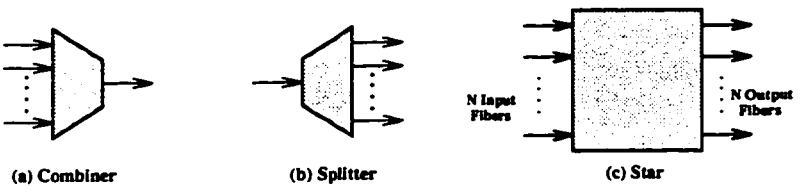


Figure 2.7: Passive Couplers and Splitters

able to generate core diameters small enough to support single mode transmission at these longer wavelengths.

Normally, a common point to point link would require that a single transmitter be directly coupled to the fiber at one end and a single receiver be directly coupled at the other end. Although this coupling is very simple, it limits the number of stations that may access a particular fiber.

2.5.2 Passive Couplers

Apart from the direct coupling used in a point to point link, passive couplers and splitters are probably the most commonly used devices in connecting a user station onto a network. In the literature, the most popular devices are combiners, splitters, and star couplers [Gre93]. These are illustrated by Figure 2.7. A N port splitter is a device that equally divides optical energy from a single input fiber to N output fibers. A N port combiner couples the energy from the N inputs onto a single output fiber. The star coupler is a $N \times N$ device. Ideally, optical energy impinging on each input fiber is equally distributed to each of the output fibers. Of course, as with each of these devices, some of the optical energy is wasted due to the insertion loss of the device [Gre93]. In all cases, the coupling and splitting are independent of wavelength and thus are regarded as wideband devices.

A device of particular interest is the Fused Biconical Taper (FBT) coupler [Sen92]. In its simplest form, two fibers are twisted together and melted under tension over some distance Z. This creates the four port device shown in Figure 2.8 which is also a wideband device as its operation is independent of wavelength. Energy impinging on input port 1 or 2 is divided between output ports 3 and 4. The coupling ratio gives the percentage division of optical power between the output ports. Define the

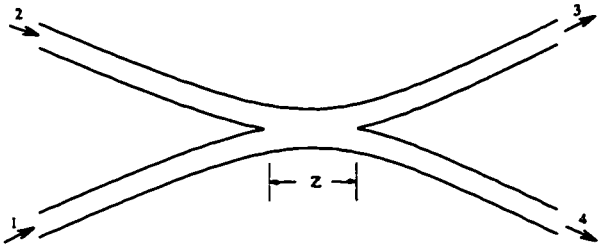


Figure 2.8: Structure of a Fused Biconical Taper (FBT) Coupler

optical power at a port x to be P_x, then the coupling or split ratio can be written as,

$$\begin{aligned} \text{Coupling ratio} &= \left(\frac{P_3}{(P_3+P_4)}\right) \times 100\%, \\ &= \left(1 - \frac{P_4}{(P_3+P_4)}\right) \times 100\%. \end{aligned} \tag{2.1}$$

It is possible to manipulate the coupling ratio of the device by varying the distance Z over which the fuse occurs. Based on this device, the $N \times N$ star coupler can be constructed by cascading multiple FBT couplers. Again, in a real device, insertion and excess losses should be considered [Gre93].

2.5.3 Tunable Sources

The choice of any optical source depends on a number of factors. When considering a WDM system, however, the choices become much more limited. Of the three main types of optical sources available, which are "continuous spectra" sources such as incandescent lamps, monochromatic incoherent sources such as light emitting diodes, and monochromatic coherent sources such as available from laser, only the laser sources can provide the required power, spectral line widths, and modulation bandwidth required for modern WDM systems [Sen92].

Viewed as a "black box", the tunable laser source is a device which is capable of providing a modulated signal on one of a set of available wavelengths. Typically, only one wavelength can be excited at any given time, while all other wavelengths remain dormant. Figure 2.9 illustrates a tunable laser block. Information is coded

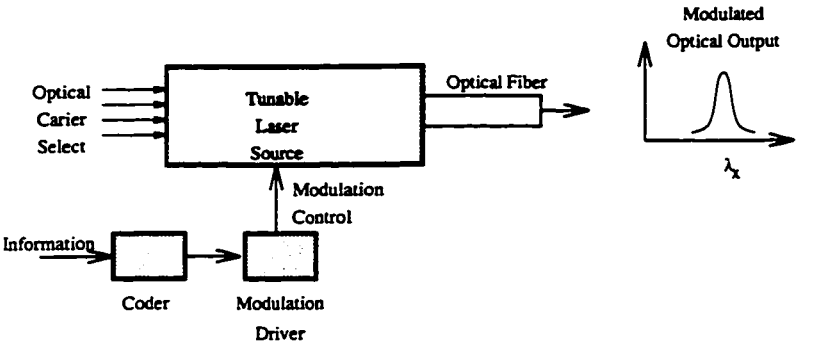


Figure 2.9: Stylized Tunable Laser Source

and passed to a modulation driver which feeds the modulation control of the laser which in normally achieved by varying the laser's power supply in an appropriate manner [Gow93]. The optical wavelength that the laser operates at is manipulated via some electronic control. Factors of primary concern are the number of resolvable channels and the wavelength tuning speed of the device.

A tunable laser can be constructed by a number of technologies. One method is to integrate a number of fixed tuned laser upon a common substrate and couple the outputs of all the lasers to a common waveguide [Gre93]. Since only a single laser will be active at a time the device will appear from an external point of view as being tunable. This type of tunable laser is known as a laser array and can be tuned at very high speeds. There are, however, efficiency concerns since possibly a large number of laser sources may need to be coupled together onto a single waveguide.

A second approach to the generation of a tunable source is called spectral slicing. In this method, a single laser with a broad spectral line width followed by a tunable filter is used. The filter allows only a subset of the laser spectrum to pass through, thus creating the desired narrow spectral line width and affording the required tunability. Tuning speed in this method is limited by the tuning speed of the optical filter. Since, however, only a portion of the laser's energy is allowed to propagate through the filter unimpeded, the available power output is severely restricted.

Finally continuously tunable lasers based upon thermal, mechanical, acousto-optic, or electro-optic effects are possible [Sen92]. The use of a particular technology will ultimately depend on how it addresses the issues of tunability range and speed.

Technology	Tuning Speed	Tuning Range
Thermal	ms	1-2nm
Electro-optical	ns	10nm
Mechanical	ms	100nm
Acousto-optical	μs	100nm

Table 2.2: Tunable Laser Range and Speed Comparison

Table 2.2 compares the minimum tuning times and the maximum tuning range for various promising technologies. For networks that are to provide packet switched modes of operation, tuning speed becomes of paramount importance. Speeds better than the microsecond and nanosecond ranges are desirable. It is the advent of these fast tunable transmitters along with the tunable receiver that has been a driving force behind the design of third generation optical networks as discussed in the next chapter.

2.5.4 Tunable Receivers

Just as with tunable sources, a number of approaches towards tunable receivers have been examined in the literature. An ideal "black box" representation of a tunable receiver is given in Figure 2.10. Such a device would receive at its input a number of wavelengths, each containing coded information. The device is capable of selectively discerning a particular wavelength out of the entire set of wavelengths at the input, and generate an electronic signal representative of the modulated wavelength in question. Similar to the tunable source, the selection of a particular wavelength would be controlled be some electronic means. Parameters which will ultimately distinguish one technology from another are the number of resolvable wavelengths and the tuning time required for the device to switch between wavelengths.

Tunable receivers are often built by placing a tunable filter element prior to a wideband receiver element [Sen92]. Such an optical filter is illustrated in Figure 2.11 where a number of wavelengths are available at the input of the filter, but only a single selected wavelength is passed to the output. The selected output is thus used to drive the receiver. Many techniques for providing the required tunability exist. Among

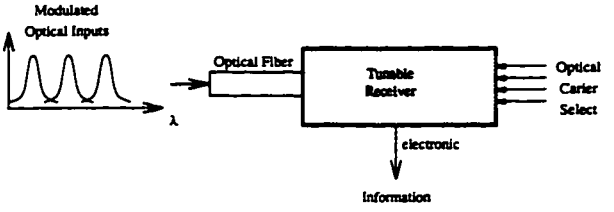


Figure 2.10: Stylized Tunable Receiver Block

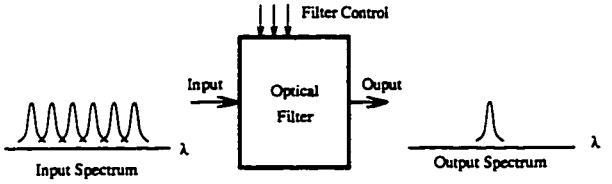


Figure 2.11: Stylized Tunable Filter Block

them are tunable Fabry-Perot and Mach-Zehnder filters [Sen92]. An interesting class of filters are those which may be considered active filters. Such devices are usually based on acousto-optical or electro-optical effects [Sen92]. Another approach towards tunable filters is to use a laser operating below its threshold which can be utilized to selectively amplify and pass desired wavelengths. Table 2.5.4 compares a number of candidate filter technologies on the basis of resolvable channels and tuning speed. Notice that with current technology there is an inverse relation between tuning speed and number of resolvable channels.

2.5.5 Wavelength Selective Couplers

Wavelength selective couplers can be considered as a hybrid between the passive coupler and the active filter. Where normal passive couplers are capable of spliting

Technology	Number Of Channels	Tuning Speed
Laser Amplifiers	10	ns
Electro-Optical	10	ns
Acousto-Optical	100	μs

Table 2.3: Tunable Receiver Resolution and Speed Comparison

a fraction of input power between its outputs over all wavelengths, the wavelength selective coupler is capable of performing the same splitting function on a controllable subset of wavelengths. In contrast with the passive couplers, these devices are refered to as active couplers in the literature [TGB95]. Additionally, the split ratio is also a controllable function of the active coupler and can range between 0% to 100% coupling. The functionality of an active wavelength-selective tap, which may be based on acousto-optics [IK92], is shown in Figure 2.12. In this case, the coupler has direct access to a specific WDM channel or set of channels. In the figure, a set of wavelengths is shown passing through a active coupler. A single wavelength, λ_i, is coupled out of one fiber and routed to the desired output fiber. Following this, λ_i', at the same wavelength but possible containing new information, is re-inserted by the coupler as shown in the figure. It is also possible that multiple channels may be simultaneously extracted and re-inserted by the tap.

Couplers of this kind have been built using acousto-optic devices as discussed in [Che90, Bra90]. In the device described in [Che90], light from the fiber is passed through a polarization beam-splitter (PBS) which spatially separates the TE and TM components as illustrated in Figure 2.13. Following this, they propagate through a waveguide which is subjected to microwave acoustic excitations. The resulting Bragg diffraction rotates the polarization at one or more wavelengths. When applied to another PBS, the selected wavelengths are routed to a separate output from the others. These wavelengths can be changed under electronic control simply by changing the microwave frequency. An attractive feature of these device is that they are "fail-safe". That is, in the event of power loss to the device, optical energy will continue to flow through the device.

2.5.6 Intensity Modulation

The modulation of a laser's output at any wavelength can be divided into two major categories: direct and external [Gow93]. Direct modulation, as illustrated in Figure 2.14a, results in an intensity modulated signal. In such a system, information is coded and passed to a modulation driver which operates the laser's power supply. By varying

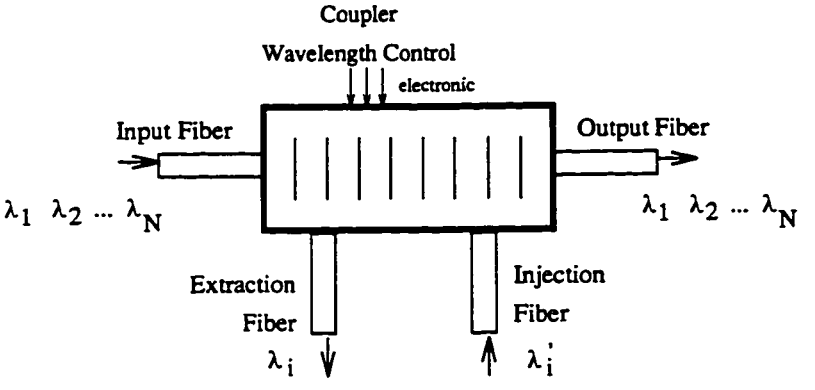


Figure 2.12: Stylized Wavelength-Selective Coupler

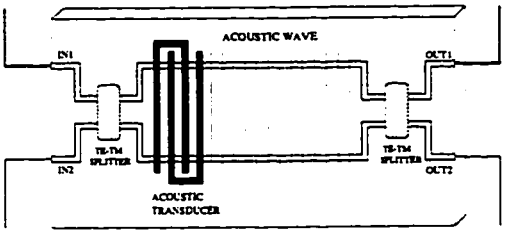


Figure 2.13: Acousto-Optic Wavelength Selective Coupler

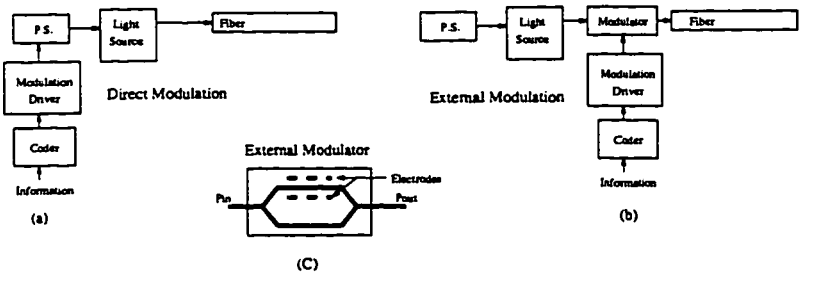


Figure 2.14: Modulation Of Laser Light

the voltage supplied to the laser in an appropriate manner it is possible to control the output intensity of the laser. The light from the laser is the coupled immediately to the optical fiber. An advantage of this modulation approach is that following the laser, light only needs to be coupled to the fiber. However, direct modulation schemes often suffer from "chirp" and have modulation bandwidth limitations [Gow93].

In an externally modulated system, as illustrated in Figure 2.14b, an external modulator is used to provide either intensity or phase modulation. In these systems, a laser is operated by means of a constant power supply such that the laser is continuously lasing at the desired output power level. The light from the laser is then coupled to the external modulator. An external modulator can be based on an electro-optical effect. Modulators such as the Mach-Zehnder inferometer [Gre93] shown in Figure 2.14c would equally split incident light through two equidistant branches and recombine the light at the output of this two port device. On one of the branches, however, would be a pair of electrodes which when excited by a voltage would manipulate the refractive index of that branch, effectively elongating or shortening the light path down that branch as desired. Of course in most materials this electro-optic effect is negligible and special care must be taken to select a material whose electro-optic effect is well known, such as lithium niobate. As a result, when the light waves traveling the two branches recombine, they do so constructively or destructively thus providing a means of intensity modulating the laser light. The modulated light is then coupled to the fiber from the external modulator. These modulators are capable of high modulation bandwidths, however, light does need to be coupled from the laser to the modulator and then to the fiber which can result in a drop in available output power.

2.5.7 Optical Amplifiers

To offset the effect of optical losses as a result of various causes including coupling losses, a regenerator can be inserted in the optical path. Often, this is done by detecting the optical signal, converting it into electronic form and decoding the digital information contained in the signal. The information is then re-coded and re-transmitted,

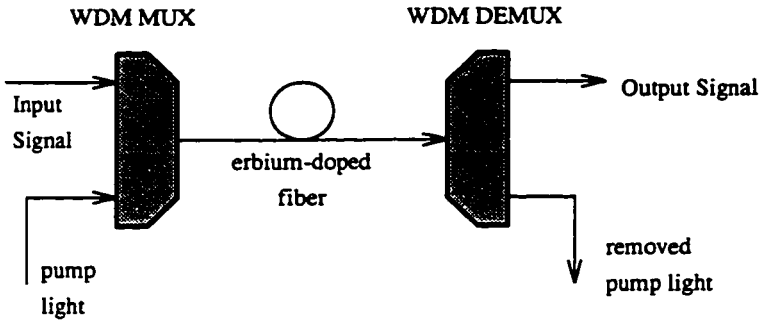


Figure 2.15: Erbium-Doped Fiber Amplifier

thus reconstructing the original optical signal. This approach is vary valuable, however these devices are expensive to construct, espically in a WDM environment. Furthermore, such amplifiers/regenerators are limited to a single modulation technique. For example, when an analog signal is used the regenerator becomes ineffective since the device has no idea as to what the signal should look like.

In order to reduce the cost of optical amplification and create a modulation independent device, various all-optical amplifiers have been examined in the literature. Such a device would require no electronics and could amplify the optical signal directly in the optical domain. Such an approach is useful when considering power-limited systems since the signal is only amplified but not reconditioned. Two approaches of interest are the traveling-wave laser amplifier [Sen92] and the resonant Fabry-Perot laser amplifier [NS90]. A travelling-wave amplifier utilizes a material with optical gain such that it is practically a laser without the mirrors. The Fabry-Perot laser amplifier is similar except that it utilizes mirrors and operates at just below the lasing threshold. In practice, achieving adquate gain with low noise has been problematic. Furthermore, the gain in these devices tend to be polarization dependent.

An alternative approach capable of high gain with low noise is the erbium-doped fiber amplifier [DSP87]. Figure 2.15 illustrates such a device. A pumping laser is used to activate the erbium-doped material into a gain state. The pumping laser is typically a laser diode with a shorter wavelength than that of the signal to be amplified. The input light and the pumping light are coupled to the erbium-doped fiber via a multiplexer. The longer wavelengths of the input signal then stimulate

the excited erbium atoms thereby amplifying the signal. A WDM multiplexer is then used to separate the pump wavelength from the signal wavelength so that the pump light does not interfere with the detection process. Such amplifiers may prove effective in WDM systems where optical power constraints have previously been a deterent to many design architectures.

2.6 Discussion

This chapter has served as an introduction to networking and the associated components used to build a photonic network. The chapter began by outlining some basic network principles and then proceeded to distinguish circuit switched networks from packet switched networks. The need for a media access protocol, as relevant to the OSI model, was discussed with particular relevance to multiaccess broadcast type networks. Finally, various optical components relevant to this thesis were described and discussed briefly. It is clear that optical device technology is sill an immature technology with much room for progress. Many of the described devices are research grade device with commercial availability of these device still several years away.

The next chapter will begin a review of current network research literature relevant to packet switched optical networks. In particular, special attention will be focused upon media access protocols. Physical level aspects of the networks are discussed in relation to the constraints it places on the media access protocols and the associated performance. The discussion begins with high speed single channel networks and then moves to optical networks and finally WDM networks which utilize those devices described earlier in this chapter.

Chapter 3

Media Access and Design Literature Review

In this chapter a literature survey of designs and media access protocols for high speed networks is put forth. Although not all of the networks covered here are of the WDM type, or even photonic, they are included as they form the historical basis for many designs that have followed. Specific attention has been placed on designs that are based on ring or bus topologies operating in a packet switched mode. The discussion begins with single channel designs which may or may not be implemented in optics. WDM networks will then be studied and new considerations associated with parallel channel architectures implemented in optics will be examined.

3.1 First and Second Generation Networks

In the past, many networks have been designed around the ring and bus. An examination of the literature suggests that these topologies make up the majority of early network designs. Such early networks are often considered to be first generation designs [Muk91a]. Loosely speaking, networks of the first generation are those based upon traditional electronic media such as copper wiring or free space radio. In a second generation network, the traditional electronic media is simply replaced by optical media. In this generation, the protocols remain identical to the first generation

networks, except that the physical level components are updated. The goal of second generation networks is to take advantage of many of the favorable properties of optical fiber such as low attenuation losses and immunity to electrical interference. This section describes a handful of the most common first and second generation networks and their associated media access protocols.

3.1.1 TDMA

Time Division Multiple Access (TDMA) techniques [Tan89] are one of the most simple methods of sharing a common communications medium. In this method, time on the medium is divided into slots. Each station is then pre-allocated a slot or set of slots in the hopes of satisfying the stations communications requirements. Systems employing TDMA as its media access control protocol require very little hardware as long as each station is synchronized to a common clock. In such a case, a user station can transmit using only a single transmitter. As a result of this simplicity, bandwidth efficiency, the usable bandwidth as a fraction of the total available bandwidth, can be 100%. Furthermore, TDMA protocols are applicable to a wide range of topologies including the ring and bus. Unfortunately, when stations do not fully utilize their allocated slots, those slots will be wasted. To remedy this situation, other media access protocols such as polling prove interesting.

3.1.2 Polling

In a polling type system [Sta93], a central station queries each user station, one at a time, as to whether or not the user station has any data to transmit. The queried station should then respond to the central poller with its current status. User stations are only permitted to transmit when the central polling station gives the user permission to transmit. Many variations of the polling media access protocol exist covering aspects ranging from the polling order, to the specific responses user stations generate. To engage in any polling algorithm requires that each station is equipped with at least a single transmitter/receiver pair in order to communicate with the polling station. Again, like TDMA systems, polling systems are applicable to a wide

range of topologies including the ring and bus. In contrast to the TDMA systems, when a polled station has no data to transmit, only a small amount of overhead is lost due to the polling process. This results in better bandwidth efficiency when user stations generate bursty traffic. However, because of the overhead associated with the polling process, many limits are placed upon the scalability of a polling system.

3.1.3 Random Access

Networks based upon the previous two media access protocols are known as contention free networks since all transmissions are coordinated such that is impossible that two or more stations will transmit on the medium simultaneously. For random access protocols there is no such coordination and as such they are often referred to as contention based networks. Many variations of random access networks have been investigated in the literature including Aloha, Slotted Aloha, and CSMA/CD [90285].

Stations in a network employing the Aloha protocol are permitted to transmit onto a shared medium whenever they have data to transmit. Normally, transmissions are broken up into packets. As long as only one station attempts to transmit a packet at any given time, the transmission will be successful. However, if two or more stations begin transmitting simultaneously, or one station interrupts another's transmission, all transmissions will likely be destroyed. Such an event is referred to as a "collision" of packets. Transmissions which are correctly received by the destination station are acknowledged. However, collided packets cause no response at all from the receiving station. Errors due to collisions are detected through the use of some amount of overhead in each packet which is used to implement an error detection scheme. Thus when a transmission is acknowledged, the process is complete. If an acknowledgment is never received after some predetermined timeout period following a transmission, the original transmission is assumed to have been destroyed and is repeated until an acknowledgment is received.

The Aloha protocol tends to be fairly efficient at light loads. That is, when there is very little packet traffic on the network only a small fraction of the total bandwidth is wasted due to the protocol itself. However, as traffic increases, overhead due to

packet collisions also increases, thus decreasing the efficiency of the network. Under the assumption of an infinite station population generating Poisson packet arrivals, the channel throughput can be written as

$$S_{aloha} = Ge^{-2G}, \tag{3.1}$$

where G is the offered load due to all transmissions [Tan89]. From Equation 3.1 it is found that the Aloha protocol can have at best a maximum channel capacity of 18% under the assumed conditions which occurs when the offered load is approximately 0.5 packets per packet transmission time.

An improvement over the basic Aloha protocol is the Slotted Aloha protocol. In this version of the protocol time is slotted and packets are only permitted to begin the transmission of a packet at the start of any slot. Here the length of a data slot is equal to the amount of time it takes to transmit a packet. This has the effect of reducing the amount of time a packet is "vulnerable" to collisions due to other stations beginning their own transmissions. Using the same assumptions as above, the channel throughput for the Slotted Aloha protocol [Tan89] can be expressed as

$$S_{slotted} = Ge^{-G}. \tag{3.2}$$

The resulting system doubles the maximum capacity of the Aloha protocol to 36% when the offered load approaches 1 packet per data slot. Still further improvement in the achievable maximum capacity is possible when carrier sensing and collision detection are added to the protocol. Such protocols are called carrier sense multiple access (CSMA) with collision detection (/CD).

Carrier sensing consists of each station monitoring the medium for transmission energy from other stations. If any such energy is sensed, the station will refrain from attempting any transmission until the medium is sensed to be free. This ability results in fewer collision occuring and improves performance. Collision detection is the ability of a station to monitor its own transmissions and to determine if that transmission has collided with another packet. These added features make the calculation of capacity in a CSMA/CD system much more complicated than for the

Aloha protocols. However, it can be found that the maximum achievable capacity of a CSMA/CD system, under the right conditions, will approach 100%. A fair approximation of the protocol's maximum capacity [Aca94], where a is the network's normalized end-to-end propagation delay, is given by

$$C = \frac{1}{1 + 2(e-1)/a}. \tag{3.3}$$

The IEEE 802 committee has chosen the CSMA/CD media access protocol as one of its standards for local area networks mainly on the basis of its high achievable capacity. However, being a random access method, the delay a station might have to endure in order to finally and successfully transmit a packet can be highly variable. In fact, this delay can be unbounded. For many, this is an unacceptable situation.

3.1.4 Token Passing

The demand for predictable and bounded access delays has made the use of random access media access protocols such as CSMA/CD inappropriate in many environments. There are many situations for which bounded, deterministic access delays are required such as Multimedia communications and other real time applications. A general approach that often satisfies the deterministic requirements of real time applications is referred to as token passing [Aca94].

With token passing protocols stations are only permitted to transmit packets when the station holds a token. Since there is usually only a single token for the entire network, the station holding the token is guaranteed that it will be the only one on the network attempting to talk. Thus there will be no chance of two packets colliding with one another. Of course the particulars of the passing of the token from one station to another and token management will vary among the various token passing protocols. In general, token passing can be applied to a wide range of topologies with minimal hardware requirements. Two of the most common token passing algorithms are found in the IEEE 802.5 token ring [80285] and the IEEE 802.4 token bus [80290a] networks. Finally, FDDI (Fiber Distributed Data Interface) [Ros89] is a second generation token passing network designed for optical media.

The IEEE 802.5 token ring network standard specifies that each station is connected to a ring via an active tap interface, the specifics of which are fully described in [80285]. By using an active interface it is possible for stations to process tokens and packets and remove or modify the bits in those elements as they propagate around the network.

Normally, when no stations have packets to transmit, the token simply rotates the ring in the "free" state. When a packet is generated at a station, that station must wait until the free token passes through the station, at which time the token is modified to be a "busy" token. Subsequent station seeing this busy token are aware that they may not modify the token or transmit any packets of their own. The station which claimed the token, however, is allowed to transmit a packet immediately following the busy token and may hold the token for an amount of time referred to as the token handling time. When the station has finished with the token, it is responsible for reseting the token to the free state. The free token then circulates to the next station on the ring.

Since there is a natural physical ordering to the stations, the access delay any station may experience is tightly bounded by the size of the network, the number of stations, and the token handling time. When these parameters are known, it is a simple matter to calculate an upper bound for the access delay. Also, because bandwidth is never lost due packet collisions, throughput and offered load are identical, and maximum capacity can be given as

$$C = \frac{N}{N + D/T}. \tag{3.4}$$

Here N is the number of user stations, D is the total round-trip delay of the ring, and T is the time it takes to transmit a packet [Aca94]. It can be shown that a token ring with a large number of user stations will always have a greater capacity than a CSMA/CD network.

A variation of the token ring called the IEEE 802.4 token bus [80290a] employs bidirectional bus and can be implemented by a wide range of technologies. Interestingly, the token bus protocol can be layered on top of bus running the CSMA/CD protocol as well. In contrast to a token ring, the token bus utilizes a logical ordering

to the user stations rather than the natural physical ordering present with the ring. As such, each station in a token bus network must know its predecessor and successor in the logical ordering to be able to accept and pass the token from one station to the next. However, in this version of the token passing protocol, when a station receives the token it does not modify it, but keeps it. The station then has the liberty to transmit as many packets as it wishes, subject to token holding constraints. When the token holding station has completed its turn with the token, it forwards the token to the station's successor station. The performance of a token bus network is similar to that of the token ring. The main advantage of the token bus, however, is that it can be implemented on simpler hardware as an active station interface is not required.

The Fiber-Distributed-Data-Interface (FDDI) [Ros89] is a network specifically designed for use with fiber optic links. The network consists of two counter rotating rings, only one of which is functional at a time. A FDDI network is illustrated in Figure 3.1. In the figure two counter rotating rings are shown even though in normal operating mode stations are only connected to one of the rings. This gives the security of having a backup link in the event of a cable or station failure and affords the ability of self healing in many situations. Stations are granted transmission permission based on a single token protocol similar to those described above. Similar to the token bus, a FDDI station employs an active station interface. However FDDI stations are capable of transmitting at a rate of 100 Mbps whereas token ring stations are limited to only 16 Mbps. Protocol efficiency is typical of all token passing protocols. A general disadvantage of FDDI and all token passing protocols is that token management can often be very complex and costly in the event that a token is lost. Furthermore, as the normalized end-to-end propagation delay of a token passing network increases, because of increased geographical span or a decrease in packet transmission time, the overhead required by the token passing mechanism begins to consume increasing amounts of the total bandwidth.

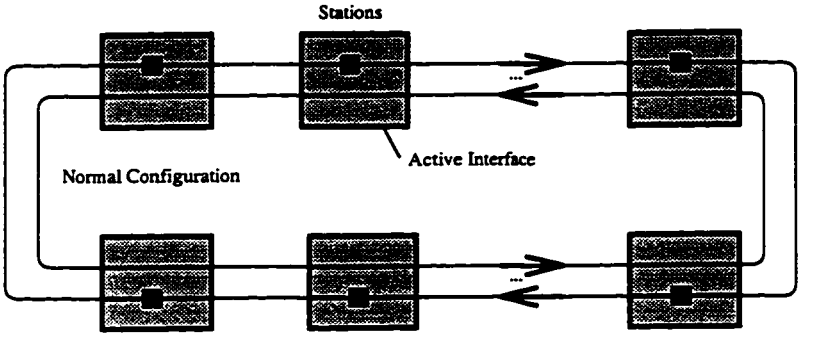


Figure 3.1: The FDDI Network Architecture

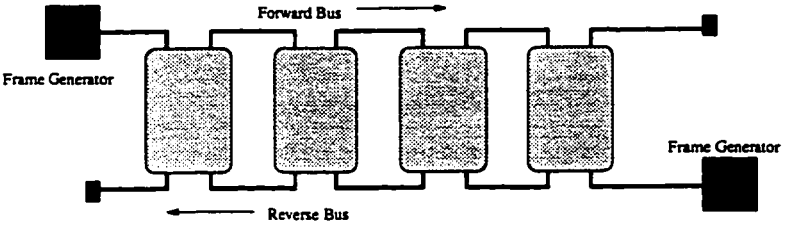


Figure 3.2: The Distributed-Queue-Dual-Bus Architecture

3.1.5 DQDB

The Distributed Queue Dual Bus (DQDB) [NBH88] is a network protocol designed to deal with the inefficiencies of CSMA/CD and token passing algorithms when operated at high data rates and covering large geographic distances. A DQDB network utilizes two uni-directional bus operating in opposing directions. The dual bus is illustrated in Figure 3.2. Unlike FDDI, each link in a DQDB network is used simultaneously. User stations can be implemented using passive taps to the network, although active interfaces are also acceptable. Each station requires a transmitter and receiver for each bus in order to execute the protocol.

At the leading end of each bus is a "frame generator" whose responsibility it is to create markers which sketch the boundaries for fixed length slots which are grouped into frames of 125 μsec each. Stations then access individual slots by executing a distributed algorithm such that the number of packets waiting upstream and downstream of each station is known by each station. Based on this information, each station knows the position of each packet it has generated in terms of a global queue

which is comprised of all packets awaiting service at all stations. Now the station can determine how many empty slots it should allow to pass by before it can write its packet into its own empty slot.

The distributed queueing algorithm is implemented by having each station maintain a number of counters. In each slot are a number of fields including a busy/free bit and a request bit. Each time a slot with the request bit set passes a station, a request counter for that bus is increased. Each time a free slot passes a station the request counter for the opposite bus is decreased. In this way, stations monitor the number of upstream and downstream packets in the global queue. When a station has a packet ready for transmission, say on the forward bus, it makes a request by setting a previously free request bit in a slot on the opposite bus. It then copies the value of the request counter for the bus it wishes to transmit on into a count-down counter and resets the request counter. The station then allows a number of free slots on the desired transmission bus to pass the station as indicated by the count-down counter, after which the station can use a free slot for its own transmission.

The capacity performance of a DQDB network can approach 100% and supply the required deterministic and bound delay performance required by many applications. However, it has been shown that under some common conditions, the distributed queueing algorithm can be quite unfair [HCM90]. To re-establish the desired fairness, the IEEE in 802.6 [80290b] has adopted bandwidth balancing as an optional feature of the standard DQDB protocol. The mechanism simply requires that stations allow an extra free slot to pass without affecting the request counter after every BWB'th transmission, where BWB is the bandwidth balancing factor. Since a station can now use up to $BWB/(BWB+1)$ of the available bandwidth, the capacity for a network with N stations then becomes [BV95]

$$C = \frac{BWB}{N * BWB + 1}. \tag{3.5}$$

Like with many networks, once a slot is used it propagates to the end of the network even though it has delivered its packet. In some designs "destination release" is implemented such that a slot is freed for use by downstream stations by the receiving station. Although stations in a DQDB network lack the active ability required for

destination release, special stations called erasure nodes [Has94] can be included in the system. Each erasure node is implemented with an active interface capable of converting used slots into empty slots. A node that has received a packet from a slot is required to write a specific bit in the slot which indicates to the erasure node that the slot can be freed. The erasure node then must also remove a request from the network so that other free slots are not wasted. The resulting system can benefit from greatly increased capacity which is dependent on the number of stations, the traffic patterns, the number and placement of erasure nodes, and the erasure algorithm.

3.1.6 Orwell

The Orwell system [FA85] consists of stations in a ring topology. The idea is that slots propagating the ring can be randomly accessed if they are marked free. Access to slots is regulated by a quota mechanism in which each station is allocated a number of slots which may be used. Once a station has exhausted its quota, it must suspend its transmissions until a quota reset is issued. In an Orwell system stations are required to implement destination release so that there will be a supply of empty slots circulating the ring. Clearly, an active station interface is required for an Orwell network.

Quota resets are generated via a distributed mechanism whereby empty slots passing a station are marked with a trial bit. Stations accessing empty slots marked with a trial bit will clear the trial bit. Thus any slots with a marked trial bit returning to the marking station indicates that all of the stations on the network have either exhausted their quotas or the network has become idle. In either case, a reset is issued by the marking station and the transmission quotas at all stations are reset.

Performance for the Orwell protocol is characterized as being deterministic due to its cycle based access mechanism. Even though quota resets require at two end-to-end propagation delays to complete, the mechanism can be quite efficient when the aggregate transmission quota of the network is large. Furthermore, the use of destination release in the protocol allows for increased network capacity relative to other first and second generation ring designs.

3.2 Third Generation WDM Networks

In the third generation, networks are characterized by the use of optical fiber for its unique properties in order to meet the requirements of high-bandwidth applications [Muk91a]. One of the most important properties of optical fiber is the ability to multiplex several independent wavelengths on a single fiber. Many third-generation networks have emerged recently covering a wide range of designs. Of primary concern here are the single-hop, multiaccess broadcast networks.

For third generation networks, the use of many traditional protocols such as contention based CSMA/CD and many non-contention protocols such as token ring are limited since they cannot be scaled to operate efficiently at high speeds [Sta87]. In contention based protocols a relative increase in end-to-end propagation delay with respect to packet transmission time results in a large proportion of bandwidth being wasted in packet collisions. Token based protocols also often suffer from bandwidth wastage because of the overhead involved in token transmission as data rates are increased [RCS89]. To overcome these limitations many new high speed protocols have been developed specific to the needs and features of WDM including the complexity of managing transmissions on multiple channels.

In the past, many single-hop optical networks have been proposed based upon passive optical star couplers [Muk92]. In these designs, each station is given a separate inbound and outbound fiber to the star as illustrated in Figure 3.3. Optical energy impinging on the star is distributed across all outputs (subject to various losses), thus creating a broadcast-and-select network. Although such networks have a simple implementation, the multi-channel environment introduces many protocol complexities when minimal station hardware is desired.

A current proposal entitled the Distributed Channel Controller Network (DCCN) is one attempt at reducing station hardware in a passive star environment. DCCN [JT93], in contrast to many current proposals, has managed to reduce the hardware requirements of a user station to a single tunable transmitter and a single fixed tuned receiver. This is made possible through the functionality of channel controllers. Each

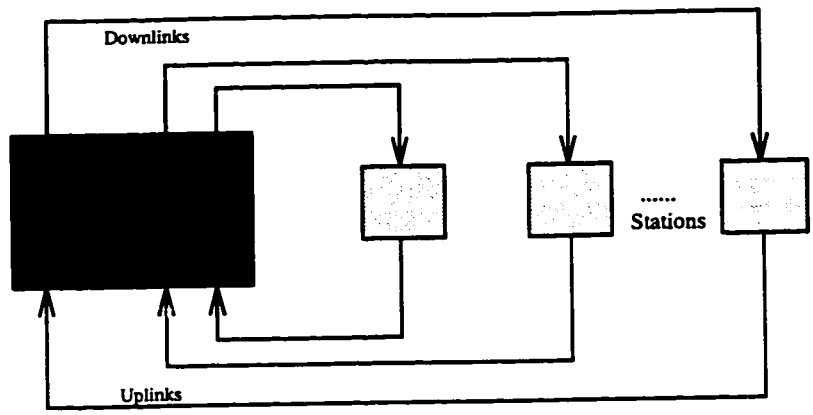


Figure 3.3: Passive Optical Star

of the controllers connects to one of the WDM channels in a one-to-one mapping. Media access is based upon communication between the user stations and the controllers which grant access to slots in a dynamic fashion.

3.3 Ring and Bus WDM Networks

Recently, more attention has been focussed on multi-channel ring and bus networks [SD93, IK92, BM93]. Although such networks have many advantages from a protocol standpoint, they have been impractical in the past because of their poor optical power budgets [Gre93]. Current interest is motivated by advances in both broadband traveling-wave and doped-fiber amplifiers [Sen92, DSP87]. In the future, these devices should permit practical photonic bus and ring LANs and MANs. The following sections summarize some of the interesting current work in multi-channel ring and bus networks.

3.3.1 EQEB

The Estimated Queue Expanded Bus (EQEB) [CM92] protocol is a multi-channel version of the Distributed Queue Dual Bus protocol. EQEB is based upon a dual bus topology with a headend placed at each end of the dual bus as depicted in Figure 3.2. User stations use a fixed channel for reception but can transmit on any of the

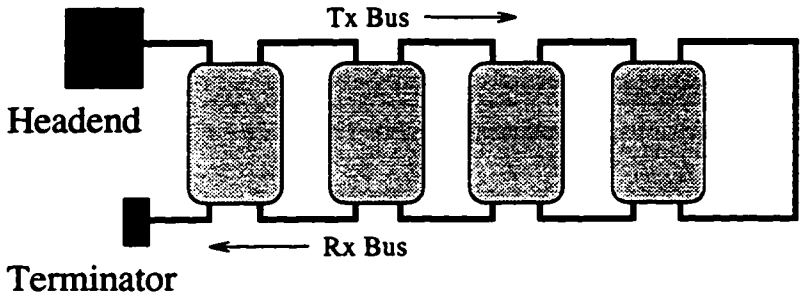


Figure 3.4: Folded Bus Topology

WDM channels. Each user station requires two tunable transmitters, two fixed tuned receivers, and two tunable sensors. Each headend requires a fixed tuned transmitter and receiver for every WDM channel. Basically, a single ordered queue is established across the network so that individual stations wishing to transmit data are granted access to the network in a first come first serve manner. Each station maintains information on the state of the network so that when a station needs to transmit, it knows how many other stations with data to transmit are queued ahead of it. Using this information, a station is able to determine when it is permitted to access the network.

3.3.2 FairNet

FairNet is a WDM-based folded bus, illustrated in Figure 3.4, where all channels are slotted and synchronized [BM93]. Each user station uses a tunable transmitter, a fixed receiver pre-assigned to a home channel, and a tunable sense receiver. At the head of the bus is a headend used for the generation of the synchronized slot structure. Each station has a set of transmission probabilities representing the probability that a station transmits on a given channel in a given slot. Backlogged stations transmit if the BUSY bit of a slot on the destination channel is not set, simultaneously setting the BUSY bit. Values of the transmission probabilities are determined from the network traffic distribution and equity constraints.

3.3.3 D-Net/C

D-Net/C [SD93] is a multi-channel extension of the D-Net [TC83] folded bus. Each user station requires a tunable transmitter, a tunable sense tap, and a fixed receiver. A generator is located at the left end of the bus which launches a locomotive on a outbound bus channel when the generator senses an end-of-train on the corresponding inbound bus channel. Backlogged stations wait for the locomotive and transmit when it is seen. A user station simultaneously uses its sense tap to detect upstream transmissions, and aborts its transmission upon detecting a carrier. If a collision is detected, stations wait for an end-of-carrier on the same outbound bus channel and re-attempt transmission.

3.3.4 B-TDMA

B-TDMA [SD93] is another folded bus protocol based on a time multiplexed multi-channel access scheme similar to AMTRAC [CG88]. Each user station on the folded bus requires one tunable transmitter, a fixed receiver, and a tunable sensor. A generator station is also required containing a transmitter for each channel. Time is divided into slots of length equal to the propagation delay between adjacent stations. Each channel cycle is then divided into $2M - 2$ slots allocating each station one slot per cycle. User stations are then allowed to begin transmission on a channel in its preassigned slot.

3.3.5 ACTA

The Adaptive-Cycle Tunable-Access (ACTA) protocol [Che93] is a multi-channel network protocol which may be applied to a bus or ring topology. Each user station requires two tunable transmitters and two fixed tuned receivers. In addition a headend containing a fixed tuned transmitter and receiver for each WDM channel is required to generate control signals for the media access protocol. Transmission occurs in a mechanism similar to Fasnet [LF82]. Any user station can write into the first N_q available slots immediately after it has seen a Cycle-Start, where N_q is an arbitrary quota decided in advance for each station according to priority. The Slot-Occupied

field is then set after a slot has been written into. Transmission is suspended until the next Cycle-Start when a station has exhausted its N_q in the current cycle, a station encounters a packet going to a different channel, or the outgoing queue has been depleted.

3.3.6 Protocol Evolution

Many other techniques for media access have been developed including hybrid schemes of all of the above described networks. For example, [LK92] developed an extension of the DQDB protocol which uses a control channel to coordinate the use of the other data channels running the DQDB protocol. Designs of this nature typically require an array of fixed-tuned optical transmitters and receivers or numerous tunable transmitters and receivers. Of primary concern is that of hardware and media access protocol complexity at the stations. It is essential to reduce the hardware required if the cost of the station interface is to be minimized. The absolute minimum transmission hardware would be a single transmitter and receiver per station, with one of the devices being tunable. The previously discussed ring and bus networks require multiple receivers or tunable sensors and because of the liberal use of transceivers they all can achieve a capacity very close to 100% per channel.

Part of the work presented in this thesis proposes multi-channel network architectures entitled BTMSC (Blind Transmitter Mini-Slot Contention) and BTMSR (Blind Transmitter Mini-Slot Reservation). The motivation behind the designs is to reduce the hardware requirements of the user stations while maintaining reasonable capacity performance. This is achieved using an intelligent headend which attaches to each of the WDM channels. Each user station requires a single tunable transmitter and one or two fixed tuned receivers for the ring and folded bus topologies respectively. The distinguishing feature between BTMSC and BTMSR is the reduction in the number of passive couplers required at each user station.

3.4 Designs Using Active Couplers

The previously discussed optical designs relied on passive coupling technology in order to connect stations onto the network. While this approach is very efficient in terms of optical losses, it fails to recognize the flexibility and properties of optical fiber and newly emerging device technologies such as the active couplers discussed in the previous chapter. The designs resulting from the use of active couplers can afford networks some very interesting and valuable properties including fault protection, wavelength reuse, and wavelength routing. Furthermore, many protocol options become available that were previously available only to electronic systems which could implement active station interfaces. The following sections summarize work done on networks utilizing active coupling technology in one form or another.

3.4.1 Self-Healing Rings

In [WL90], the survivability due to segment faults in ring architectures was examined. In the paper, the use of add/drop multiplexers to create a class of self-healing ring architectures for use with SONET (Synchronous Optical NETwork) [SHP91] applications was examined. In this case links employing SONET, which is a physical layer and signaling standard, are afforded redundancy switching in the event that a particular link is damaged. Essentially the add/drop multiplexers, which could be implemented using acousto-optic devices, are used as protection switches. Normally, the switch will allow data to flow through it and to the default output link. In the event of a link failure, the switch is set so that one of the backup links is used. Alternatively, in the case of multiple link failures, the topology can be manipulated so as to isolate the faults from the rest of the network. Although this approach can significantly improve the reliability of a network, it does not improve the capacity of the system, nor does it exploit the potential WDM uses of active couplers.

3.4.2 Single Channel Designs

In the paper [Muk91b], the author asks "Why not make the interface at the various stations active?" The resulting open ring and active bus designs are single channel networks which utilize active add/drop multiplexers which too may be implemented with acousto-optics. The new structures have the advantages of suffering from fewer synchronization constraints and boast higher capacity because of the ability of stations to implement spatial reuse in the protocols. In a similar attempt, the Bellcore METROCORE network [AGI88] uses active station interfaces in conjunction with optical fault bypass devices to achieve the survivability and protocol performance advantages of active couplers.

3.4.3 Monolithic Wavelength Routed Structures

Many have recognized that active coupler technology such as acousto-optics is capable of routing wavelengths between a set of output ports. Based on such functionality, a number of networks have been proposed. In [ea93a] a scalable multi-wavelength, multihop optical network is proposed. To achieve scalability into the WAN environment, the proposal exploits wavelength reuse which is made possible through the use of wavelength cross-connect devices which are implemented using acousto-optic couplers. The network fabric consists of many such wavelength-routing cross-connects managed via a common network control which provides the ability to dynamically rearrange the fabric. The goal of this approach is to meet changing traffic, service and performance requirements as well as provide fault-tolerant service.

Another approach utilizing wavelength routers is proposed in [ea93b]. In this paper, a three-level hierarchy is proposed wherein extensive reuse of wavelengths between levels is achieved via the use of wavelength routers and wavelength converters and/or switches. Initially, it is assumed that these wavelength selective devices will operate in a static or quasi-dynamic mode until efficient configuration algorithms are developed. Other designs incorporating wavelength routers investigated in the literature include [Jos96, Jan96, AT95].

3.4.4 Evolution Into LAN Designs

A multichannel ring network using wavelength selective acoustically tunable active couplers as coupling devices within stations was examined in [IK92]. This paper demonstrated that a migration of active coupler technology into the design of a WDM network station was natural. Each station on the network consisted of a single active coupler, one transmitter fixed tuned to a particular wavelength, and a number of receivers equal to the number of wavelengths in the system. The active couplers are configured to select a proportion of optical energy from the fiber for each wavelength. For the wavelength that the station uses to transmit on, the coupler is set so that all incoming energy is coupled off the fiber. Because of device symmetry, all optical energy generated at the station on the transmitting wavelength is coupled back onto the network, thus replacing the previous information on that wavelength. For all other wavelengths, the coupler is configured to extract only the proportion of energy from the fiber that is required to drive each receiver in the station. Thus each station coupler is configured once and remains static during normal network operation.

It was demonstrated that for networks of this class, the capacity available to each station was 100% of the achievable bit rate. This is possible since the designs are connection-oriented where users access the medium without any contention or necessary protocol. Of course in connection oriented designs, similar to circuit switched networks, when a particular connection is not being used by its dedicated station, the capacity of that link is wasted. Furthermore, the economics of requiring multiple receivers at each station may prove prohibitory in LANs and MANs.

The second part of this thesis proposes alternative designs using active couplers in a packet switched network. The goals of the new designs is to improve the economics of the stations by reducing hardware requirements without sacrificing capacity performance. To do so will require that the tunable aspects of active couplers be fully exploited. Two protocols are explored which are applied to the ring, folded bus, and dual bus topologies. Also, constraints due to the tuning process of the active couplers are formalized and protocols which satisfy the tuning constraints are introduced.

3.5 Discussion

This chapter performed a literature survey regarding ring and bus network protocols and designs. The survey began with first and second generation networks and discussed many media access protocols and designs including their capacity performance. The chapter then proceeded to discuss third generation photonic WDM designs for which capacity performance far exceeded the designs of the first and second generation. The literature presented represents the important works and issues for the state-of-the-art WDM ring and bus networks that are now being researched. Particularly interesting is the division between work conducted on passively coupled networks and actively coupled networks. It is obvious that although passive couplers represent a mature technology, progress in active coupling devices can lead to designs which exploit properties of wavelength and spatial reuse which could ultimately render passive couplers obsolete.

This thesis will explore designs employing both passive and active couplers, with the particular goal of reducing the hardware requirements for user stations in ring and bus topologies. Also of concern are the relative strengths of passive designs versus active designs and the type of media access protocols which can be performed with each approach. Specific issues arise due to the desire to reduce hardware counts. The next chapter begins the original content for this thesis and deals with the design and performance of ring and bus networks employing passive couplers.

Chapter 4

Passively Coupled Architectures

This chapter begins the original content for this thesis. It leads with an overview for passively coupled networks and the motivations for the presented designs. The proposed network architectures are then described in terms of their physical implementation aspects and comparisons are made with other similar WDM networks. Following this, the functions of the headend will be described along with some of the partitioning algorithms and HOME channel reconfiguration schemes. Specific operating modes pertaining to each of the topologies will then be given. Two separate protocols, BTMSC and BTMSR, will be described, giving details of the frame and slot structures.

Finally the performance of the ring and folded bus networks based upon the passive coupling technologies discussed in this chapter is examined. Approaches will include both analytical and simulation techniques. Capacity and delay performance indices are derived for the various systems followed by an exploration of the optical link budgets for the systems. Results give an indication of the relative strengths of the designs.

4.1 Overview

Recently, increased attention has been placed on multi-channel ring and bus networks [SD93, IK92, BM93]. Although such networks have many advantages from a protocol

standpoint, they have been impractical in the past because of their poor optical power budgets [Gre93]. Current interest is motivated by advances in both broadband traveling-wave and doped-fiber amplifiers [Sen92, DSP87]. In the near future, these devices should permit practical photonic bus and ring networks.

The research quoted in [Muk92] deals with a number of issues specific to single-hop WDM networks. Of primary concern is that of hardware complexity at the stations. It is essential to reduce the hardware required if the cost of the user station interface is to be minimized. The absolute minimum complexity in transmission and reception hardware at a user station would be a single non-tunable transmitter and receiver. However, many current photonic bus/ring designs require multiple fast-tunable transmitters and/or receivers in order to achieve dynamic packet-switched operation [SD93, Che93, CM92, CSR90]. Also of importance with bus and ring networks is the number of taps placed onto the network at each station. It is these taps that are the limiting factor to the optical link budget when considering a LAN design. A reduction in the number of taps will increase network extensibility and delay the point at which photonic amplification is required.

In this chapter, four different WDM bus/ring designs are considered. All four have user station hardware designs with various reductions in the number of tapping points, number of transmitters and receivers and their tunability requirements. With each reduction in hardware, protocol complexities and performance reductions are introduced which will be considered. The designs thus give an indication of the cost/performance tradeoffs which are possible as the amount of hardware is reduced at the user stations.

The first pair of user station designs share a protocol referred to as Blind Transmitter Mini-Slot Contention or BTMSC [GT94]. The second pair use a protocol referred to as Blind Transmitter Mini-Slot Reservation (BTMSR) [GT95] which is a modification of the DCCN [JT93] protocol to support ring and folded-bus topologies. The use of this second protocol is the necessary result of a reduced number of fiber taps at each user station. All of the designs use a system headend containing a set of channel controllers, each of which attaches to one of the WDM channels. The controllers are required in order to provide synchronization and assist in station media

access in various ways. The objective in all cases is to simplify the hardware and protocol requirements of the user stations as much as possible while moving a large portion of system complexity to the headend.

4.2 Network Architectures

In the networks considered, the BTMSC and BTMSR protocols are each applied to the folded-bus and ring topologies, resulting in four different user station designs. In all cases, it is assumed that there are C WDM channels and N user stations per channel where typically $N \gg 1$. For convenience, we assume that the number of stations per channel is equal, but this is not a requirement.

4.2.1 Station Design

The user station designs considered are applicable to the folded bus and ring topologies. Common to all user station designs is a single fast-tunable transmitter and fixed-tuned or slowly-tunable receiver. The folded-bus designs require an additional receiver per station which is also fixed-tuned. Note that none of the designs require any additional fast-tunable receivers or fast-tunable "sense taps". The tunable transmitter is thus the only fast-tunable device at each station.

In Figure 4.1 the interconnection of system components is illustrated for all of the systems. To reduce the number of figures, the station configurations for each of the four designs are shown on the same figure. In an operating network of course, all stations would have only one of the station hardware configurations shown. It is envisioned that the optical fiber used in the designs will be single mode fiber and that each wavelength channel will operate in the Gbps range. As indicated in the diagram, all network attachments are made passively, using fused biconical taper couplers (FBTs) [Sen92]. It can be seen that the BTMSR architectures have reduced the required number of FBT couplers from what is required in the BTMSC systems. This reduction comes at the price of more complex media access protocols which will be discussed.

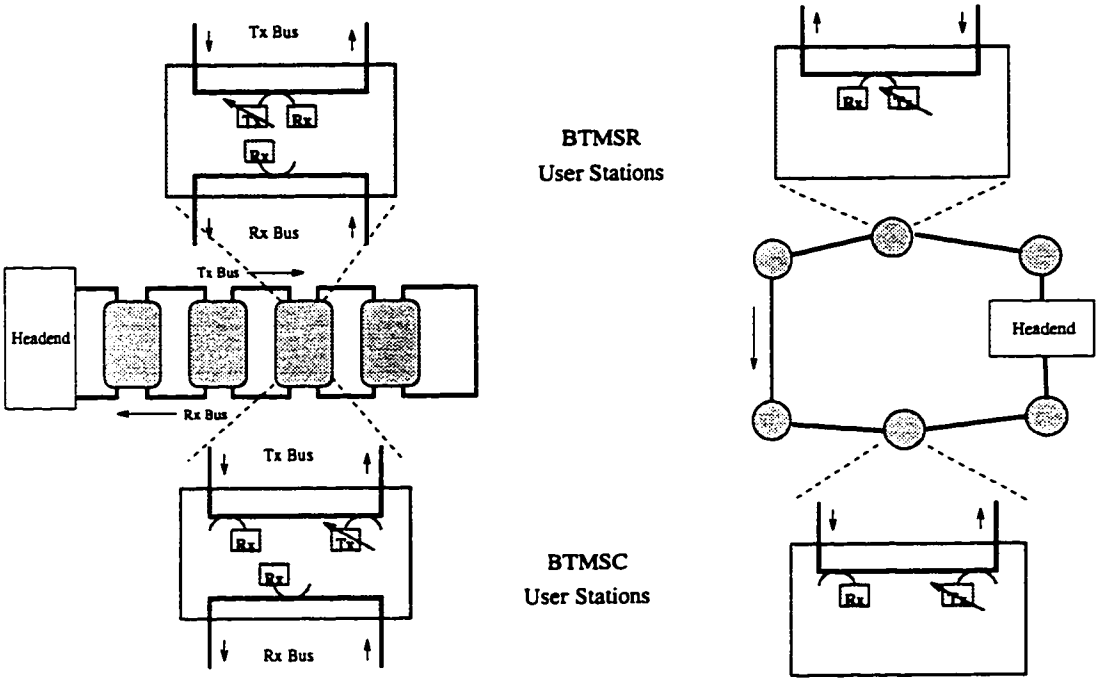


Figure 4.1: Physical Network Topologies

To extend the optical power budget of the system, it may also be necessary to perform some optimization of the coupling coefficients for the FBT couplers. Normally, each FBT coupler would be configured as a 3 dB device such that optical power at an input port is equally divided between both output ports, subject to any insertion losses. Clearly, this would not be an optimal case for all situations since each receiver in the system requires only a fixed amount of optical power. Here it becomes convenient to think of each coupler tapping off some fraction of power from the network and then coupling power back onto the network from its transmitter. Whenever more power than required is tapped from the network, it is lost to downstream stations. Since the coupling coefficient is a configurable parameter at device construction, it would be prudent to build couplers appropriate for the station. That is, each station should only take the fraction of power from the network that it requires to drive its receiver(s). However, due to the symmetry of the FBT couplers, the fraction of power that is coupled from the network is also the same fraction of power that the station can couple onto the network from its transmitter. Also, due to positional variations

in the stations, it turns out that each station would require a unique coupler configuration. In practice this may prove to be prohibitively expensive. A better approach may be to produce a single coupler configuration which is used at all stations. Under this constraint, the configuration should then made to be optimal. The exact calculations for determining coupling coefficients are given in Section 4.4.3 where optical link budgets are discussed.

A station will be able to transmit via its fast-tunable transmitter on an arbitrary wavelength channel, which is generally different from the channel its receiver(s) is tuned to. The station thus must be able to acquire synchronization information for the channel that its transmitter has "blindly" tuned to elsewhere. Thus the resulting designs have been named "blind transmitter" as in *Blind Transmitter* Mini-Slot Contention (BTMSC) and *Blind Transmitter* Mini-Slot Reservation (BTMSR). In order to derive the needed synchronization information, each station will continuously monitor one of its fixed tuned receivers from which the required timing can be extrapolated. For the ring systems this is done with its only receiver. On the other hand, the folded bus systems use the receiver on the Tx bus. This ability to extrapolate timing information is achieved with the aid of the headend.

In all cases, the headend generates a synchronized data-frame across all channels. On the ring, this frame propagates in one direction passing all user stations, finally returning to the headend. On the folded bus, the frame is released on the Tx bus and folds back along the Rx bus before returning to the headend. Since all stations are passively coupled to the network, the synchronized data-frame will remain aligned across all channels in time as it propagates the network. The result is that a station can monitor a single channel for synchronization information which is valid for all channels.

In the BTMSC networks, user stations have the ability to transmit while simultaneously receiving from upstream. This is possible since each station's receiver is coupled to the network via its own FBT coupler which is located upstream of the station's transmitter. As a result, any transmission emanating from a particular station will not interfere with any data being simultaneously detected by the station's receiver. In practice this may require the use of a fiber loop between the receiver and

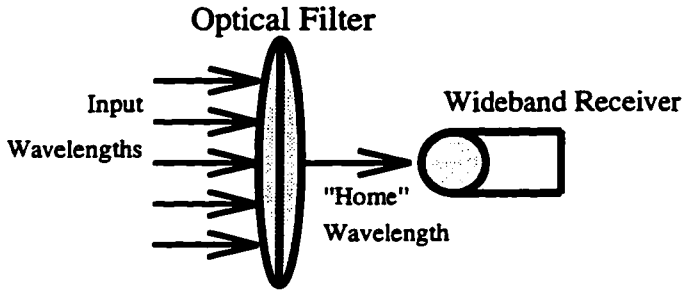


Figure 4.2: Receiver Module

transmitter couplers in order to provide enough delay so that the electronics have enough time to respond [NBH88] to any necessary protocol information. As a result, a BTMSC station will be able to interpret information in a slot and simultaneously use that same slot for its own transmissions.

In the BTMSR cases, the reduced number of couplers no longer permits this type of operation. Now, since only a single coupler is used to connect both the transmitter and receiver to the network, it becomes necessary to locate the transmitter upstream of the receiver. Whenever the station transmits, the station's receiver would detect both the station's own transmission plus any data from upstream. The resulting reception would be useless since the transmissions from upstream stations would be overwritten/garbled before reception. This limits the type of media access protocols which may be used in those designs.

In all four networks, all of a station's receivers are assumed to be slowly tunable and are pre-assigned to the same fixed "HOME" wavelength channel. The use of slowly tunable receive modules allow for the manufacture of identical components which are then configured on site. It is expected that, in order to provide the required wavelength selectivity for the receive modules, a tunable optical filter will be employed prior to a wideband receiver as illustrated in Figure 4.2. Because of the home channel construction, the filter needs to pass only one of the input wavelength channels to the receiver at the output of the filter. As an alternative to using tunable optical filters, fixed tuned bulk optic filters may be employed, resulting in a cheaper, but less flexible design.

4.2.2 Home Channel Assignment

Since each user station is assigned a HOME channel for reception, it is prudent that some policy be used for this configuration. The HOME channel assignment policy is flexible, however for convenience, a round-robin allocation scheme will be assumed for this thesis. In this scheme, station A is assigned the home channel $HOME$ as defined by

$$HOME = A \bmod C \tag{4.1}$$

where $1 \leq A \leq N$ and C is defined as the total number of wavelength channels. Recall that N is the number of user stations in the system. Using this scheme, a source station can easily determine the home channel of any destination station in a decentralized fashion. This results in two conceptual views of the network topology: a physical topology in which each station is physically connected to a single fiber, and a virtual topology for which each receiver and transmitter may be connected to a different wavelength channel. The virtual topology resulting from the round-robin allocation scheme for the folded bus and ring is shown in Figures 4.3 and 4.4 respectively. Each station in the figures is directly connected to its HOME channel via its slow tunable receiver(s). The tunable transmitter allows access to any of the wavelength channels on the system. Thus to reach any destination channel, a station would determine the destination's HOME channel, tune its transmitter to that channel and transmit, all in accordance with some media access protocol. Certain skewed traffic cases may, however, require alternative assignment policies or dynamic HOME channel reconfiguration schemes. For example, in a case where many "busy" stations exist, it may be necessary to spread those "busy" stations over the available channels rather than have them concentrated on only a few channels. The optimal solution for an arbitrary traffic skew distribution is a problem which is out of the scope of this thesis. Any such scheme, however, would be implemented within the network headend.

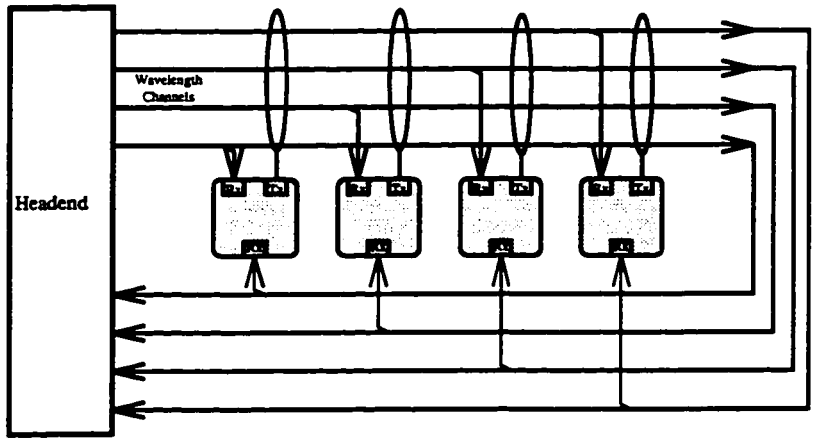


Figure 4.3: Virtual Folded Bus Topology

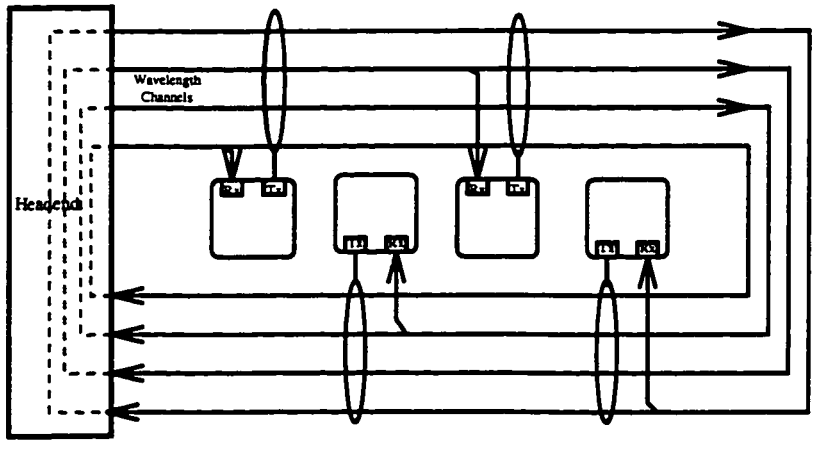


Figure 4.4: Virtual Ring Topology

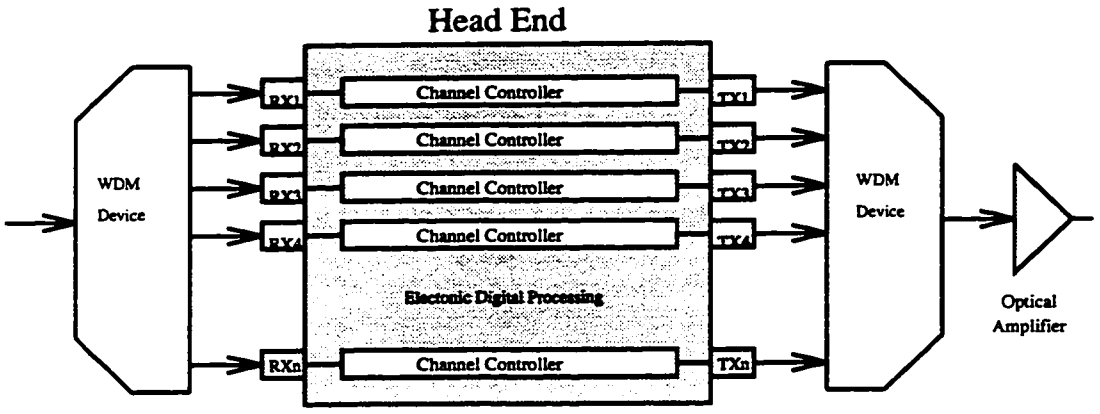


Figure 4.5: Headend Architecture

4.2.3 Head End Architecture

The headend is a device which initiates one end of the fiber and terminates the other. Within the headend is a set of channel controllers (CCs), each attached to a single WDM channel as illustrated in Figure 4.5. As shown in the figure, a WDM demultiplexer can be used to separate the wavelength channels for use by each CC receiver while a multiplexer is used to combine the wavelength channels from each tuned transmitter onto a single outgoing fiber. Alternatively, simple FBT optical splitters may be used to divide the optical power among tuned receivers similar to those used in the user stations. Such a device is shown in Figure 4.6. In this device, a number of FBT splitters are cascaded in order to provide a set of optical outputs. Each output is then passed through an optical filter in order to extract the desired wavelength for each channel controller. In this way, each channel controller can utilize a single wideband receive module. A similar device can also be used as a combiner for the transmitters. In this case, no special optical filtering is required. In order to overcome the splitting losses introduced at the transmitter side of the headend due to the WDM multiplexing device, an optical amplifier may be introduced. This amplifier would be an all-optical device, such as the erbium-doped device discussed in Chapter 2.5.7, and would follow the combiner within the headend. Other all-optical amplifiers may be placed elsewhere in the network as required.

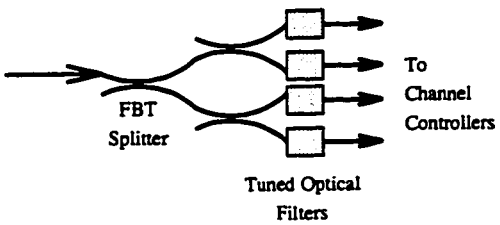


Figure 4.6: WDM Demultiplexer Implemented with FBTs

Within each channel controller are electronics and algorithms which may be implemented in hardware for making bandwidth allocation, generating and freeing slots, and doing system management and maintenance. Each channel controller is also responsible for synchronizing each channel in order to provide the functionality required for the blind transmitting that occurs at the user stations. Furthermore, in the ring designs the headend is also responsible for forwarding data slots around the headend. This is required in order to maintain full connectivity in the ring. For example, from Figure 4.4 it can be seen that if a station near the terminating side of the headend wishes to transmit to the station immediately following the headend, its transmissions must be able to pass through the headend without being removed. The headend protocol functions which are specific to each design will be discussed in the next section. It is envisioned that these protocols will be implemented in hardware using ASICs (Application Specific Integrated Circuit) in order to provide the high-speed functionality required for Gigabit per second communications. This high-speed functionality is further supported by the channel controller concept since each CC will operate independently, exchanging information with other CCs only infrequently and as needed. The result being that each CC needs only to be as fast as the channel which it is controlling. Even so, it still may be necessary to limit the complexity of the protocols for the sake of implementation concerns. In terms of cost, since it is expected that $N \gg 1$, the cost of a channel controller will be spread across a number of user stations.

The use of slow tunable transmitters and receivers for each channel controller's transmit/receive module allows for the manufacture of identical components which are configured during installation. This modular approach may lead to overall cost

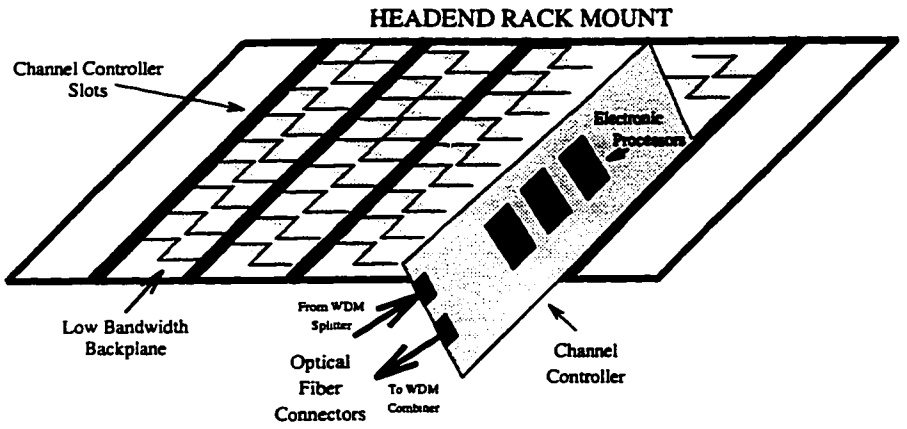


Figure 4.7: Rack Mounting of Channel Controllers

savings in the long run. In an actual implementation, the CCs would typically be connected via a low bandwidth backplane bus within a common rack so that they can share synchronization signals and configuration information. A visualization of the rack mount assembly is shown if Figure 4.7. A minimal arrangement for any of the passively coupled networks described would consist of a single channel, the headend, and user stations. Network evolution is thus accommodated through the addition of wavelengths and the modification of the headend to accommodate the additional wavelengths. Using the channel controller design, this can be achieved by simply plugging in an additional CC into the headend rack.

4.2.4 Optical Devices

The design of the BTMSC and BTMSR networks has kept in mind the need to reduce the required number of optical devices for user stations and the headend. Furthermore, it was identified that an important and influential factor in the design of passively coupled networks is the number of coupling devices employed. Table 4.1 shows the hardware requirements for BTMSC, BTMSR, and a number of other network designs in terms of the number of required fast-tunable transmitters, fixed-tuned receivers, fast-tunable sense taps, and fixed generation transmitters and receivers. N is defined as the total number of user stations and C is the total number of WDM

Protocol	Topology	Tunable Tx	Fixed Rx	Tunable Sense Tap	Fixed Gnr.Tx	Fixed Gnr.Rx
BTMSC, BTMSR	Ring	N	N	0	C	C
BTMSC, BTMSR	Fld Bus	N	$2N$	0	C	C
D-Net/C [SD93]	Fld Bus	N	N	N	C	C
Fairnet [BM93]	Fld Bus	N	N	N	C	0
ACTA [Che93]	Dual Bus	$2N$	$2N$	0	C	C
EQEB [CM92]	Dual Bus	$2N$	$2N$	$2N$	$2C$	$2C$
DCCN [JT93]	Star	N	N	0	C	C

Table 4.1: Network Transceiver Requirements

channels on the network. The first three hardware categories of Table 4.1 are components of the user stations whereas the generator transmitters and receivers are system headend components. Currently, a major cost in a WDM network design is that of the tunable components. It can be seen that D-Net/C [SD93], Fairnet [BM93], and EQEB [CM92] require fast-tunable sense taps in addition to their tunable transmitter requirement. Such sense taps are of economic concern in that although they do not require the same bit receiving abilities of a true tunable receiver, they still must have the costly faculty of tunability. ACTA [Che93] which does not require a tunable sense tap does require additional tunable transmitters with respect to BTMSC and BTMSR, also making ACTA a more costly network. Lastly, a new generation of star networks are under development which also reduce the number of tunable devices in the user station design. DCCN is a good example of these networks. However, these star networks have difficulties in synchronization which needs to be addressed. Clearly BTMSC network designs have managed a reduction in total network cost over many alternative designs through the reduction of the tunable hardware and coupling requirements for a user station.

4.3 Media Access Control

The BTMSC and BTMSR media access protocols are the rules which stations must follow in order to successfully transmit and receive data in their respective network architectures. To do so, the concept of a HOME channel plays a central role [LT87], allowing stations to derive timing information and claim or reserve data slots. The

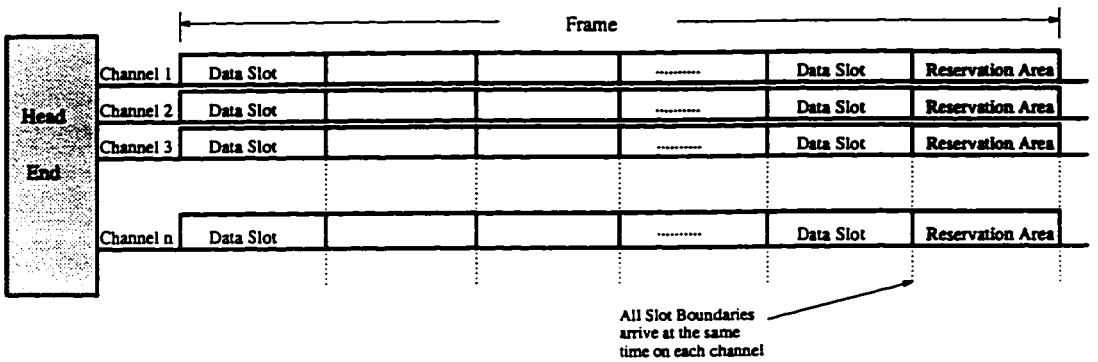


Figure 4.8: Slot Boundary Synchronization Across WDM Channels

resulting protocols form a packet oriented, contention free, broadcast and select type network.

In order to accommodate the lack of receiver tunability at a station, access to slots on a given channel is resolved on the station's HOME channel through a set of "mini-slots." A transmitting station then tunes its transmitter to the destination channel and transmits using slot-timing derived from the station's fixed-tuned receiver on its HOME channel.

4.3.1 Headend Operation

The headend has two main responsibilities as far as protocol operation is concerned. First, it must format a synchronized frame that is sent out across all channels. A number of data slots are contained in the frame. Each data slot on a channel is consecutively numbered and the numbering is consistent across all channels. Since each slot is released simultaneously at the headend, slot timing can be derived from any HOME channel and is valid across all channels as illustrated in Figure 4.8. Secondly, the headend must determine channel-to-channel bandwidth partitions and provide the necessary media access information to the user stations. This consists of determining for each channel, the number of data slots allocated for incoming traffic from each of the other channels. To facilitate understanding, a two level media access model is used. At the higher level, a movable-boundary asynchronous TDMA mechanism is used. Contiguous blocks of data slots on every channel are allocated for exclusive use

by stations on each channel. Many algorithms are possible for doing this and vary from being highly adaptive to static. The requirements of any such algorithms are briefly discussed in Section 4.3.2. Bandwidth partitioning may be implemented independently by each channel controller as in [JT93]. This information is communicated amongst all CCs over the low-bandwidth backplane of the headend as the partitions are changed.

Figure 4.9 shows an example of a data-slot partitioning. Each row in the figure represents a frame appearing on a given channel, which contains some amount of overhead and a number data slots. The data slots in the frame are sectioned into blocks as dictated by the bandwidth partitioning algorithm. Each block is dedicated to carrying packets transmitted by stations on a specified source channel. In the figure, four channels are shown. On channel 3 for example, the first block of slots has been assigned to stations on channel 1. This is referred to as B_{13}. Dynamic access to this block of slots occurs amongst those stations on HOME channel 1 that wish to transmit on channel 3. Using this arrangement, the use of slots in any block B_{ik} is determined by activity occurring solely on HOME channel i. This eliminates the need for a fast-tunable receiver since all transmissions are resolved locally.

4.3.2 Bandwidth Partitioning Algorithms

A bandwidth partitioning algorithm is that set of rules used in determining the number of slots assigned on each channel to other channels. For a constant uniform load, where the amount of traffic from each channel destined to each channel is equal, a fixed partition is sufficient. The same is true for skewed load scenarios as long as the traffic distribution remains constant or at least similar over time.

In this thesis, it will be assumed that the partitioning is such that blocks are allocated proportional to the steady-state traffic flows between channels. If γ_{ij} is the mean steady-state packet generation rate from stations on HOME channels i to j, then the fraction of data slots allocated to B_{ij} is given by

$$B_{ij} = \gamma_{ij} / \sum_{k=1}^{C} \gamma_{kj}, \tag{4.2}$$

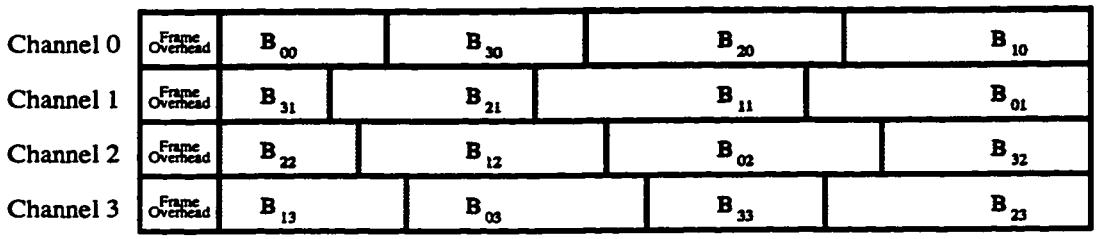


Figure 4.9: Inter-Channel Bandwidth Allocation

subject to integer roundoff. It should be noted, however, that the slot partitioning is a feature which is common to all of the proposed passive designs and is not in itself the focus of this thesis.

When traffic patterns are subject to many significant changes over some period of time, it becomes necessary to use some type of dynamic bandwidth partitioning algorithm in order to maximize capacity. Such an algorithm should attempt a best match schedule for a channel to channel traffic matrix. The traffic matrix is generated at the headend by monitoring the use of mini-slots at the end of the bus or ring. Matrix 4.3 illustrates just one possible traffic matrix for a three channel system. Each element of the traffic matrix represents a specific source to destination load where a row indicated the source channel and a column indicates a destination channel. In this example, source channel 0 requires 3 data slots for destination channel 0, 3 slots for destination channel 1, and 5 slots for destination channel 2.

$$\begin{bmatrix} 3 & 3 & 5 \\ 3 & 5 & 3 \\ 5 & 3 & 3 \end{bmatrix} \tag{4.3}$$

The channel to channel traffic matrix can be derived and updated from the utilization of the previous allocation in a manner similar that examined in [Che93]. A source to channel element is updated by the following expression:

$$element(src, dst) = \frac{element(src, dst) * Utilization}{controlledloadLc}, \tag{4.4}$$

where Utilization is the fraction of mini-slots in which reservation had been placed for a given connection. The controlled load parameter specifies the desired throughput under heavily overloaded condition.

A bandwidth partitioning algorithm then works upon this matrix to build a time-wavelength schedule of slots that the headend will release in the channel allocation table (Figure 4.12) of the next frame or use in granting slot allocations to individual stations. The partitioning algorithm must be sure to satisfy the constraint of the system. That is, a given channel may have only one channel sourcing data to it at a time.

The decomposition of the traffic matrix into a time-space schedule can be approached in terms of Bellman's principle of optimality [SHP91]. Bellman's principle says that if an optimum sequence of modes is decomposed into a subsequence taking the system from its initial state to an intermediate state, followed by another subsequence taking the system to the final state, the subsequence of modes to reach the intermediate state must be the optimum way to reach this state and the subsequence to go from there to the final state must be the optimum way to do this.

An ideal schedule breaks down any given traffic matrix into a minimum number of mode matrices and uses a minimum amount of time (slots). Many algorithms have been developed in the area of SS/TDMA [BCW81, Gan92, GBW83, Inu78, Inu79, IUMY77] which is similar to this problem. Luckily, the traffic matrix is an estimate of overall channel to channel utilization and thus we have some leeway in creating schedules for the channels. The only requirement is that only one channel may source a given destination channel.

Home Channel Reconfiguration

Finally, it may occur that during normal network operation traffic loading may become skewed. That is, much more traffic may be generated from or destined to a subset of channels. Such occasions may arise when there is a concentration of servers or busy stations on one or more channels. To optimize network performance it is desirable to evenly distribute the load over all channels. This is possible through a reconfiguration phase in which the headend instructs user stations to change home channels. Such a command can easily be placed in the start of frame or be a broadcast packet placed in a data slot. Since each station is only equipped with a slow tunable receiver this reconfiguration should be initiated only rarely. Once the commands to

reconfigure have been issued, normal network activity would have to be suspended until such time when the re-tuning at all user stations can be guaranteed complete.

4.3.3 Access Control Hierarchy

The lower level of media access is used to provide station access to individual data slots. Two separate protocols are used. The first protocol (BTMSC) requires a minimal amount of headend processing, but requires an extra FBT coupler in each user station. The second (BTMSR) requires fewer FBT couplers in each user station, but places more of the scheduling operations at the headend.

The two protocols will be applied to the folded bus and ring topologies. In each case a different hardware configuration is required as discussed in the previous section. In each protocol, access to slots is gained via the use of mini-slots. A mini-slot is similar to a normal data slot except that it contains far fewer bits and is thus much shorter in time. This discussion begins by differentiating the operations of the folded bus and ring versions of the protocols.

4.3.4 Folded Bus Operation

The folded-bus offers the simplest operation of the two topologies. Frames are aligned and transmitted onto the bus by the headend. In these networks, a station's Tx bus receiver is fixed-tuned to the HOME channel and is used to derive slot synchronization and the slot access information as discussed below. Reception of data slots occurs on a station's HOME channel using the receiver on the Rx bus. That receiver is also fixed-tuned to the HOME channel. Transmissions are achieved via the fast tunable transmitter located on the Tx bus. This transmitter is capable of accessing any of the wavelength channels on the system.

4.3.5 Ring Operation

On the ring networks, frames of slots are also generated by the headend and propagated around the ring. In contrast with the folded bus, each frame is split into

sub-cycles which are equal to the slot latency of the ring. It is assumed that in a LAN, frame lengths would normally be greater than the end-to-end propagation delay of the network so that multiple sub-cycles will be present in each frame. It should also be noted that the cycle and sub-cycle frame lengths are equal for each channel in order to maintain synchronization across the wavelength channels.

In the protocols, each of the sub-cycles is required to complete two rotations around a ring in order to ensure that all packets are properly passed to their destinations. This is a necessary result of the use of a single receiver at each station. In order for stations to transmit to upstream stations, all of their transmissions must somehow make their way via the downstream route past the headend. Since it is not known apriori which transmissions will be required to round the headend, it is assumed that potentially all transmissions will have to. We refer to the first rotation as the INITIAL rotation while the second is called the RESIDUAL rotation. It is not possible to interleave slots destined to an upstream station with new slots because, again, the required information is not known to the headend apriori. All INITIAL rotation slots must be converted to RESIDUAL rotation slots. It is also not reasonable to have sub-cycles shorter than the ring latency because of the interleaving of INITIAL and RESIDUAL data slots that would occur. Although this would not be a fatal condition, the interleaving would add much complexity to the headend which would have to track the slots in order to know when to release new ones.

In Figure 4.10 the ring sub-cycle mechanism is illustrated. In the first frame of the figure, the headend releases an INITIAL sub-cycle which is completed in the second frame. The third frame shows the conversion of the INITIAL slots into RESIDUAL slots as the INITIAL slots round the headend. Once those RESIDUAL slots return to the headend, they are removed and INITIAL slots for the next sub-cycle are released.

The fact that frames must do double rotations in the ring designs results in a reduction in capacity. To permit a frame shorter than the ring latency would require that the subsequent frame be postponed until the current one has completed both INITIAL and RESIDUAL rotations, wasting additional bandwidth. However, the ensuing designs require one less receiver at each station. Just as in the folded bus, slot synchronization is obtained from the fixed-tuned receiver. Packet reception

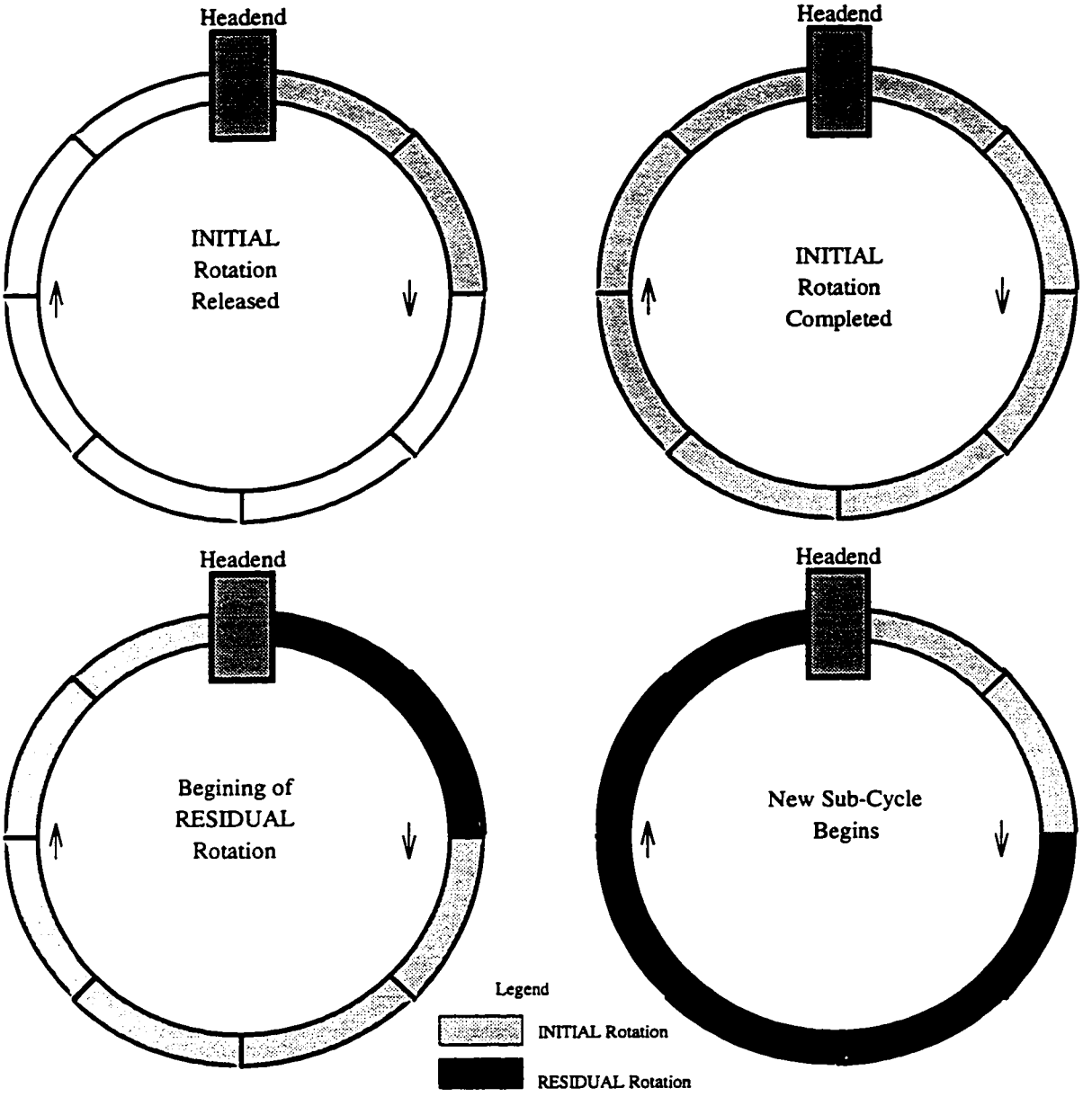


Figure 4.10: Ring Sub-Cycle Mechanism

also occurs via this attachment. Station transmission occurs via the fast tunable transmitter.

4.3.6 Mini-slot Contention Access

The BTMSC protocol will now be described for the ring topology. Since the ring protocol is more complex, it is a simple matter to apply it to the folded bus. Basically, all references to INITIAL and RESIDUAL rotations can be ignored for the folded bus. Once a slot has returned to the headend, it has completed its job and is removed from the system.

As described above, channel-to-channel data blocks are allocated by the headend. Stations contend for these slots on their HOME channel using contention mini-slots. To eliminate the need for a tunable receiver, there must be one mini-slot on each HOME channel for each data slot available to those stations on each HOME channel. For example, all stations with HOME channel x will observe contention mini-slots corresponding to the data slots available only to that group of stations. This allows the HOME channel receiver(s) to determine which slots are available on the various channels, and allows the tunable transmitter to reserve those slots while tuned to the HOME channel. Thus once per frame, the tunable transmitter is required to tune to the station's HOME channel in order to reserve data slots, regardless of the channel those slots will be on. Having then reserved the required data slot(s), the transmitter is tuned to the channel for which it has reserved its first slot.

In this protocol, the contention mini-slots appear at the beginning of each cycle and do not normally make a RESIDUAL rotation of the ring. Under heavy load conditions, it would not be necessary for the contention mini-slots to ever make multiple rotations of the ring since all data slots would have already been successfully contended for, for that frame. However, during low loads it may prove useful in reducing low load delay to have the contention mini-slots rotate the ring with each sub-cycle. In this way, data slots for the frame could be dynamically reserved not only at the beginning of each frame, but during each sub-cycle of the frame. However, the resulting system would suffer from somewhat reduced capacity because of the

increased overhead of transmitting the contention mini-slots during each sub-cycle. Clearly, this option is not available for the folded bus version of the protocol. Because of the added complexity of such a system, the remainder of this discussion will assume that the contention mini-slots only rotate the ring once.

Each data slot released by the headend is marked as an INITIAL rotation. This is done by means of the ROTATION field contained in each data slot. As slots marked INITIAL round the headend they are remarked as RESIDUAL and placed back onto the ring. The frame format is shown in Figure 4.11. Each frame begins with a START OF FRAME section containing some amount of guard-band in order to provide synchronization and account for timing jitter. Also contained within the START OF FRAME are the quota resets used for providing fairness as discussed in Section 4.3.6. The next portion of the frame is the CHANNEL ALLOCATION TABLE which is used for the reservation of data slots. Finally, are a number of data slots. Each data slot consists of a SLOT NUMBER, ROTATION, DESTINATION and PAYLOAD fields. The SLOT NUMBER field is required to identify the slot with respect to the CHANNEL ALLOCATION section, and the ROTATION field identifies the slot as being on its INITIAL or RESIDUAL rotation as mentioned earlier. The DESTINATION field consists of the address to which the data slot is intended. This field may also contain the address of the messages creator. Finally, the PAYLOAD field contains the message for transmission. Each PAYLOAD field is fixed in length and the same for every data slot. It is envisioned that the PAYLOAD field will be on the order of 1000 bits long. Given a transmission rate of 1 Gbps, each data slot will take 1 μs to transmit. Each mini-slot will take significantly less time to transmit. It may be possible to make mini-slots as short as 1 bit a piece.

Channel Allocation Table

Since there is a one-to-one correspondence between mini-slots and data slots, mini-slots are used by stations on each HOME channel to reserve data slots for use during that cycle. The number of mini-slots is a function of the number of data slots in the current frame and the number of channels in the network. Since each cycle length is the same for all channels, the headend may have to release contention slots that will

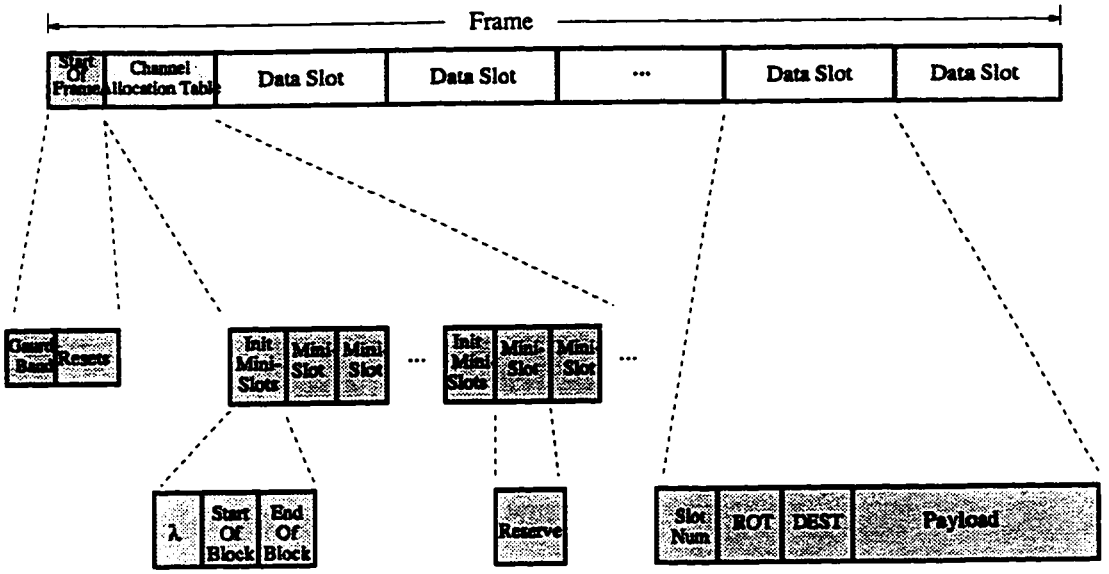


Figure 4.11: BTMSC Frame Structure

never be used. This type of contention slot is referred to as a NULL allocation. The headend uses the bandwidth allocations it has previously determined as "optimal" to set the number of slots available to stations on a given HOME channel.

Stations are synchronized to the mini-slots following an initialization. Each station must read information contained in the initialization and determine which mini-slots correspond to the block of slots that the station wishes to access. An example is shown in Figure 4.12. The initializations are shown as three digits, where the first refers to the destination channel, and the other two show the data slot number range of the block. In the Figure, stations on HOME channel 0 will see that they have access to channel 0 for 5 slots, channel 3 for 3 slots, channel 2 for 12 slots, and so on.

Note that the headend does not release the same allocation table on every channel. While there are the same number of mini-slots, their allocations vary and may contain NULL allocations. As in Figure 4.12, the other channels of the system will see similar bandwidth allocation tables with the exception that the destination channels and block ranges will be rearranged.

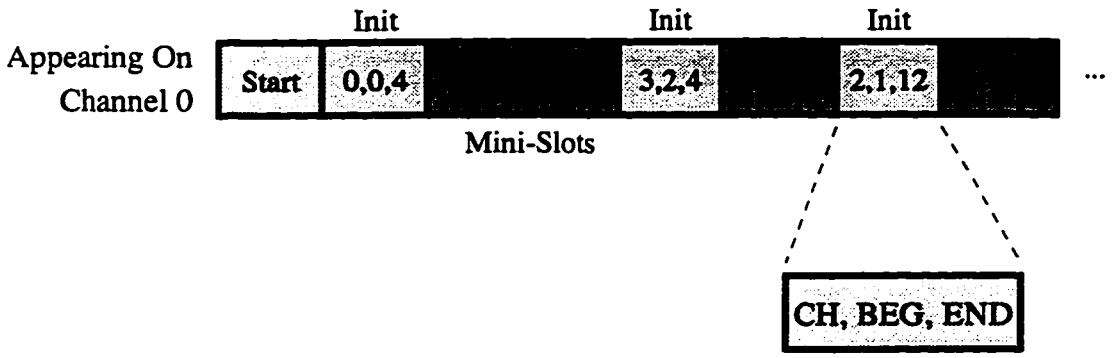


Figure 4.12: Channel Allocation Table

Transmission and Reception

Stations wishing to transmit may do so only when they have successfully contended for slots. For each data slot desired on a particular channel, a previously empty mini-slot must be successfully written into. This action reserves the associated data slot and prevents stations downstream from using it. This is possible as a direct modulation scheme is assumed for the mini-slots. Thus a free mini-slot consists of an absence of optical energy during that mini-slot, whereas a reserved mini-slot contains optical energy. A station wishing to reserve a data slot will write optical energy into the mini-slot while simultaneously determining if the mini-slot already had optical energy within it. Only when the mini-slot has been determined to have been previously empty will a successful reservation be acknowledged on behalf of the station. Having reserved a data slot, or slots, the station will then tune its transmitter to the appropriate channel and await its reserved data slot(s).

Station reception is very simple. A station only receives data on its HOME channel which it continually monitors for packets. When a properly addressed packet arrives at the station, it is copied into the station's local buffers and the slot is allowed to continue propagating on to the next station. All other slots are permitted to pass by the station without any protocol specific action being performed.

Fairness Mechanism

In order to avoid unfair reservations by stations upstream of others, a fairness mechanism is required. This is because the base mini-slot contention mechanism is unfair, favoring stations closest to the transmitting side of the headend. As mini-slots are released, the station closest to the headend will encounter no resistance in trying to reserve a data-slot. Subsequent stations, however, will see the reservations made by upstream stations. As a result, downstream stations may starve for data slots, while the upstream stations satiate themselves.

In this thesis a technique first proposed in [FA85] is modified for use with multiple channels. The mechanism is a cycle based approach where each station is given a transmission quota. In the BTMSC scheme, each station is given a separate transmission quota, N_q, for every channel in the system. Each time a station transmits, the transmission quota for that channel is decremented in the station. If the count reaches zero for a given destination channel, the station must suspend transmission to that channel until the quota is reset. The length of a fairness cycle is then dependent on the total number of stations, the transmission quotas granted to the stations, and the total load placed on the network.

Resets are generated by the CCs and are placed within the start of frame on a station's HOME channel. Resetting is independent for each block and is initiated whenever the headend sees an un-reserved mini-slot in the block. Using this scheme, individual stations can be given higher priority access to the network simply by increasing the stations' quota with respect to other stations. At low loads, then, resets will be generated frequently and the fairness cycle will be short. At increasingly higher loads the frequency of resets will decrease resulting in longer fairness cycles. The reader is referred to [FA85] for a more detailed discussion of the choice of N_q. In general, the choice of N_q involves a balance between efficiency and the duration of time over which fair operation is achieved. This thesis does not contribute to the subject of fairness any further.

4.3.7 Mini-slot Reservation Access

In the previous section it was shown that the station design for the BTMSR protocols lacked the ability to transmit while simultaneously receiving transmissions from upstream. This was a result of the reduced number of passive couplers at each station. As a result, a protocol other than BTMSC is required as the mini-slot contention mechanism requires that stations be able to manipulate bits in a mini-slot while reading the mini-slot's previous state.

The BTMSR protocol uses a centralized technique similar to DCCN [JT93] which passes the task of data slot allocation to the channel controllers. In this case, stations are given a dedicated request mini-slot on their HOME channel. A station makes a request for data slots by transmitting into its mini-slot. All requests are then received by the station's HOME channel CC, which generates slot allocations based upon the channel-to-channel blocks allocated by the headend as discussed in Section 4.3.2. These are then transmitted back to the stations in a set of allocation mini-slots and appear on a station's HOME channel. It will be assumed that mini-slots are of the same length regardless of them being request, allocation, or contention mini-slots.

As with the BTMSC protocols, the BTMSR protocols may be applied to either a ring or folded bus topology. Again, data slots on a ring must perform both an INITIAL and RESIDUAL rotation in order to ensure that all packets are passed to their destinations. In this case, however, the BTMSR ring does not have the option of allowing reservations to be made in each sub-cycle. To do so would be overly complex as the headend would then be required to generate the slot allocations much more frequently.

The frame structure is shown in Figure 4.13. The request mini-slots contain two fields. The λ field indicates the destination channel that a station wishes to communicate on, and the NUM field indicates the number of data slots requested. Each allocation mini-slot is dedicated to each station as done with the request mini-slots. An allocation mini-slot consists of three fields. The λ field indicates the channel a station may transmit over, the START field indicates which data slot to begin transmission into, and the NUM field indicates how many contiguous data slots may be transmitted.

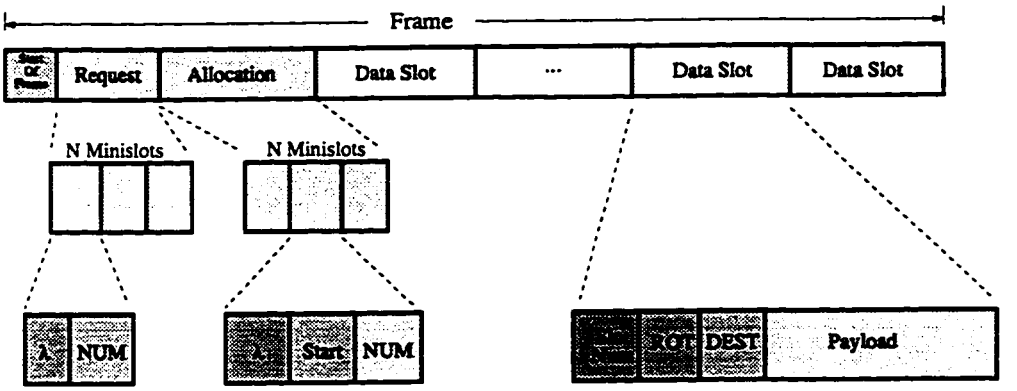


Figure 4.13: BTMSR Frame Structure

Transmission and Reception

To transmit data on a channel a station must tune its transmitter to its HOME channel once per frame and transmit in its request mini-slot. The station then examines its allocation queue for a received allocation from the previous frame. If such an entry exists, it tunes its transmitter to the channel and transmits the specified number of slots.

Reception is identical to the BTMSC case. A station only receives data on its HOME channel which it continually monitors for packets. When a properly addressed packet arrives at the station, it is copied into the station's local buffers and the slot is allowed to continue propagating on to the next station. All other slots are remitted to pass by the station without any protocol specific action being performed.

Headend Request Scheduler

The details of how allocations are generated in response to requests are a function of the headend request scheduler located in each channel controller. Each CC must be able to continuously accept requests and generate appropriate slot allocations for all stations on its HOME channel once per frame. Note that because of propagation delays, there would normally be a pipelining delay of one frame between requests and allocations. Of concern is the ability of the headend to generate the required

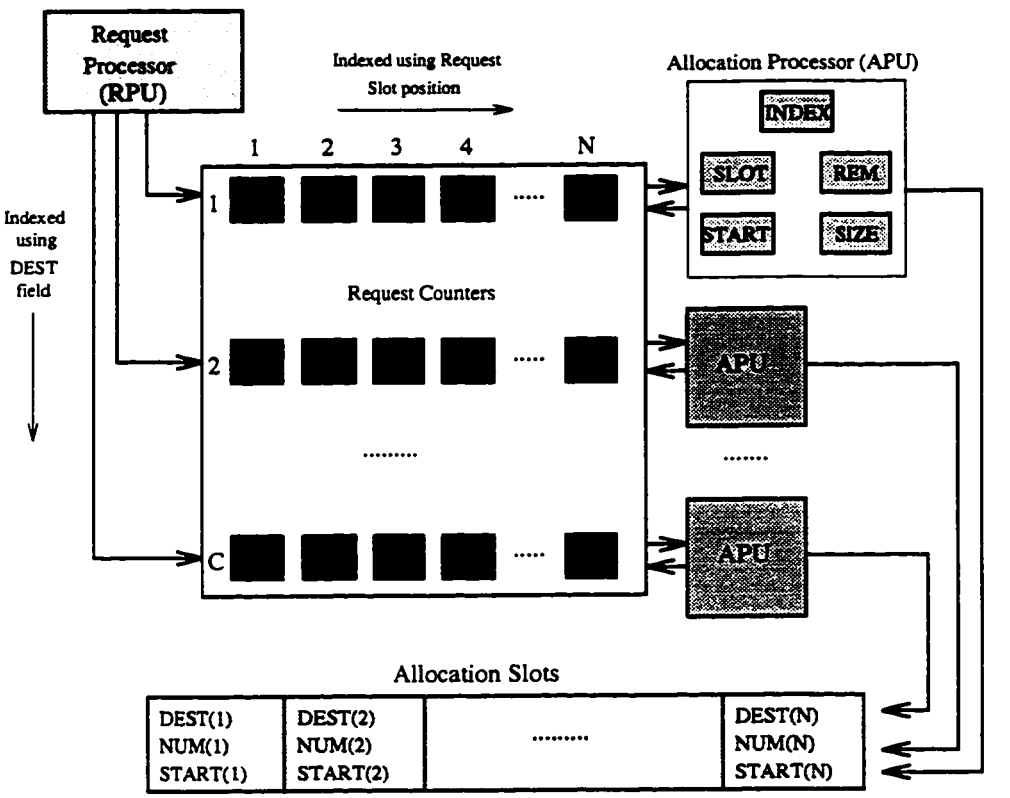


Figure 4.14: Headend Processing

slot allocations within this frame period. Because of the data-slot block partitioning, each channel controller can independently generate allocations for the stations attached to its HOME channel. This feature greatly simplifies the task of allocation generation, which may be easily done using a custom hardware scheduler (one per channel controller).

An example of one possible design is shown in Figure 4.14 [JT94]. The scheduler consists of $C \times N$ request counters, a request processing unit (RPU) and C allocation processing units (APUs). The RPU is responsible for receiving and processing incoming request mini-slots. Once a request has been received, the RPU must select the request counter specified by the λ field of the request mini-slot and index the request according to the position of the request slot within the request sub-frame. The counter is then incremented using the NUM field of the request mini-slot. Once per frame the headend generates an allocation sub-frame based upon the data in the request counters using the set of APUs. An APU is dedicated to each channel in the network and is used to fill in the fields of the allocation mini-slots.

The operation of an APU can be explained by considering it as a set of five registers and counters. The INDEX register indicates which request counter and allocation mini-slot to access. The START and SIZE registers indicate the start position and size respectively of the data slot block on the channel for which the APU will allocate slots. These two registers are set by the block partitioning algorithm. The REM counter indicates how many data slots remain to be allocated within the block and the SLOT counter is used to hold the position of the next free data slot within the data slot block. REM and SLOT are loaded with the contents of SIZE and START respectively at the beginning of the allocation process for each frame. All APUs operate in parallel using the following sequence of operations [JT93]. The variable *val* is a temporary variable required in the allocation process.

1. Access the request counter specified by INDEX register.
2. $val = min(REM, request\ counter)$
3. $if(val \neq 0) then$
   $NUM(INDEX) = NUM(INDEX) + val$

$START(INDEX) = SLOT$
$DEST(INDEX) = channel$
$request\ counter(INDEX) = request\ counter(INDEX) - val$
$REM = REM - val$
$SLOT = SLOT + val$

4. $INDEX = INDEX + 1$

5. Repeat as long as $REM \geq 0$.

Step 2 determines the number of slots that can be allocated. If slots are available then step 3 is used to update the allocation mini-slot of the appropriate station and the counters and registers of the APU. Using this type of design, each CC can easily generate allocations when channel bit rates are well above 1 Gbps using off-the-shelf digital hardware. Finally, no separate fairness mechanism is required since station position is no longer a factor in the request procedure or in the allocation generation algorithm.

4.4 Performance Analysis

In order to gauge the merits of differing architectures and protocols, designers often justify their networks in terms of capacity and delay performance. Many analytical techniques have been developed to do this, among which are mathematical capacity and mean delay analysis [Hay84]. To augment the simplifying assumptions made for most analytical models, computer simulation models based on discrete-event programming are also used [Abu88]. These simulations are useful for the validation of the mathematical model and for testing the design under more complex operation conditions. Such conditions may include the use of real traffic traces to drive the simulations.

In many analytical capacity models, it is assumed that an infinite population of contending stations exists on the network, each with data to transmit to a destination. Equivalent to this assumption is that of a finite number of contending stations, there will always be at least one station with an infinite backlog of packets to transmit. In

reality such conditions rarely exist. However, such analysis is useful when comparing differing designs.

Analytical delay models attempt to measure the mean time difference between when a packet is generated at a station and the time that the packet is received at its destination. Such calculations are almost always much more complicated than capacity calculations for the same system. Clearly, many statistics are involved in the determination of delay. For example, a commonly used statistic is the characterization of the packet generation process at a station. A mathematically convenient and accepted method of characterizing this arrival process is as a Poisson process. Such an assumption often limits the applicability of the analytical models to the real world. As such, simulation plays a big role in fine-tuning system parameters to a specific application. Before becoming too pessimistic of the value of the Poisson arrival process assumption in system modeling, it should be noted that its use is very common and it does provide a solid basis for comparing performance between systems.

The discussion of the systems performance now begins with an exploration of capacity, defined as the fraction of bandwidth that is available for the transmission of user station data.

4.4.1 Capacity

Maximum throughput or capacity can be interpreted as the total peak user data rates as a fraction of the total. For frame based protocols such as BTMSC and BTMSR, the analysis of capacity is quite simple. Any protocol can be considered as frame based if at some predictable time a regenerative event occurs. Such an event can be viewed as marking the end of the previous frame and the start of a new frame. In view of the frame structures discussed earlier in the chapter, the maximum capacity can be easily calculated using

$$C_{frame} = \frac{\text{Used Data Slot Time Per Frame}}{\text{Total Frame Time}}. \tag{4.5}$$

The following variables are used in the capacity relations.

L = Time to Transmit a data slot
M = Number of data slots per frame
C = Number of channels
N = Number of user stations per channel

The discussion of analytical capacity models begins with the capacity formulation for the passive systems.

Analytical Models

For the passive systems, capacity derivations are straight forward. Each protocol contains markers which identify the start and end of frames. It is clear that within each frame there is a certain amount of operating overhead which degrades the maximum achievable capacity performance. Putting all this together under Equation 4.5, the derivation becomes simple. To facilitate writing of capacity expressions, the following assumptions are made for the BTMSC and BTMSR systems.

1. There is at least one backlogged station with packets destined for each channel, thus ensuring that each block of data slots will be fully utilized on every channel.

2. The network is operating under steady state conditions. Interchannel blocks have been allocated proportional to interchannel traffic flow.

3. The effects of preamble and guard time are taken into account by the size of the mini-slots and data slots.

4. Round-trip propagation delay is assumed to be less than frame transmission time. This is typically a valid assumption for LANs and will be used for the analysis of both protocols.

In light of the frame based capacity formulation given in Equation 4.5, the per channel capacity of the BTMSC protocol can be written as,

$$C_{BTMSC} = \frac{\text{Used Data Slot Time Per Frame}}{\text{Info Time} + \text{Minislot Time} + \text{Init Time} + \text{StartOfFrame Time}} \quad (4.6)$$

Similarly, the capacity of the BTMSR protocol becomes,

$$C_{BTMSR} = \frac{\text{Used Data Slot Time Per Frame}}{\text{Info Time} + \text{Reservation Time} + \text{Allocation Time}}. \quad (4.7)$$

For the BTMSR and BTMSC protocols, the following additional parameters will be required for a complete formulation.

$$\begin{aligned} T_{ms} &= \text{Time to transmit a contention mini-slot} \\ T_{conf} &= \text{Time to transmit an allocation mini-slot} \\ T_{req} &= \text{Time to transmit a request mini-slot} \\ T_{mi} &= \text{Time to transmit a mini-slot initialization} \\ T_{st} &= \text{Time to transmit a start of frame} \end{aligned}$$

Furthermore, it will be assumed that traffic is uniformly distributed causing bandwidth to be partitioned into C equal blocks on a destination channel (1 block for each source channel). For the BTMSR protocol, a mini-slot request and a mini-slot confirmation is released for every user station on a particular channel. Since the data slots on each channel will be fully utilized, the capacity calculations for BTMSR simply become

$$C_{BTMSR-ring} = \frac{LM}{2LM + N \cdot T_{conf} + N \cdot T_{req} + T_{st}}, \quad (4.8)$$

$$C_{BTMSR-bus} = \frac{LM}{LM + N \cdot T_{conf} + N \cdot T_{req} + T_{st}}. \quad (4.9)$$

For the BTMSC systems, the number of contention mini-slots released on a channel is determined by the particular block partitions for that frame. For convenience, the number of mini-slots in a frame is defined by the terms S_{ms} and S_{mi}, which represent the number of contention mini-slots and mini-slot initializations that are produced per frame. Since the data slots on each channel in a frame will by fully utilized, the per channel capacity calculations for BTMSC simply become

$$C_{BTMSC-ring} = \frac{LM}{2LM + S_{ms} \cdot T_{ms} + S_{mi} \cdot T_{mi} + T_{st}}, \quad (4.10)$$

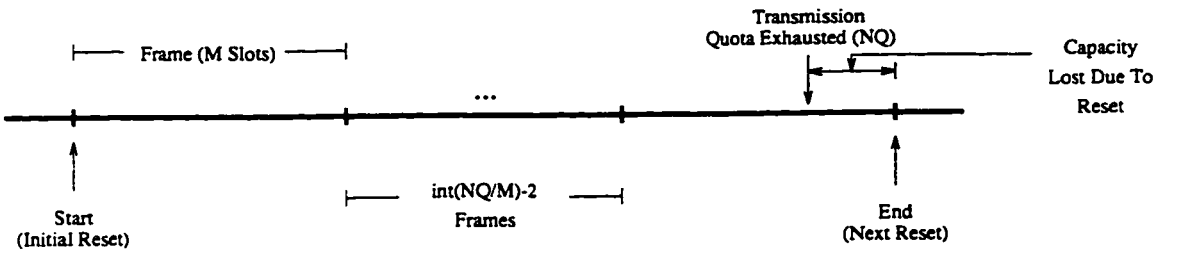


Figure 4.15: Effect Of Quota Reset Mechanism On Capacity

for the BTMSC ring system and

$$C_{BTMSC-bus} = \frac{LM}{LM + S_{ms} \cdot T_{ms} + S_{mi} \cdot T_{mi} + T_{st}}, \tag{4.11}$$

for the BTMSC folded bus.

Now consider the effects of finite transmission quotas on the BTMSC systems. In view of the steady-state block partitions, the utilization of each block on any channel may be easily found. Define the quantity N_Q to be the sum of all contending station's assigned N_q for the block in question. If the number of data slots per frame for the block is B, then it can easily be seen that the utilization of the block is given by

$$U_B = \frac{N_Q}{\lceil \frac{N_Q+1}{B} \rceil \cdot B}. \tag{4.12}$$

An example is shown in Figure 4.15 where the first $int(N_Q/M) - 1$ frames are fully utilized. The numerator of Equation 4.12 represents the number of packets occupying data slots between consecutive resets. The denominator gives the total number of data slots (used and unused) between consecutive resets. As the value of N_Q runs from 0 to ∞, the capacity experiences local maxima where $(N_Q + 1)$ is an integer multiple of B as depicted in Figure 4.16. Using Equation 4.12, it is very straightforward to calculate the total capacity of a channel with different size blocks and transmission quotas.

Under the assumption that all blocks on a given channel are equal, the capacity per channel of the ring and folded bus is given by

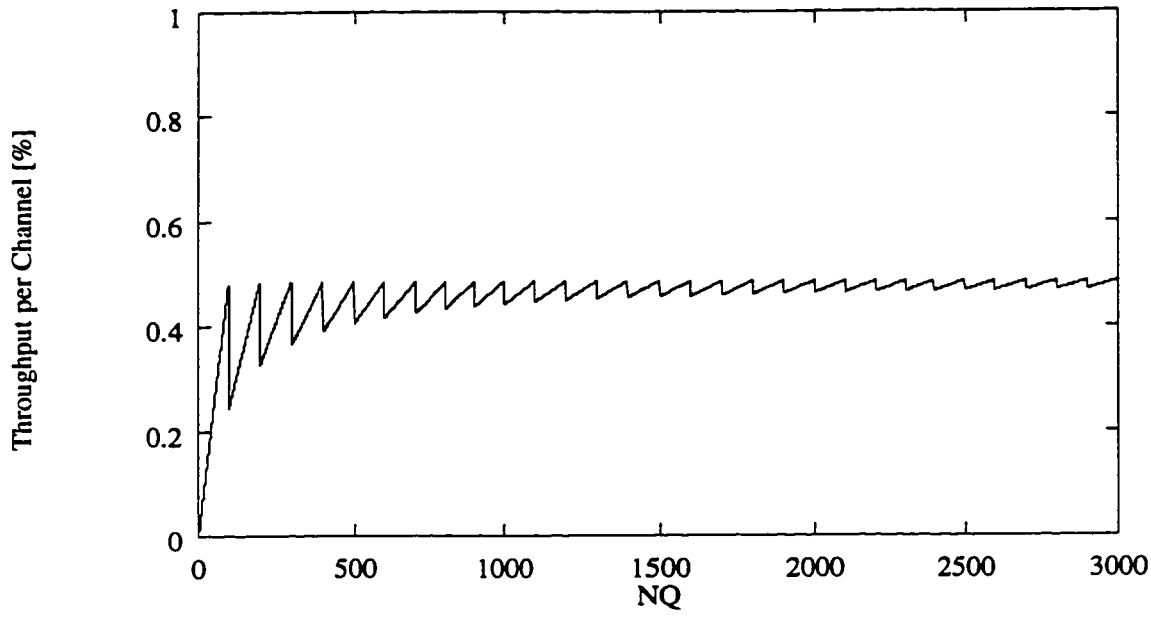


Figure 4.16: BTMSC Capacity versus N_Q for the Ring Topology, L=20, B=100, C=10

$$C_{BTMSC-ring} = \frac{LM}{(2LM + S_{ms}T_{ms} + S_{mi}T_{mi} + T_{st})} \cdot \frac{N_Q}{\lceil \frac{N_Q+1}{M/C} \rceil M/C}, \tag{4.13}$$

and

$$C_{BTMSC-bus} = \frac{LM}{(LM + S_{ms}T_{ms} + S_{mi}T_{mi} + T_{st})} \cdot \frac{N_Q}{\lceil \frac{N_Q+1}{M/C} \rceil M/C}. \tag{4.14}$$

To complete the BTMSC capacity model we must consider the effect of propagation delay (τ measured in mini-slots) for the folded bus. In the ring system, it is assumed that the round trip propagation delay is less than or equal to the frame transmission time. For the folded bus topology, Figure 4.17 illustrates that a large τ will simply delay the generation of a reset by a number of frames. This occurs since the headend must detect the exhaustion of the transmission quota before a reset can be sent out.

From Figure 4.17 we see we must wait $\lceil \frac{\tau}{NL} - 1 \rceil$ frames for the next reset following the frame in which the transmission quota has been exhausted. As before, the capacity can be easily written as

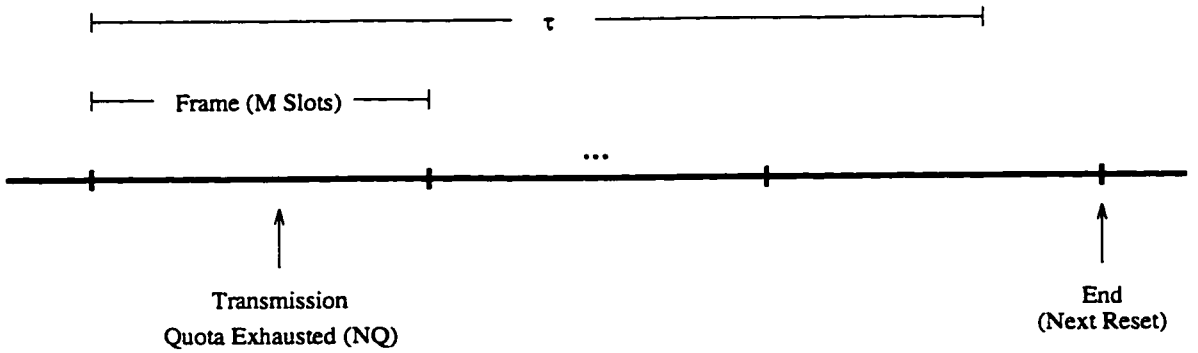


Figure 4.17: Normalized Propagation Delay's Effect On Capacity with Passive BTMSC Systems.

$$C_{BTMSC-bus} = \frac{LM}{LM + S_{ms}T_{ms} + S_{mi}T_{mi} + T_{st}} \cdot \frac{N_Q}{(\lceil \frac{N_Q+1}{M} \rceil + \lceil \frac{\tau}{ML} - 1 \rceil)M}. \quad (4.15)$$

Under the assumption that traffic is uniformly distributed causing bandwidth to be partitioned into C equal blocks on a destination channel (1 block for each source channel) it is obvious that $S_{mi} = M$ and $S_{mi} = C$. This requires that on each channel a mini-slot is released for each data slot in a frame and the number of mini-slot initializations required is equal to the number of blocks in the frame.

Results

The above equations are an exact analysis of the BTMSC and BTMSR systems capacity. This arises from the deterministic frame structure of the passive system protocols. The results of Figure 4.18 plot the above derived equations and clearly illustrate the lost performance when moving from the folded bus topology to the ring. Also of interest is the sensitivity of throughput performance to the number of channels in the system for the BTMSC network. In the modeled systems it was assumed that τ is less than the framelength, L=20, M=100, and N_Q=150. Larger values of N_Q would increase throughput, but exceedingly large values may not be appropriate since any sense of short term fairness would be lost. Furthermore, T_{ms}, T_{conf}, T_{req}, T_{mi}, and T_{st} are all normalized to unity. As a result of Figure 4.18, the BTMSC protocol should be limited to a system with only a handful of wavelength

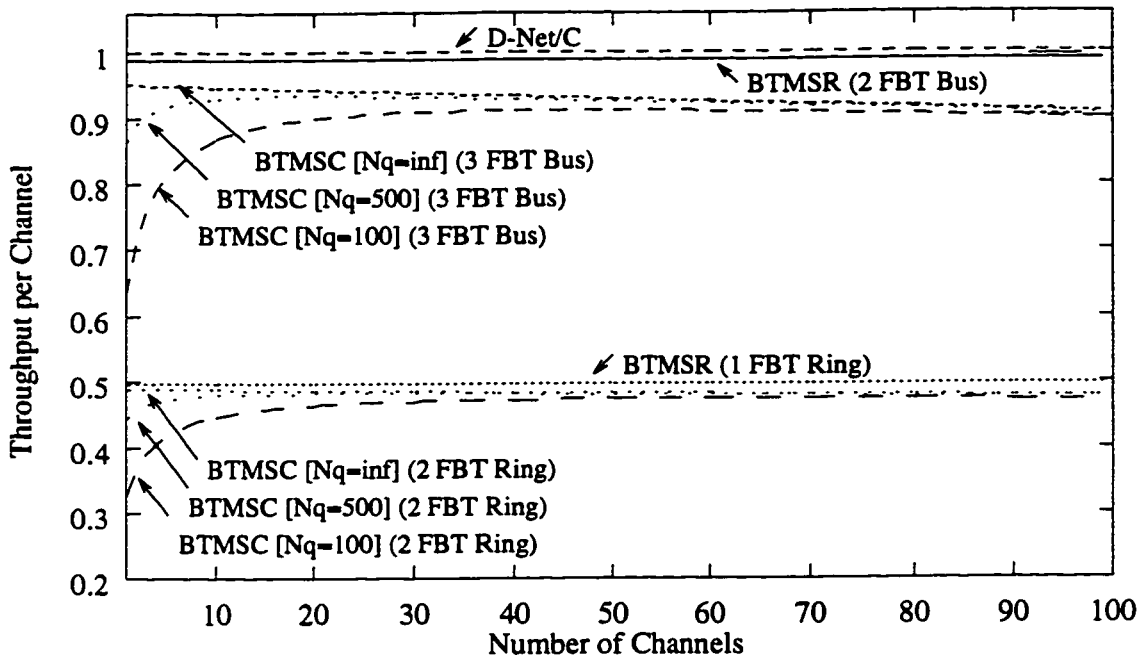


Figure 4.18: Capacity as a Function of Number of Channels for Passive Systems

channels. On the other hand, from Figure 4.19, it is seen that the BTMSR protocol is sensitive to the number of user stations per channel. As such, the BTMSR protocol is limited to a small number of user stations per wavelength channel. In the near future it is expected that designers will have only a small handful of wavelength channels to work with. Correspondingly, the number of user stations per user channel will be relatively large. In such cases, it can be argued that the BTMSC protocol (using more user station hardware) is a more suitable interim candidate.

It is interesting to note that much improved performance is possible if wavelength agility is also included in the Tx channel receiver. Using the same parameters as in Figures 4.18 and 4.19, a capacity very close to unity is possible using the D-net/C scheme discussed in [SD93]. By adding the additional hardware requirement, contention for remote slots can be performed directly on the remote channel. Finally, Figure 4.20 depicts the deteriation of capacity performance for the BTMSC bus system. In this case, a BTMSC bus system is sensitive to network size for a fixed frame length. The BTMSC ring is not immune to the effect, but due to protocol operation, it was assumed earlier that the frame length would always be larger than the network

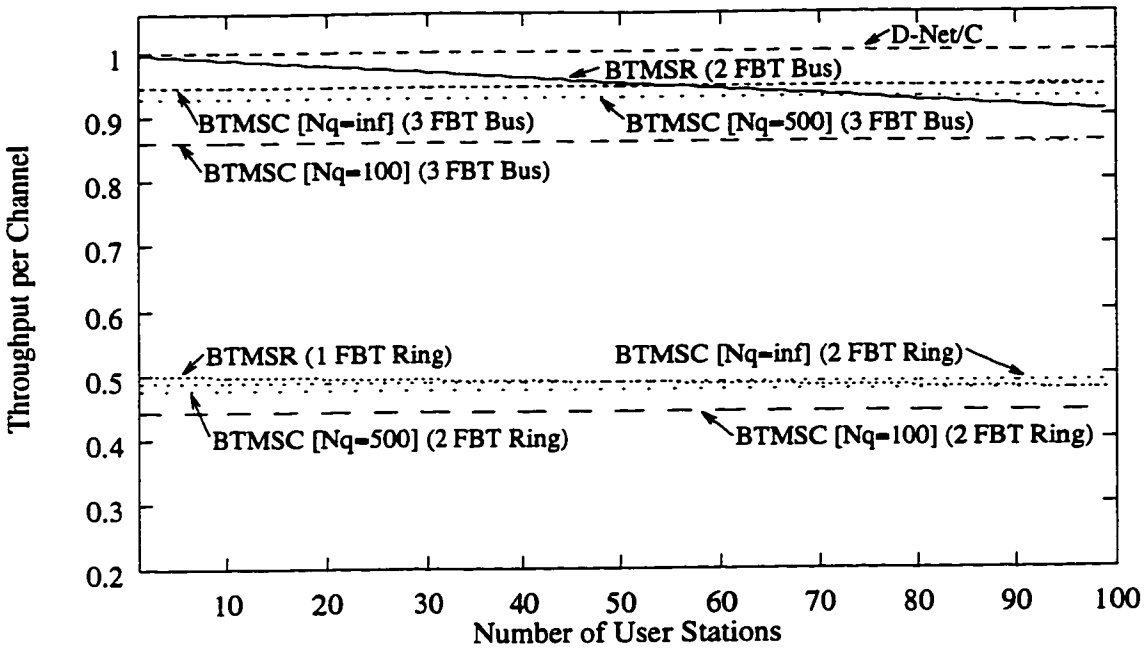


Figure 4.19: Capacity as a function of Number of User Stations per Channel for Passive Systems

size. This is reflected in the constant capacity of the passive ring systems. It is also shown that the BTMSR systems have no significant capacity relation to the network size. For interest, two active coupled networks discussed in the next chapter are also plotted on the same graph. Both simulation and analytical results are depicted. It can be seen that in general passive systems have less dependency on network size than do active systems. The results obtained here thus give an indication of the complexity and performance tradeoffs associated with reducing the hardware capabilities of the stations.

4.4.2 Delay

In this section we consider the mean delay experienced by packets generated at the stations. This is an average delay experienced by a packet between its generation at a source station and its reception at a destination. Both simulation and analytic models are given. The objective is to obtain simple closed-form analytic results which can be used to validate the simulation results. Stations are assumed to be uniformly tuned

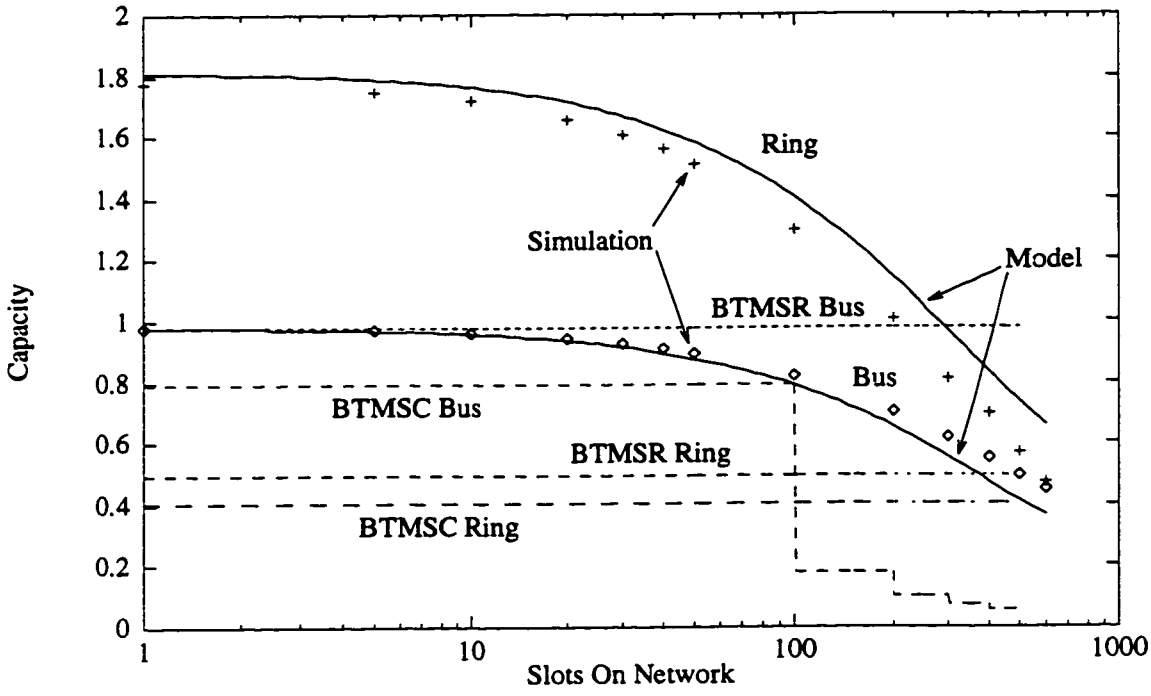


Figure 4.20: Capacity vs. Slots on Network for Passive Systems

to HOME channels across all C wavelengths. The number of stations per channel is again defined to be N. Consider a particular queue at some average station i which is generating arrivals destined only to a particular channel. We consider the steady-state mean delay experienced by these packets. The buffer is assumed to be infinite and is fed by packet arrivals from a Poisson process at rate $\gamma_i = \gamma$. Packets are fixed in length and occupy one data slot each. Packet destinations are random and chosen independently and uniformly across those stations on the destination channel.

Gated Model

This model is intended to predict the delay of the passive system protocols: BTMSC and BTMSR. These systems can be considered to provide "gated" access to slots in a frame. That is, a station can only acquire transmission rights to a slot at the beginning of a frame. A slot arriving at a station must always wait until the next frame before even attempt is made to transmit. In this model it is assumed that a

subset of stations on channel i are generating traffic to stations on channel k at a combined Poisson rate of λ_{ik} slots per frame. The number of slots available for use by the transmitting stations is B_{ik}, which remains fixed.

The analysis embeds a Markov chain at the end of each frame, for which a state equation regarding the number of customers held in a global queue can be found. The state equation is then written as a probability generating function which can be solved with the aid of Rouché's theorem and inverting the probability generating function. The delay experienced by a customer is then found in terms of the mean number of frames required to service the customer. Finally, the mean delay experienced by a customer can be written in terms of mini-slots. This was derived in Appendix A and for BTMSC, expressed as

$$\overline{D}_{BTMSC} = \frac{T_f}{2} + T_f(\overline{m} - 1) + \frac{T_f}{4}, \tag{4.16}$$

where $\overline{m}$ is the mean queuing delay, expressed in frames, experienced by an arrived packet, and T_f is the length of a frame in mini-slots. Similarly, for BTMSR, mean delay can be expressed as

$$\overline{D}_{BTMSR} = \frac{T_f}{2} + \lceil \frac{\tau}{T_f} \rceil \cdot T_f + T_f \cdot (\overline{m} - 1) + \frac{T_f}{4}. \tag{4.17}$$

Using Equations 4.16 and 4.17 it is possible to predict the mean queuing delay experienced by a packet destined to a specific block in a frame. To do so requires that we know T_f, the number of data slots in the block, and the aggregate arrival rate of packets destined to that block. Each of these parameters are easily known.

Simulation Environment

The simulations of packet delay are based on stochastic, self-driven, discrete-event models. The simulators were also written in the C programming language with the aid of library routines that provide discrete-event and random variable facilities as described in Appendix C. Each simulation program is capable of modeling various parameters of network operation. For comparison and validation with the analytical

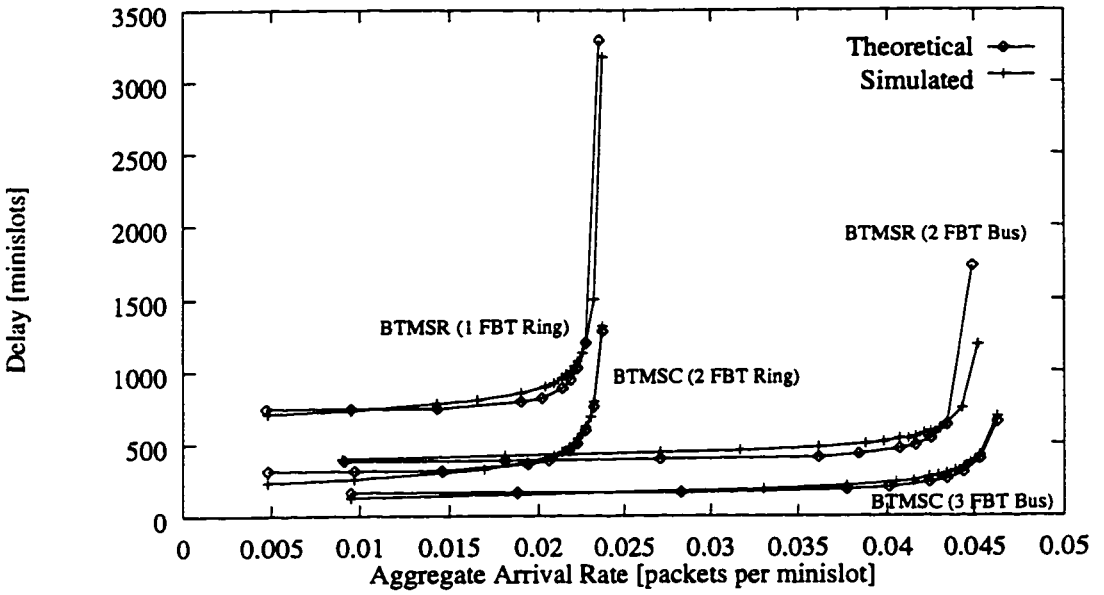


Figure 4.21: Mean Network Queuing Delay for Passive Ring and Folded Bus Systems.

delay models, the simulation was run such that station traffic was uniformly distributed and packet arrivals were generated by a Poisson process. A fixed schedule was used for the bandwidth partitioning algorithm and an end-to-end network length of five data slots was assumed.

Results

Figure 4.21 gives a typical example of the queueing delay characteristics of the BTMSC and BTMSR protocols under steady state, uniform load conditions. Again it was assumed that $C = 1, L = 20, M = 10, N = 10$, $N_Q = \infty$ and $T_{conf} = T_{req} = T_{st} = T_{mi} = T_{ms} = 1$. The analytical and simulation methods produce results which are in close correspondence with each other. The major difference between the two approaches occurs during low traffic loads. This is due to the assumptions made in the analytical model regarding packet departure time. Notice that for each protocol, the folded bus implementation outperforms the corresponding ring. Not only will the folded bus accommodate higher arrival rates, but at low loads the delays experienced by packets in a folded bus network are nearly half of what is experienced in the ring.

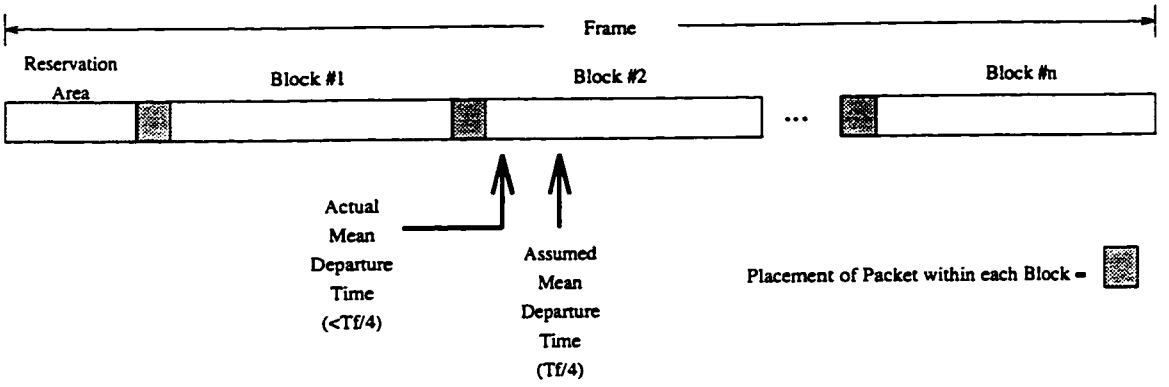


Figure 4.22: Distribution of packets for low load conditions

It was assumed that departures in the analytical model occur at a time equal to $T_f/4$ into the frame. However, for low load, steady state conditions, this assumption is inaccurate. At low traffic loads, it is true that packets are uniformly distributed across transmission blocks. However, the packets are not uniformly distributed within those blocks. Figure 4.22 illustrates this principle. In this case, packets are placed at the front of each block (whereas the analytical model assumes that the packets are uniformly distributed). This shifts the actual mean departure time down, causing the analytical model to generating slightly higher delays than what can be expected. A similar effect occurs at high loads where mean packets departures actually occur in the middle of any given frame. Since the analytical model assumes that packet departures occur at $T_f/4$ the analytical model will slightly underestimate the mean delay.

In Figure 4.21 it can also be seen that the capacity is slightly higher for the BTMSC systems. This result is contrary to that shown in Figure 4.18 and is due to the shorter frame length used in generating Figure 4.21. Also of interest is the discrepancy between low load performances of the two passive system protocols. In fact for the BTMSR protocols the mean delay at low loads is about a full frame larger than in BTMSC. This can be explained by the fact that in BTMSR, after the generation of a packet to transmit, a station must wait until the headend receives the request for a data slot and generate the corresponding allocation in the next frame, before transmission can proceed. For BTMSC, the generation of a packet occurs on

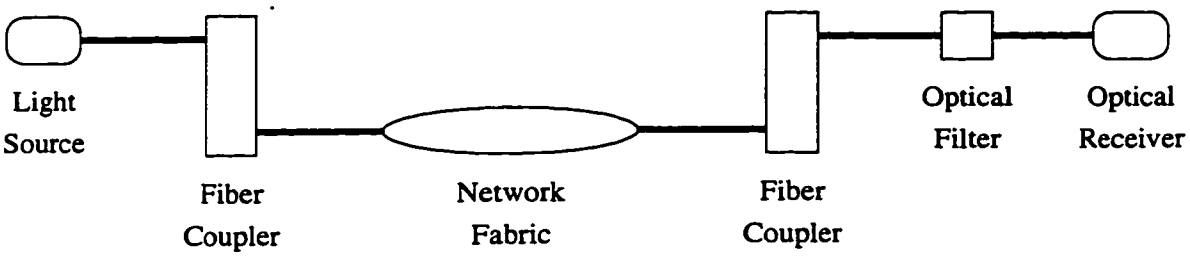


Figure 4.23: Components of the link-budget.

average in the middle of a frame, after which a data slot can be contended for in the next frame and transmitted on average in the middle of that next frame saving the delay of a full frame over BTMSR. It should be noted that the improved performance of the BTMSC protocol comes at the cost of more elaborate user station hardware requirements.

4.4.3 Power Budget

A network's optical power or link budget generally refers to the number of user stations that can be placed on the network so that there is still enough optical energy to drive every receiver in the system. In principle this easily calculated by summing the optical losses through a path from the first transmitter to the final receiver on the network. Figure 4.23 shows a block diagram of a path from optical source to receiver including various components which affect the link-budget calculation.

The problems associated with link-budgets for bus and ring topologies are mostly due to the splitting losses of these couplers or taps. The splitting loss may be supplemented to a point by optimizing the splitting ratio between coupler outputs and by using photonic amplification along the optical path [Gre93].

The analysis assumes the following quantities are defined in decibels. P_{TX} is the optical power transmitted from a tunable laser or led, P_{RX} is the optical power required at a receiver for detection at an acceptable bit error rate, I_{fbt} is insertion loss of each FBT coupler, I_s is loss due to a fiber splice, and I_{filt} is loss due to filtering at the receiver. Finally, the following will also be required in the calculations. N is the number of stations on the system including the headend, and C_{fbt} is the coupling

coefficient of the FBT expressed as a fraction of power leaving the outport as compared to the power entering the inport.

The power available at the input of headend receiver, P_{RX}, for the BTMSC ring is given by

$$P_{RX} = P_{TX} + 10(N-1)\log(C_{fbt}) + 20\log(1-C_{fbt}) - (N+1)I_{fbt} - 2NI_s \quad (4.18)$$

Similarly, for the BTMSC bus the available power can be written as

$$P_{RX} = P_{TX} + 20(N-1)\log(C_{fbt}) + 20\log(1-C_{fbt}) - 2N \cdot I_{fbt} - (4N-2)I_s \quad (4.19)$$

The optical receiver sensitivity, P_{RX} is given by [Gre93]

$$P_{RX} = 10\log(\frac{hcER}{\lambda} \cdot 10^3) + I_{filt} + 2I_s, \quad (4.20)$$

where h is Planck's constant, c is the speed of light in free space, R is the bit rate, E is the receiver sensitivity in photons per bit, λ is the received wavelength, and I_{filt} is the loss due to optical filtering. By equating 4.18 and 4.19 with 4.20 we can find the maximum bit rates at which each network can be run as done in [IK92].

$$R_{BTMSC-ring} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}+20(N-1)\log(C_{fbt})+20\log(1-C_{fbt})-2NI_{fbt}-2(N+1)I_s-I_{filt}}{10})}. \quad (4.21)$$

$$R_{BTMSC-bus} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}+30(N-1)\log(C_{fbt})+20\log(1-Cfbt)-(3N-1)I_{fbt}-(4N)I_s-I_{filt}}{10})}. \quad (4.22)$$

Similar analysis on the BTMSR systems yield the following equations

$$R_{BTMSR-ring} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}+10(N-1)\log(C_{fbt})+20\log(1-C_{fbt})-(N+1)I_{fbt}-2(N+1)I_s-I_{filt}}{10})}. \quad (4.23)$$

$$R_{BTMSR-bus} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}+20(N-1)\log(C_{fbt})+20\log(1-Cfbt)-(2N)I_{fbt}-(4N)I_s-I_{filt}}{10})}. \quad (4.24)$$

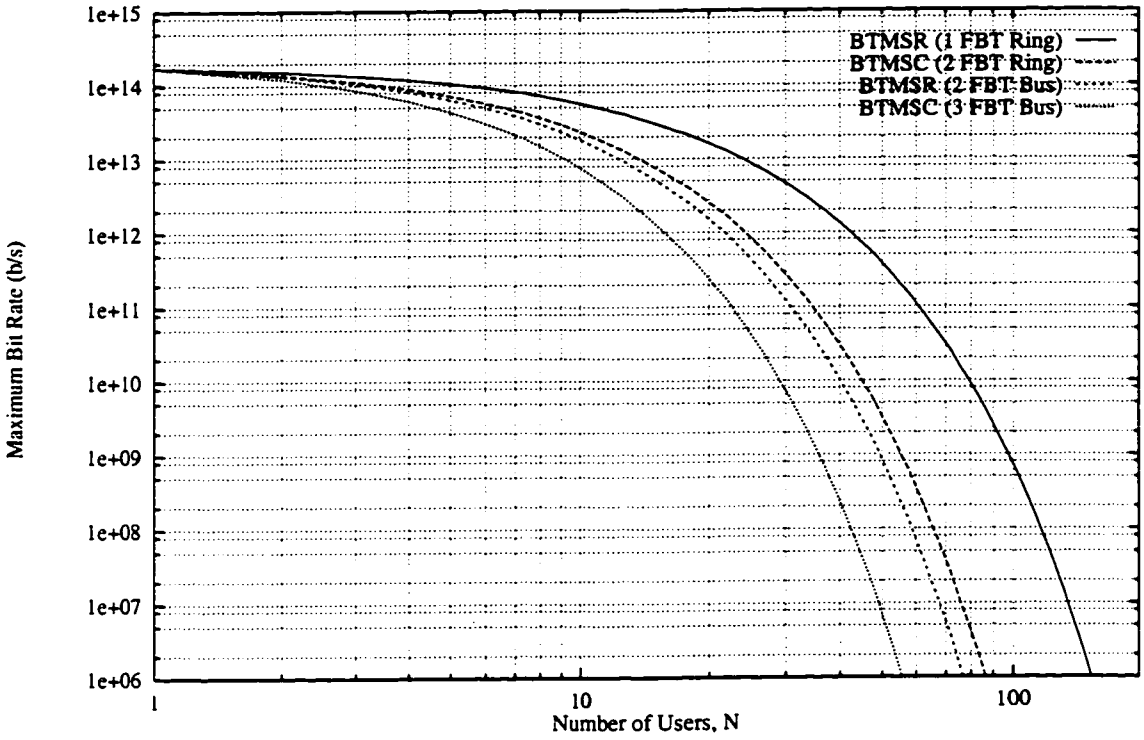


Figure 4.24: Network bit rates as a function of the Number of User Stations in Passive Systems.

Figure 4.24 shows the network bit rates given by Equations 4.21, 4.22, 4.23, and 4.24. The values of λ, P_{TX}, and E were assumed as 1.5μm, 1mW, and 750 photons/bit as assumed in [IK92]. The values for I_s, I_{fbt}, and I_{filt} were assumed as 0.06 dB corresponding to fusion splices, 0.2 dB which is a typical low value, and 1 dB for a typical low filter loss and are described as typical in [Sen92].

Immediately evident from Figure 4.24 is the benefit, other than reduced user station cost, of reducing the number of station taps. The BTMSR design supports a greater number of users for a given bit rate when compared to the corresponding BTMSC design. Another conclusion from Figure 4.24 is that optical amplification is needed for the expansion to more than a handful user stations if reasonable channel rates are to be used for either system.

4.5 Discussion

Four different network designs with reduced station hardware were presented in this chapter. The incentive of the designs was to consider the effects of reducing the complexity of the user stations by a degree not previously considered in the literature. This includes not only a reduction in the number of transmitters and receivers and the degree of wavelength agility required, but the number of passive couplers used to connect the components to a network. Each hardware reduction at a station introduces operating complexities which will be dealt with via media access protocols. For example, as a result of the reduced number of couplers in a BTMSR system, stations must be careful not to transmit while data is being received. On the other hand, BTMSC systems, which utilizes extra couplers, can safely transmit while receiving data from upstream. The architecture of the headend was also discussed and hardware count comparisons were made with other WDM ring and bus networks from the literature.

The media access protocols required for collision free, packet oriented, broadcast and select network operation were then described. It was illustrated how, by means of protocol, a reduced number of fixed or slowly tunable receivers and a reduced number of coupling devices could be supported. With each hardware reduction comes

an associated protocol difficulty. It was shown that in the ring, data slots must perform an INITIAL and RESIDUAL rotation to ensure that payloads reach their destinations. It was then shown that a reduction in the number of passive couplers no longer permitted a mini-slot contention scheme and a more complex reservation scheme had to be introduced.

From the performance section, it was found that performance of passive systems is degraded as the number of users or channels is increased. Delay results indicate that the passive system protocols deliver different low load delay performances with the better performance afforded by the more hardware intensive designs. Finally, power budgets were presented for the designs and indicate that passive coupled networks can support a modest number of stations at fairly high bit rates. These results thus give an indication of the performance versus hardware complexity tradeoff.

Chapter 5

Actively Coupled Architectures

This chapter moves the focus of a station's design from the use of passive couplers to active couplers. It leads with an overview for actively coupled networks and the motivations for the presented designs. The proposed network architectures are then described in terms of their physical implementation aspects. Designs for a ring, folded bus, and dual bus are given. The protocols required for an actively coupled network architecture are considered next. Two station media access protocols are exported from single channel high speed networks to the WDM domain by the use of the active couplers and a set of proposed Station Insertion Gap (SIG) protocols.

Finally, the performance of the ring, folded bus, and dual bus networks based upon the active coupling technologies discussed in this chapter is examined. Approaches will include both analytical and simulation techniques. Capacity and delay performance indices are derived for the various systems followed by an exploration of the optical link budgets for the systems. Results give an indication of the relative strengths and weaknesses of the designs.

5.1 Overview

Normally a single-hop optical LAN is based upon a wideband broadcast-and-select architecture. To ease the need for multiple transmitters and receivers at each station, tunable components are used to provide the station with wavelength-agility. This

allows each station to dynamically access the available wavelengths with a modest amount of transmission hardware. In single-hop LANs, architectures based upon passive star couplers and ring/bus topologies have been considered [Muk92, Ram93]. In terms of optical power budget, passive star networks have an advantage compared with passively tapped rings and busses, since device losses accumulate only logarithmically [Gre93]. However, since there is no positional ordering amongst the stations, synchronization requirements often add a significant degree of complexity to the design. In future networks, wideband optical amplification may help compensate for the poor power performance of rings and busses [Sen92, DSP87].

In most ring/bus LAN designs, [SD93, BM93, GT95], stations passively tap the fiber using a fused biconical taper (FBT) coupler [Gre93]. In this arrangement, some fraction of the upstream power is extracted from the fiber and brought into the station where filtering accesses the various channels. A similar arrangement can be used to couple power from the station onto the fiber. However, because of the passive nature of the taps, an equal fraction of the power must also be coupled out of the fiber [Gre93]. In a typical station design, several such couplers may be required per station as demonstrated in the previous chapter. However, an advantage of this system is that coupling losses are fixed regardless of channel re-tuning activity at the stations. An unfortunate aspect of passively coupled networks is that at high transmission rates it is difficult for stations to synchronize to upstream transmissions since packets are not regenerated at the stations. In the previous chapter a number of different designs are considered using a wide range of station attachment arrangements, couplers and protocols.

A ring or bus can also be designed using active wavelength-selective couplers. The functionality of an active wavelength-selective tap, which may be based on acousto-optics [IK92], is shown in Figure 5.1. In this case, the coupler has direct access to a specific WDM channel or set of channels. In the figure, a set of wavelengths is shown passing through a user station. A single wavelength, λ_i, is coupled out of the fiber and into the station for processing. Following this, λ_i is re-inserted by the coupler as shown in the figure. It is also possible that multiple channels may be simultaneously extracted and re-inserted by the tap.

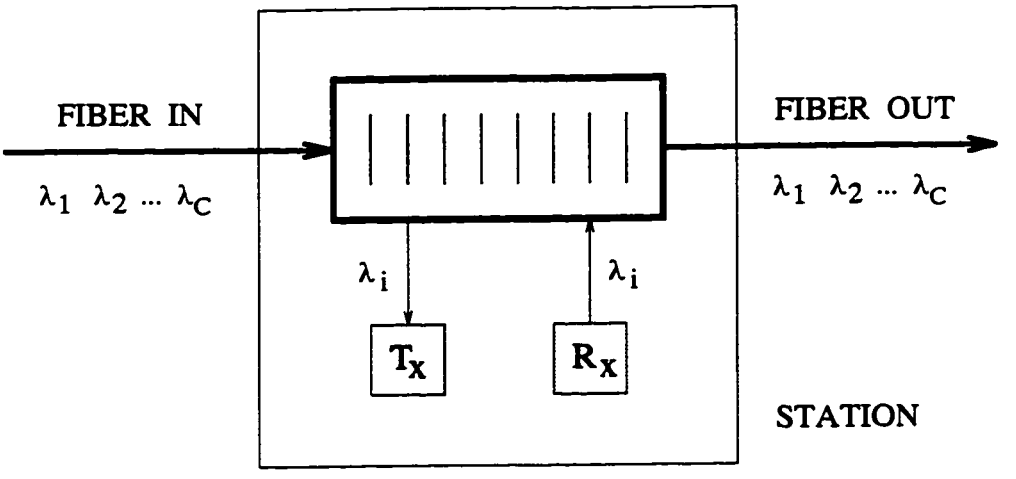


Figure 5.1: Wavelength-Selective Station Coupler

In [IK92] these devices were used as drop-and-add multiplexors to implement a set of rings to which stations have full attachment. In the design, each station requires C transmitters and receivers, one for each WDM channel accessed. Such a requirement would obviously be too expensive in a LAN where the cost of the user station attachment is of primary concern. Others have proposed the use of active couplers to provide fault tolerance [WL90] and wavelength routed networks [ea93b, ea93a].

5.2 Network Architectures

In the designs considered, it is assumed that there are C WDM channels per fiber and a total of N user stations where typically $N \gg C$. The folded-bus and ring systems are illustrated in Figures 5.2 and 5.3 respectively. In the figures, stations are shown with their wavelength-selective attachments. The ring design requires a single coupling point at each station where both transmission and reception will occur. It can be seen that the folded-bus design requires two station coupling points. The upper one is used to write the bus when transmitting packets, and the lower one is used for reception. Although various combinations of passive and active tapping are possible, we will assume that wavelength-selective taps are used in both places. In

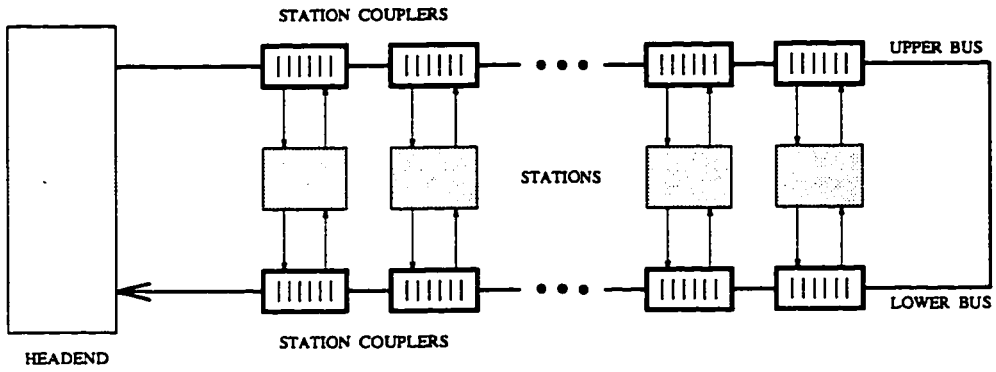


Figure 5.2: Active Folded-Bus

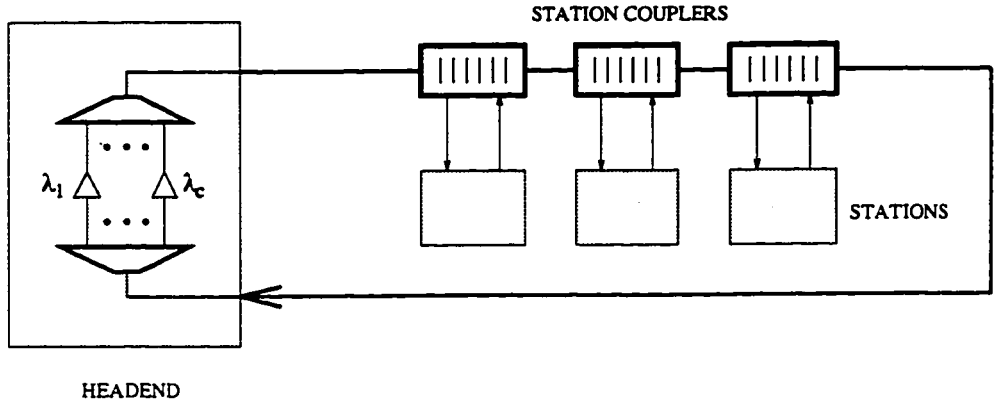


Figure 5.3: Active Ring

Figure 5.4, the dual-bus version is shown. In this case, each station must attach to the forward and reverse bus and must have access to two wavelengths on each fiber. Since a DQDB-type protocol is used in this design, each station must be able to send and receive on each accessed forward and reverse bus wavelength. When configured properly, the couplers will act as "fail safe" devices. That is, in the event of power loss to the station, optical energy will continue to propagate through the station subject only to the insertion losses of the devices.

5.2.1 Headend Architecture

In each of the proposed designs, a system headend is required for each wavelength channel. Because the operation of each channel is independent, every channel can be

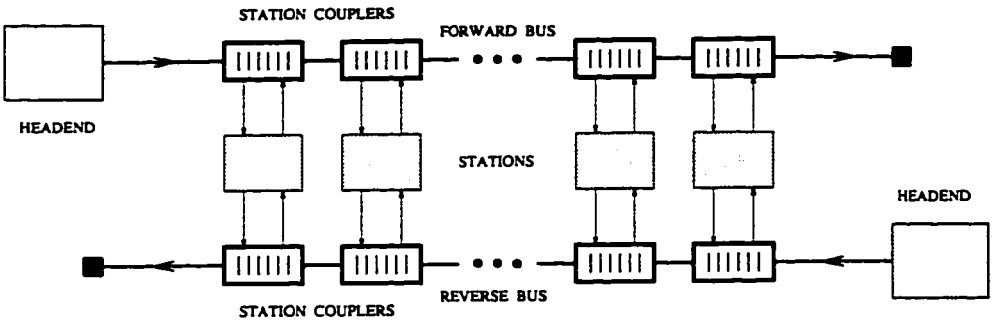


Figure 5.4: Active Dual-Bus

thought of as a separate bus or ring. Thus a single fiber will have many virtual buses or rings implemented via separate wavelengths. In future discussions the terms bus or ring and wavelength channel will be used interchangeably. The headend is attached to all C wavelengths and is responsible for a number of protocol functions discussed in the next section. Since the number of stations is expected to be larger than the number of channels, the cost of the headend is spread across the total number of stations.

For the ring and folded bus designs, the headend is architecturally similar to the headend of the passive BTMSC and BTMSR systems. That is, it is envisioned that a channel controller model will be used and that each channel controller will be interconnected via a low bandwidth backplane bus. Each channel controller will be responsible for the operation of the channel to which it is attached. As will be seen in the next section, the complexity of headend operation is greatly reduced in the active station architectures.

For the dual bus system two headends are required, one for each bus. It is expected that in an operational system both headends would be co-located in order to share timing and synchronization information. The functions of the headend can be broken into two parts; the slot generator and the bus terminator. The slot generator portion of the headend is located at the beginning of each bus and is used in the creation of empty slots and other synchronization events. The bus terminator, located at the end of each bus, is used for certain bus latency measurements. Each terminator and generator should consist of an array of receivers and transmitters in addition to the

required control circuitry. As a result, a channel controller construction may be used, incorporating a bus terminator and slot generator in a single location.

5.2.2 Station Channel Operation

Stations access various channels by re-tuning their active couplers. Additional details of the station design are shown for the folded-bus and ring in Figures 5.5 and 5.6. Each station inserts a transmitter/receiver pair into the station's "HOME" channel wavelength. In the folded-bus system, this is done on the lower bus. This attachment is pre-configured and stations are distributed across the available channels in some reasonable fashion. Each station thus has a fixed presence on its HOME channel and is always an active part of that particular ring or bus. In addition to the HOME channel transceiver, an additional transmitter/receiver pair at each station is used to effect transmission to other stations. In the ring, we will assume that the additional channel is provided by the same station coupler as shown in Figure 5.6. Alternately, the ring design could use two couplers per station.

A similar arrangement is shown for the dual-bus in Figure 5.7. In this case, two wavelengths are extracted from each fiber. We will assume that this is performed by a single tap on each bus. Each station has a HOME channel on each bus to which it fix-tunes one transmitter/receiver pair as shown. The other Tx/Rx pair is used to dynamically access the other dual-bus wavelengths. This pair is always tuned to the channel currently being accessed (shown as λ_x). Media access is attained using a variation of the IEEE 802.6 DQDB protocol [80290b] as discussed in the next section.

Station transmission is accomplished by first tuning the appropriate active coupler(s) to the desired destination channel. In Figures 5.5 and 5.6 the stations are shown currently accessing the bus or ring at wavelength λ_X. A transmitting station is thus actively inserted into the desired destination bus at wavelength λ_X. Once inserted into the appropriate bus, a station must execute the media access protocol discussed in the next section. In the dual-bus system, a station transmits by tuning both couplers in order to insert itself into the desired wavelength (shown as λ_X) on both forward and reverse busses. This is necessary since a station must be able to

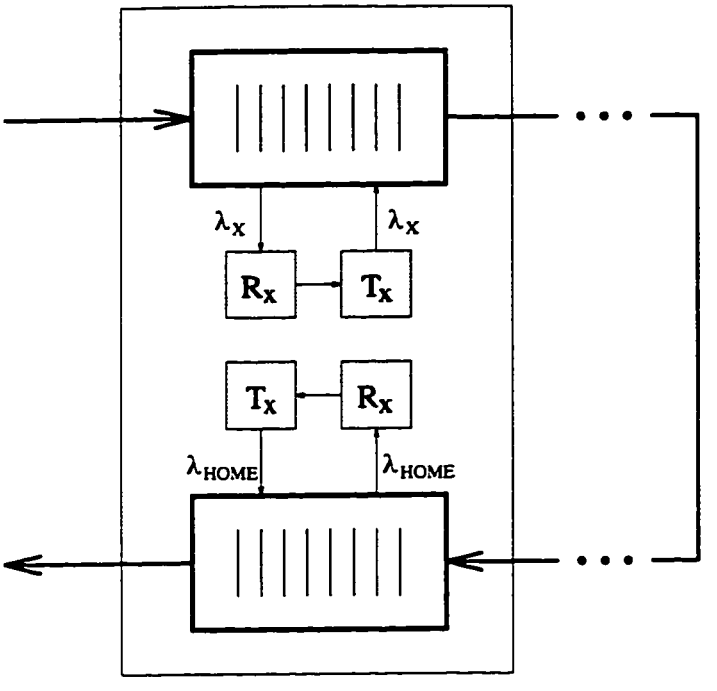


Figure 5.5: Folded-Bus Station

access both forward and reverse bus wavelengths in order to execute the IEEE 802.6 DQDB protocol. However, a station must tune each of its forward and reverse bus taps separately, and in accordance with the tuning constraints discussed in the next section.

5.2.3 Home Channel Assignment

As in the passive systems, each user station is assigned a HOME channel for reception. Similarly, some policy should be used for this configuration. For convenience a round-robin allocation scheme will be assumed for this thesis as was done for the passive systems. Again, in this scheme, station A is assigned the home channel $HOME$ as defined by

$$HOME = A \bmod C \tag{5.1}$$

where $1 \leq A \leq N$ and C is defined as the total number of wavelength channels.

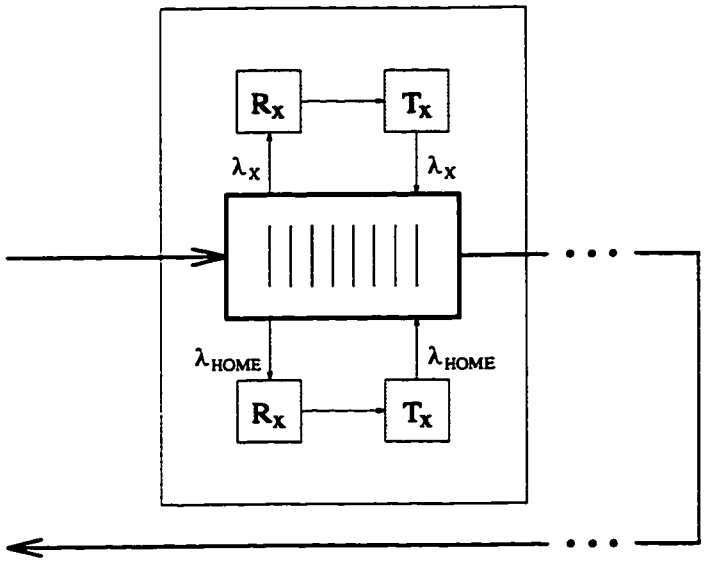


Figure 5.6: Ring Station

Recall that N is the number of user stations in the system. Using this scheme, a source station can easily determine the home channel of any destination station in a decentralized fashion just as with the passive systems. This results in two conceptual views of the network topology: a physical topology in which each station is physically connected to a single fiber, and a virtual topology for which each receiver and transmitter may be connected to a different wavelength channel. The virtual topologies resulting from the round-robin allocation scheme for the folded bus, ring and dual bus systems are similar to those of the passive designs.

Certain skewed traffic cases may, however, require alternative assignment policies or dynamic HOME channel reconfiguration schemes. For example, in a case where many "busy" stations exist, it may be necessary to spread those "busy" stations over the available channels rather than have them concentrated on only a few channels. The optimal solution for an arbitrary traffic skew distribution is a problem which is out of the scope of this thesis. Any such scheme, however, would be implemented through the network headend.

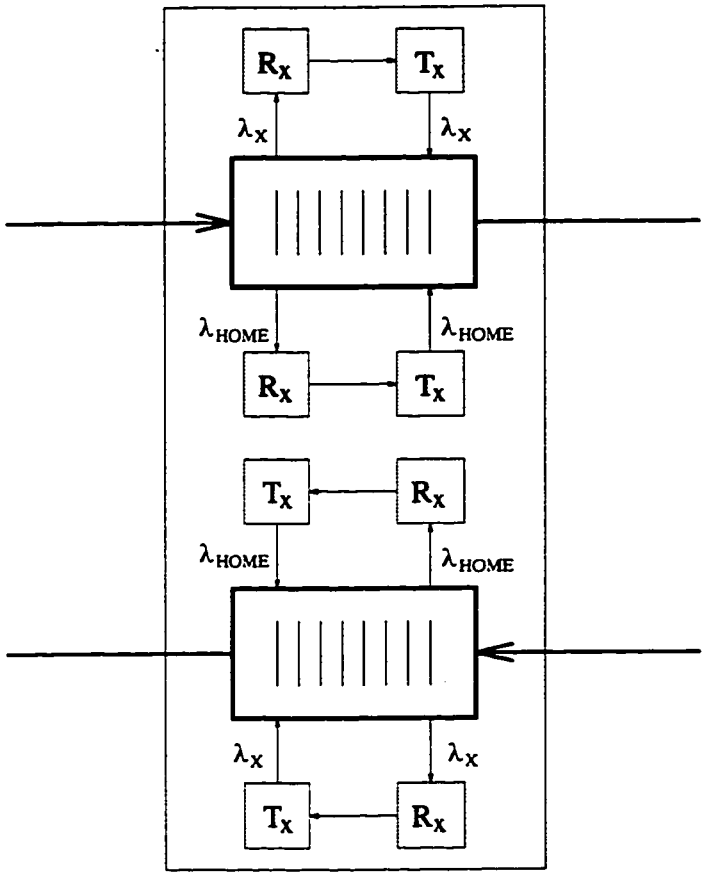


Figure 5.7: Dual-Bus Station

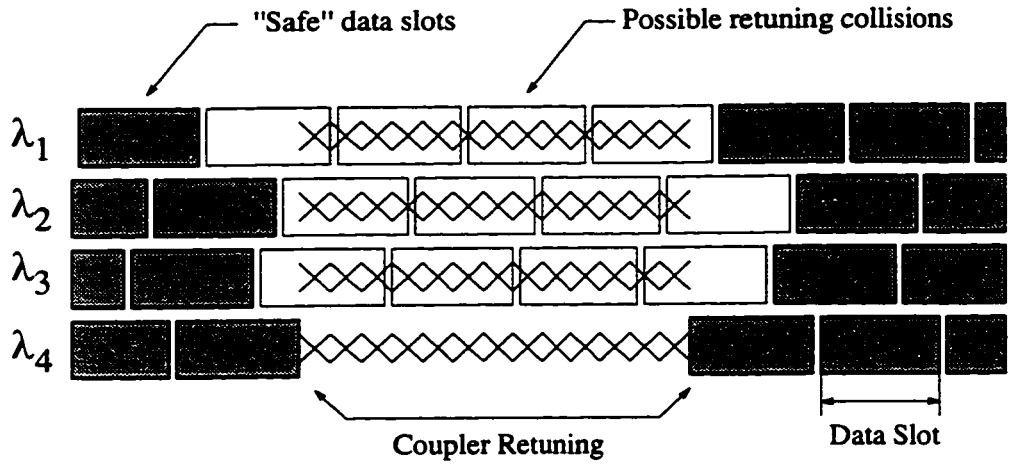


Figure 5.8: Affected data-slots due to re-tuning

5.2.4 Re-Tuning Collisions

In all systems, if the station re-tuning process is undertaken at an inappropriate time, it may interfere with packets propagating through the station. This action is referred to as a "re-tuning collision" and may result in the destruction of packets. Normally, the active couplers would require some non-zero re-tuning time. During such an interval, the status of transmissions propagating through the coupler may not be determinate. An example of this is shown in Figure 5.8 where four wavelengths are shown passing through a station coupler. Although λ_4 is clear throughout the re-tuning interval, full data slots passing through the coupler on the other wavelengths may be destroyed.

This would occur as follows. Each station is an active part of the wavelength that it is currently connected and thus introduces a bit latency to that channel. In the example this would be λ_4. At some point during the tuning process, the station is removed as an active component of that channel. Similarly, looking at the new destination channel of the station, we find that at some point during the coupler re-tuning, the station will enter that channel introducing a station latency that was not present prior to the tuning. This results in two sources of packet destruction. On the original channel, bits would be lost in passing slots, thus shrinking them. Whereas on the new channel, slots would be padded and elongated as the new station

latency is introduced. In each case, vital slot information could be lost. Since the destination channel may not be known a-priori, it is reasonable to assume that the destruction of packets could occur on any of the channels. Notice that this source of packet destruction is not unique to acousto-optic devices, but may occur in any active coupling device including couplers utilizing opto-electro-opto conversions. A second source of loss for acousto-optic filters should also be considered. As the RF signal used to tune the acousto-optic device is sweeped to a new frequency, we should find that optical power levels in not only the original and new channels will be affected. These power fluctuation would be very undesirable for downstream stations.

Re-tuning collisions are to be avoided and protocols which achieve this are discussed in the next section. The problem of avoiding re-tuning collision is also compounded for another reason. It is very difficult for a station to derive global slot-timing information from any one observed channel. This is because the latencies of a particular channel change depending upon the number of stations currently inserted into it. In Figure 5.9 the latency introduced onto the channel to which a station is currently inserted is illustrated. Thus slot boundaries on different channels do not remain aligned as they propagate along the bus. For the purposes of this thesis , it will assumed that if a coupler is re-tuned under these conditions, any transit packets within the coupler at the time of tuning are destroyed.

5.2.5 Optical Devices

For the active architectures, the focus is not so much on the reduction of hardware as on how protocols and system operation can be simplified over equivalent passive designs. Table 5.1 shows the hardware requirements for BTMSC, BTMSR, and the active network designs in terms of the number of required fast-tunable transmitters, fixed-tuned receivers, coupling points, and fixed generation transmitters and receivers. N is defined as the total number of user stations and C is the total number of WDM channels on the network. The first three hardware categories of Table 5.1 are components of the user stations whereas the generator transmitters and receivers are system headend components. In comparison with the passive designs, the active

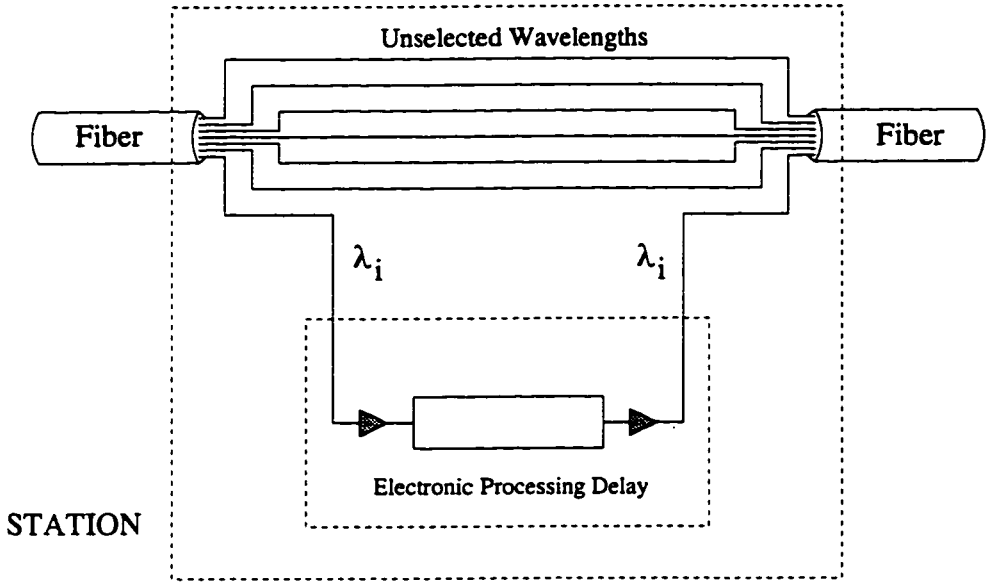


Figure 5.9: Stations Introduce a Bit Latency onto the Inserted Channel

Protocol	Topology	Tunable Tx	Fixed Rx	Coupling Points	Fixed Gnr.Tx	Fixed Gnr.Rx
BTMSC	Passive Ring	N	N	$2N$	C	C
BTMSC	Passive Fld Bus	N	$2N$	$3N$	C	C
BTMSR	Passive Ring	N	N	N	C	C
BTMSR	Passive Fld Bus	N	$2N$	$2N$	C	C
Morewell	Active Ring	$2N$	$2N$	N	C	C
Morewell	Active Fld Bus	$2N$	$2N$	$2N$	C	C
DQDB-like	Active Dual Bus	$4N$	$4N$	$2N$	C	C

Table 5.1: Active Network Transceiver Requirements

systems utilize a competitive number of coupling points. It should be noted, however, that the use of active couplers is much more expensive than the alternative passive FBT coupler. We also find that because of the active coupling, the number of required receivers is at least doubled from the passive systems. As a result, the use of active couplers in a network design should be viewed not as a station cost reduction method, but as a way to simplify system operations for high speed networks. For example, because signals will be regenerated at stations, synchronization can be easier. Furthermore, the active coupling allows many protocol options to be implemented which are not possible in the passive networks.

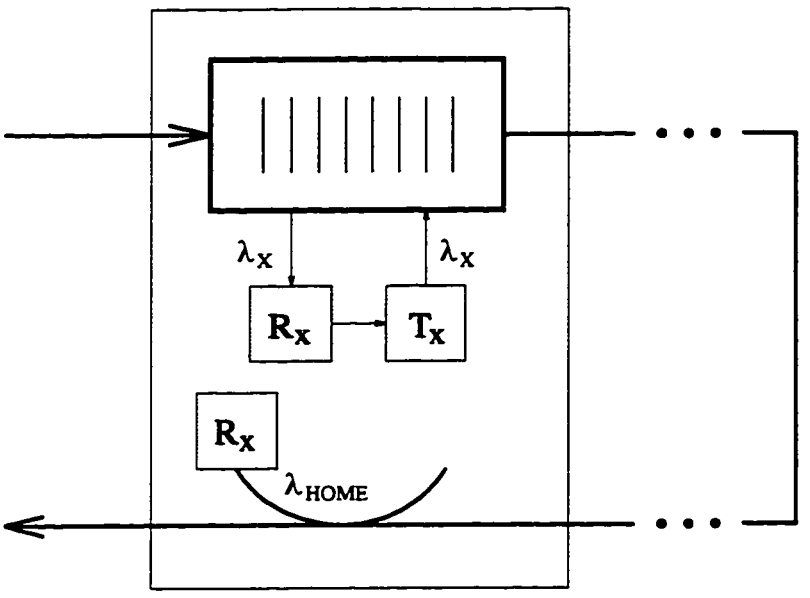


Figure 5.10: Hybrid Active/Passive Coupled Folded Bus Station

5.2.6 Hybrid Coupled Architectures

An option in replacing the active couplers with passive couplers exists. In particular, passive FBT couplers can be integrated into the design of the stations at the network access points which do not require frequent tuning. Such an access point would be at the HOME channel transceivers. For example, in the folded bus station design, the active couplers on the lower bus at each station may be substituted with passive FBT couplers as shown in Figure 5.10. In this case, only a receiver tuned to the HOME channel is required. The HOME channel transmitter can be eliminated without affecting system operation.

A similar approach can be used in the ring station design. In the ring case however, the active station coupler is still required for transmission access. A passive coupler is simply added to the station design to which a tuned receiver is attached. This is illustrated in Figure 5.11. An identical configuration is available for both the forward and reverse bus access points in the dual bus architecture. In the dual bus and ring hybrid coupled designs, a savings of two and one HOME channel transmitters is achieved respectively.

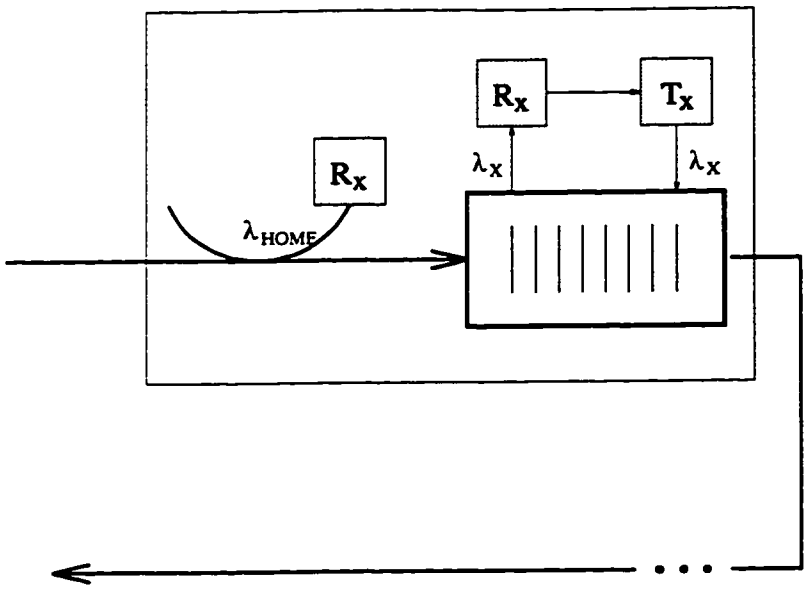


Figure 5.11: Hybrid Active/Passive Coupled Ring Station

5.3 Media Access Control

Since wavelengths may be completely dropped and added by an active coupler, stations can dynamically insert themselves into their channels of interest. As a result, certain protocol functions are very similar to those found in simple single-channel LANs. However, because coupler tuning will take place much more frequently than in other proposals using active devices [WL90, ea93b, ea93a], care must be taken to prevent re-tuning collisions which can occur if a station inappropriately re-tunes its coupler, thus destroying slots in transit through the station. This can be avoided using a simple protocol which eliminates re-tuning collision by synchronizing all station coupler re-tunings. However, the overhead associated with this action can severely restricts the physical parameters over which efficient network operation may be achieved.

In this section, new protocols are introduced which allow for the efficient operation of LANs with active station taps over a wide operating range. Here, a two level model is used. The first level is responsible for giving stations a chance to re-tune their active couplers in order to access various wavelength channels. Several first level protocols, or flavors, are introduced. In the first flavor, the headend simply

ensures that the entire network is cleared of all transmissions which may be destroyed due to re-tuning collisions. While this protocol is very simple, system scalability into the metropolitan area is severely curtailed. Alternate flavors which significantly extend the scalability are also discussed. Flavor B is based on a-priori knowledge of station insertion delays and the second makes use of simple latency measurements made by the system headend. The resulting systems can scale much better into the metropolitan area compared to a simple minded approach. These first level protocols are applied to export some common single channel protocols, which compose the second level of protocol, for the ring, folded-bus, and dual-bus architectures to WDM. For the dual bus, a multi-channel DQDB-like protocol is used for channel access. However, to accommodate the coupler tuning constraints, the forward and reverse bus station couplers must be tuned independently. The ring and folded-bus actively coupled networks utilize an Orwell-like protocol entitled Morewell. To generate cost-effective designs, each station has a small number of transmitters and receivers regardless of the number of WDM channels.

5.3.1 Active Media Access Protocols

In this section we consider the media access protocols required for the ring, folded-bus and dual-bus networks described in the previous section. To prevent re-tuning collision, the headend issues information that synchronizes all station tunings in the first level of protocol. This is achieved by having stations re-tune their active couplers during a time period referred to as a Station Insertion Gap or SIG. SIGs are issued to the stations every t_{SIG} seconds by having the headend write into an appropriate packet header bit simultaneously on all wavelengths. It is envision that t_{SIG} will be on the order of tens to hundreds of microseconds. An example is illustrated in Figure 5.12. In the figure, a SIG of length T_{SIG} is shown on each wavelength. The intent is that during this idle period, station coupler re-tunings can occur safely, thus avoiding re-tuning collisions. As mentioned in the previous section, system timing is complicated by the fact that the propagation latency on a given bus or ring is not fixed but is dependent upon the number of stations currently inserted. As a result,

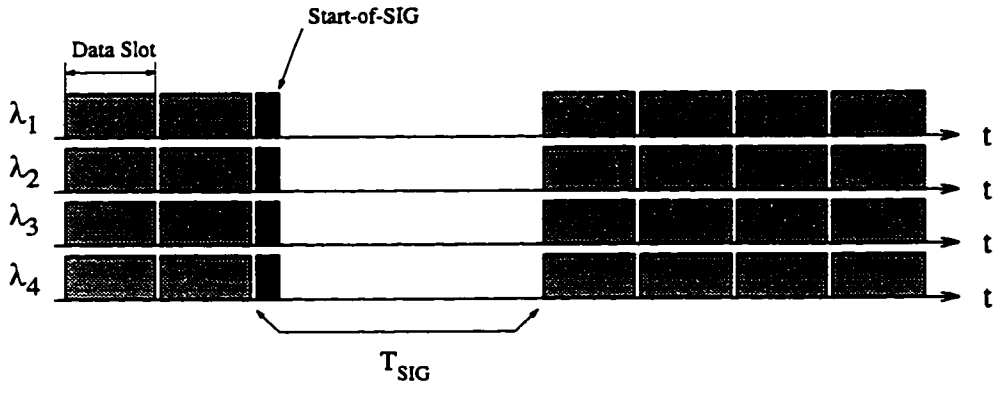


Figure 5.12: SIGs as they are released from the Headend

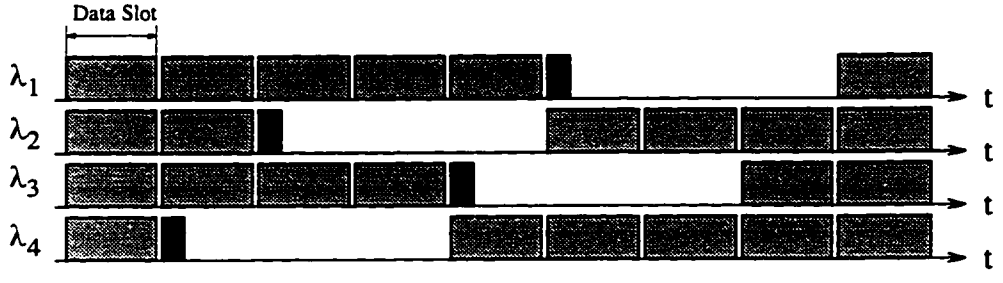


Figure 5.13: SIGs Skew in Time as they Propagate the Network

the SIG on each channel will expand, contract and shift depending upon latency and re-tuning activity. A sample of the SIG timing skew over a number of channels as the SIG propagates the network is shown in Figure 5.13. Thus the value of T_{SIG} must be large enough to accommodate these changes. When not in the process of re-tuning during a SIG, stations execute the second level of media access.

5.3.2 The Morewell Protocol

A particular second level protocol for actively coupled networks is the Morewell protocol which is a multi-channel derivative of the Orwell protocol [FA85]. In Morewell, however, each station maintains a separate queue for each destination channel and the protocol can be applied to either the active folded bus or active ring network designs. During normal operation, the headend issues a continuous stream of empty slots onto each wavelength channel. Each slot contains a BUSY/IDLE header flag

Figure 5.14: Morewell Slot Format

indicating whether it is currently in use or not. The format of a Morewell slot is given in Figure 5.14.

Transmission and Reception

Idle slots on a particular channel may be only used by a station whose tunable transmitter/receiver pair is inserted into that channel. A station wishing to use a particular slot must first ensure that it has tuned its transmitter/receiver pair to the correct channel. Recall that channel re-tuning can only occur safely during a SIG. The station then will examine a slot's BUSY/IDLE flag on the channel to which it has tuned. If the slot is marked as BUSY, the station is not permitted to use that slot an it is allowed passed the station without modification. However, a slot marked as IDLE may be used by the station for transmission. In order to use a slot, the station rewrites the BUSY/IDLE flag to read as BUSY, thus ensuring that subsequent station will not overwrite the slot's payload. The transmitting station then sets the packet's destination address in the DEST field and transmits its data into the payload section of the slot.

Reception of packets occurs only on a station's HOME channel which it continually monitors for packets. When a properly addressed packet arrives at the station, it is copied into the station's local buffers. In the folded bus case, the slot is allowed to continue propagating on to the next station. In the ring case, however, destination release is performed, thus freeing the slot for use by other stations. This action is required in order for a continual supply of free slots to be present on the ring. Notice that in the folded bus system, the headend performs the release of all data slots. All other slots are permitted to pass by the station without any protocol specific action being performed.

Fairness Mechanism

A cycle-based mechanism is used to prevent unfair use of data slots by stations having a positional advantage over others in the network. The protocol is motivated by the single-channel technique first proposed in [FA85]. In the Morewell protocol, station i is given a quota of D_i slots for *each* channel and quota resets are controlled by the headend.

For the folded bus version of the protocol, stations closest to the transmitting side of the headend are favored. As slots are released, the station closest to the headend will encounter no resistance in trying to use a data slot. Subsequent stations, however, will see the full slots as they are used by upstream stations. As a result, downstream stations may starve for data slots, while the upstream stations satiate themselves. In the ring version, because of destination release, things tend to be more fair. However, there can arise pathological traffic patterns which will cause stations to starve for slots.

In the fairness scheme, each time a station transmits, the transmission quota for that channel is decremented in the station. If the count reaches zero for a given destination channel, the station must suspend transmission to that channel until the quota is reset or the station tunes away from that channel. The length of a fairness cycle is then dependent on the total number of stations, the transmission quotas granted to the stations, the total load placed on the network, and the particular traffic mix.

Resets are generated by the CCs and are placed within the QUOTA RESET field of a slot on a station's HOME channel. Resetting is independent for each channel and is initiated whenever the headend sees an empty data slot. For the folded bus system, as soon as the headend detects an empty slot, it releases a slot containing a quota reset. As a precaution, the headend does not allow multiple quota resets to be propagating on any one channel.

For the ring system things are a little more complicated. Because of destination release, slots that appear empty when they return to the headend may have been indeed used at some point. One method of detecting such slots would be to have a USED field in each slot that is set the first time a slot is used. This field would

then remain marked as used even after it has delivered its data and been re-marked as empty. Upon returning to the headend, the USED field would be examined and those slots which were never marked as USED would cause a quota reset to be generated for that channel. An alternate method has the headend utilize a TRIAL bit which is set for each empty slot that arrives at the headend. When a slot is used for transmission by a station, the TRIAL bit is cleared, thus indicating to the headend that the slot has been used, regardless of where the slot has been released. This second approach is a version of that used in the Orwell protocol except that in Orwell the trial bit can be set by any station detecting an empty slot. In order to simply the user station designs, Morewell utilizes the centralized version of the quota resetting mechanism. In either approach, protocol performance would be approximately the same. For this thesis the TRIAL field is implemented for the ring network. Again, the headend should limit the number of quota resets propagating any channel at a particular time to one.

Using this fairness scheme, individual stations can be given higher priority access to the network simply by increasing the stations' quota with respect to other stations. At low loads, then, resets will be generated frequently and the fairness cycle will be short. At increasingly higher loads the frequency of resets will decrease resulting in longer fairness cycles. The reader is referred to [FA85] for a more detailed discussion of the choice of N_q. In general, the choice of N_q involves a balance between efficiency and the duration of time over which fair operation is achieved.

5.3.3 A Modified DQDB Protocol

In the dual-bus network, the headend for each bus performs a similar function as that described above. In addition, it is assumed that the headends are synchronized so that they release SIGs on each bus at the same time. It is expected that the headends would be co-located in an operational system. In this case, the forward and reverse bus station couplers must be tuned independently at a time which depends upon the arrival of the SIG on each bus. When accessing a particular dual-bus wavelength, a variation of the IEEE 802.6 DQDB protocol [80290b] is used. The slot format for the

Figure 5.15: Modified DQDB-Like Slot Format

modified system is given in Figure 5.15.

Transmission and Reception

Transmission occurs using the basic DQDB mechanism. That is, each station maintains a number of counters that implements a virtual distributed global queue. When a station desires to transmit a packet, a COUNTDOWN counter is loaded with the position that the packet belongs in the global queue. The station then allows that number of empty slots to pass by unaffected. Having done this, the station then is permitted to transmit into the very next empty slot. Please refer to Chapter 3.1.5 for more details of the basic DQDB mechanism or see [NBH88].

Reception, like all other systems previously discussed, occurs only on a station's HOME channel which it continually monitors for packets. When a properly addressed packet arrives at the station, it is copied into the station's local buffers. Now, however, since each station is an active part of the particular wavelength it is inserted, destination release can be implemented.

Reuse Protocol

Many options are available for DQDB with erasure nodes. For this thesis a simple basic counter mechanism [BV95] is selected, although many algorithms which exhibit superior fairness exist [BV95, Has94]. Using the basic counter mechanism, each time a data slot is received at a station, it is released for use by downstream stations. However, to maintain global queue counter consistency, for each data slot released a corresponding slot reservation must be removed from the opposite bus. This is achieved by means of a pair of ERASED counters, one for each bus. Each time a data slot is released the counter for the opposite bus is incremented. Each slot with a REQUEST bit set will cause a station to examine its ERASED counters and, if

non-zero, clear the REQUEST bit for that slot and decrement the ERASED counter for that bus. Thus, in most cases, each released slot will be reflected to upstream stations by having a request removed from the network.

Fairness Mechanism

It has been demonstrated that the basic DQDB protocol can be unfair [HCM90]. For this multi-channel version of the protocol a Bandwidth Balancing (BWB) mechanism is employed. The mechanism simply requires that stations allow an extra free slot to pass without affecting the REQUEST counter after every BWB'th transmission, where BWB is the bandwidth balancing factor. Notice that each time a station tunes to a new channel, all of its counters are reset complicating all fairness issues. It is not the purpose of this thesis to optimize fairness in these networks and does this discussion not contribute to the subject any further.

5.3.4 The Station Insertion Gap

In the folded-bus, ring and dual-bus systems, the headend includes a parameter in the start of SIG indicator which is used for certain SIG timing functions. We will denote this parameter as L_{WAIT}. Upon seeing the start of SIG indication, a station that wishes to re-tune waits for a time period given by L_{WAIT} before starting to re-tune its coupler. Typically the value for L_{WAIT} would be contained in the payload portion of a slot. The delay indicated by L_{WAIT} is necessary to ensure that at any point on the network, the SIG indicator has passed on all wavelengths before any coupler re-tunings occur, ensuring the safety of all slots and the SIG indicator. The length of the SIG itself is denoted by T_{SIG}. Therefore, after the start of SIG indication, the headend waits T_{SIG} seconds before issuing new slots on all channels.

SIG Length Derivation

In the ensuing discussion, we consider the timing associated with one particular SIG. The purpose is to determine the minimum possible values for L_{WAIT} and T_{SIG} so that wastage can be minimized. In Table 5.2, some variables are defined for this analysis.

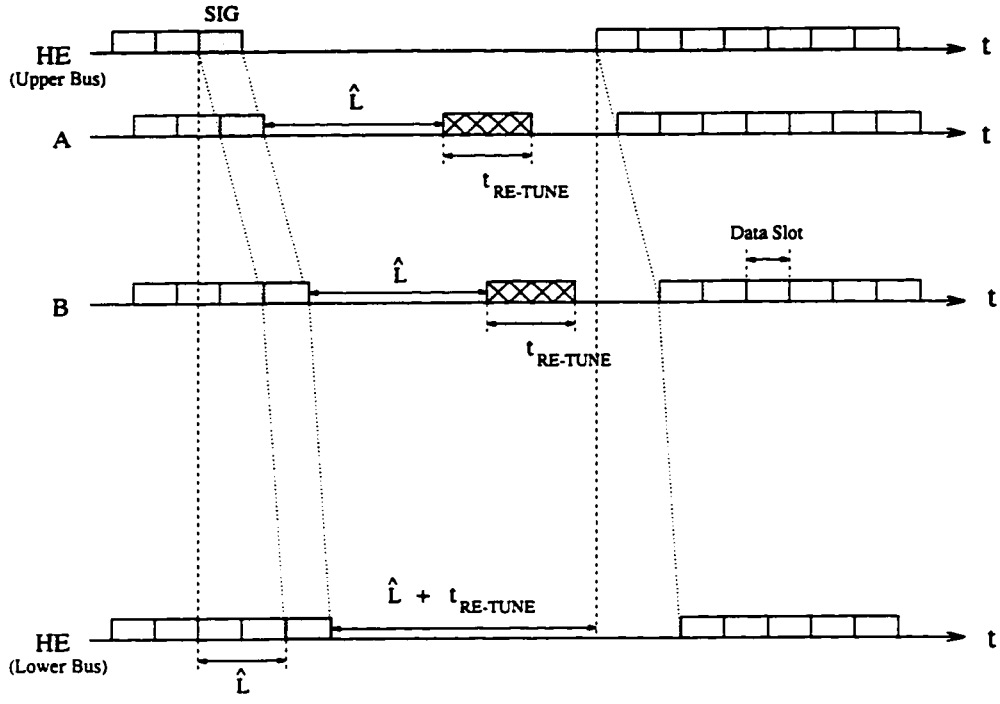


Figure 5.16: SIG Protocol Timing

For the purposes of this development, we consider various activities at a particular reference point x on the folded forward bus. In the dual-bus network, the following discussion applies independently to each bus. The ring network will be discussed separately.

As discussed above, SIGs of length T_{SIG} are released at the headend simultaneously across all channels as depicted in Figure 5.12. However, due to uneven latencies introduced by inserted stations on varying channels, these SIGs do not remain aligned as they propagate along the bus. It is clearly not sufficient for the headend to simply generate SIGs whose length exceeds the worst-case coupler tuning time, $t_{re-tune}$. Instead, the SIGs must be long enough so that a "dead" period, during which no slots are present, will be available on all channels at the same time and place. Using the definitions from Table 5.2, it is obvious that the following two constraints must be satisfied.

$$t_i^{Tune-Start}(x) \geq \max_i t_i^{Start}(x), \tag{5.2}$$

t_o	SIG release time at headend
t_e	Headend release time of new data slots, indicating the end of a SIG
T_{SIG}	The length of the SIG ($= t_e - t_o$)
$\tau(x)$	Propagation latency from headend to position x, not including insertion delay
$t_i^{Start}(x)$	Start of SIG arrive time at position x on channel i
$t_i^{End}(x)$	End of SIG arrive time at position x on channel i
$t_i^{Tune-Start}(x)$	Tuning start time for a station at position x on channel i
$t_i^{Tune-End}(x)$	Tuning completion time for a station at position x on channel i
$t_{re-tune}$	Coupler tuning time (Device Parameter)
$I_i(x)$	Accumulated insertion delay on channel i at position x

Table 5.2: Variable Definitions

$$t_i^{Tune-End}(x) \leq \min_i t_i^{End}(x). \tag{5.3}$$

Constraint 5.2 states that a station at position x on a bus cannot begin tuning until the start of SIG (SOS) on every wavelength has passed that point. This ensures that tuning will not destroy the SOS or any (possibly full) data slots on any channel. Constraint 5.3 states that tuning activity at a station must be completed before any new data slots propagate past the station. Now, note that the arrival time of the SOS is given by

$$t_i^{Start}(x) = t_o + \tau(x) + I_i(x). \tag{5.4}$$

That is to say, the arrival of a SOS will be at a time corresponding to the network propagation delay plus the inserted delay due to the active station interfaces, after the release of the SOS. On wavelength i, consider the smallest time that a station at position x must wait, following the SOS, before re-tuning its coupler. This is denoted by $L_{wait}^i(x)$, and is given by

$$L_{wait}^i(x) = \max_j t_j^{Start}(x) - t_i^{Start}(x). \tag{5.5}$$

This ensures that a station on channel i will refrain from tuning long enough to allow the SOS to pass on all other wavelengths. Using Equation 5.4, we have that

$$L_{wait}^i(x) = t_o + \tau(x) + \max_j I_j(x) - t_o - \tau(x) - I_i(x)$$

$$= \max_j I_j(x) - I_i(x). \tag{5.6}$$

This represents the timing skew between an arbitrary channel i and the channel with the maximum amount of inserted delay. Practically, a common minimum value of $L^i_{wait}(x)$ for any channel at position x is therefore given by

$$L_{wait}(x) = \max_i I_i(x) - \min_i I_i(x). \tag{5.7}$$

Finally, since the waiting time parameter issued by the headend is the *same* for all stations, it must be chosen so that it accommodates its worst-case value. This is given by

$$\begin{aligned} \tilde{L}_{wait} &= \max_x L_{wait}(x) \\ &= \max_x \{\max_i I_i(x) - \min_i I_i(x)\}. \end{aligned} \tag{5.8}$$

This expression is equal to the largest station insertion skew measured across all channels anywhere along the bus.

To find the minimum required SIG length given a value for L_{Wait}, notice that station tuning at position x on channel i will start at

$$t_i^{Tune-Start}(x) = t_o + \tau(x) + I_i(x) + L_{Wait}. \tag{5.9}$$

Similarly, tuning must be completed in accordance with Equation 5.3. Now notice that

$$\begin{aligned} t_{re-tune} &= t_i^{Tune-End}(x) - t_i^{Tune-Start}(x), && (5.10)\\ &\leq \min_i t_i^{End}(x) - t_o - I_i(x) - L_{Wait} - \tau(x), && (5.11)\\ &\leq t_e + \tau(x) + I_i^*(x) - t_o - I_i(x) - L_{Wait} - \tau(x), && (5.12)\\ &\leq t_e - t_o + I_i^*(x) - I_i(x) - L_{Wait}, && (5.13) \end{aligned}$$

where $I_i^*(x)$ is the value of station insertion delay on channel i *after* all station tunings have completed. In this result, we have used Equations 5.2 and 5.3. The length of a SIG required to satisfy the tuning constraints for channel i at position x is therefore

$$T^i_{SIG}(x) = t_e - t_o \geq L_{Wait}(x) + t_{re-tune} + I_i(x) - I_i^*(x). \tag{5.14}$$

Maximizing over all channels we now have that

$$T_{SIG}(x) \geq L_{Wait}(x) + t_{re-tune} + \max_i I_i(x), \tag{5.15}$$

where we have also assumed the worst-case value of $I_i^*(x) = 0$, which assumes that the station in question tunes to a channel where there is currently no upstream station insertion latency. Finally, maximizing over all bus positions gives

$$T_{SIG} \geq \tilde{L}_{Wait} + t_{re-tune} + \max_{i,x} I_i(x) \equiv \tilde{T}_{SIG}. \tag{5.16}$$

Equations 5.8 and 5.16 give the minimum values of L_{wait} and T_{SIG} possible under the proposed protocol. Unfortunately however, it is not possible to determine these values in practice, since the $I_i(x)$ values are not known to any one station. A complicating factor is that these values *change* as stations tune from one bus to another. For this reason, various bounds on these parameters must be used instead. Equations 5.8 and 5.16 will however be used as a proof of correctness for the subsequently described approximations of $\tilde{L}_{wait}$ and $\tilde{T}_{SIG}$.

In this thesis, three different upper bounds for $\tilde{T}_{SIG}$ are considered. In the first approach, (A), no a-priori knowledge of fixed propagation or station insertion delays is assumed. Instead, the headend measures the total latency of each WDM bus prior to the SIG, and uses the maximum value obtained to compute the bounds. Note that the maximum measured bus latency is given by $\tau + I_{MAX}$ where τ is defined to be the total forward and reverse bus latency not including station insertion delays. I_{MAX} is defined to be the maximum inserted delay across all busses and includes both forward and reverse bus components. In this case,

$$L_{WAIT} = \tau + I_{MAX}, \tag{5.17}$$

$$T_{SIG} = 2(\tau + I_{MAX}) + t_{re-tune}, \tag{5.18}$$

which are clearly larger than $\tilde{T}_{SIG}$ and $\tilde{L}_{WAIT}$. This protocol was the one originally proposed in [TGB95] and amounts to ensuring that the bus has been cleared of all data slots before re-tuning is attempted. Slots are only released by the headend when it can be guaranteed that all tuning is complete. In the second approach, (B), we

assume that the headend knows the fixed insertion delay of each station beforehand. A worse-case upper bound on $\tilde{T}_{SIG}$ can then be computed from

$$T_{SIG} = 2I_{FBtotal} + t_{re-tune}, \tag{5.19}$$

and

$$L_{WAIT} = I_{FBtotal}, \tag{5.20}$$

where $I_{FBtotal}$ is the sum of the bus insertion latencies for *all* stations. In the folded-bus, only the forward bus latencies are used. In this case, a system may be designed where the estimate is based on a maximum station count and thus no timing information need be included in the SIG indication. The third approach, (C), has the headend continuously measure the current bus latencies, then compute I_{FBmax}, which is used as the value of L_{WAIT}. Here, I_{FBmax} is defined to be the maximum insertion delay on the forward bus only. Then,

$$T_{SIG} = 2I_{FBmax} + t_{re-tune}, \tag{5.21}$$

and

$$L_{WAIT} = I_{FBmax}, \tag{5.22}$$

which is easily seen to be an upper bound on $\tilde{L}_{WAIT}$ and $\tilde{T}_{SIG}$. In this scheme, the headend must also know the value of τ and the home channel station assignments so that I_{FBmax} can be determined from the total latency measurements. This is the most sophisticated method and requires that the headend collect and disseminate the required information. In Table 5.3, the bounds used for each SIG protocol flavor is summarized. Complexity refers to the effort with which timing information for the SIG is collected.

5.3.5 SIGs on a Ring Network

For a ring network things are once again more complex. SIG timing on a ring is complicated by the fact that many data slots will routinely have to pass around the headend in order to reach their destinations. This adds another dimension to the SIG procedures discussed in the previous sections. As in the bus, the headend acts

Protocol Flavor	L_{WAIT}	T_{SIG}	Complexity
A	$\tau + I_{MAX}$	$2(\tau + I_{MAX}) + t_{re-tune}$	low
B	$I_{FBtotal}$	$2I_{FBtotal} + t_{re-tune}$	medium
C	I_{FBmax}	$2I_{FBmax} + t_{re-tune}$	high
DELUXE	calculated	calculated	very high

Table 5.3: Protocol Parameter Bound Summary

as a source of slots on each of the WDM channels. However, since stations have only a single tapping-point, full slots which arrive to the headend are passed through unaffected. That is, the headend functions as a conventional ring station thus enabling full ring connectivity.

As with the bus systems, SIGs are used in the ring in order to synchronize the re-tuning activity. For a SIG to be successful and preserve all data slots from re-tuning collisions, it becomes necessary to ensure that all data slots that are to pass the headend do. This is achieved by extending the length of each SIG to contain a phase during which all transmission rounding the headend can be guaranteed to be complete, after which the previous discussion dedicated to the bus networks may be applied. In particular, the time a station delays from when it receives the SOS to the timing tuning occurs is lengthened. It turns out that L_{WAIT} should be twice the largest end-to-end propagation delay of the network. For this thesis,

$$L_{WAIT} = 2(\tau + I_{MAX}), \tag{5.23}$$

$$T_{SIG} = 4(\tau + I_{MAX}) + t_{re-tune}, \tag{5.24}$$

which is analogous to protocol A for the bus systems. In this case however, the ring endures two clearing cycles, the first of which guarantees the delivery of packets rounding the headend. The second clearing cycle is used to preserve all data slots including the SOS. What happens is that after the SOS is released, the first delay of the end-to-end propagation delay will ensure that the ring is cleared of all data slots that do not round the headend. During the second part of the delay, all other data slots will be guaranteed to be delivered as they will not make a second trip around the headend and the problem is reduced to that of the folded bus. For the active ring

network, only SIG flavor A is considered as it is never possible remove the SIG length dependency on the end-to-end propagation delay.

5.3.6 Deluxe SIG Flavor

The previously describe "flavors" of SIG protocols assumed that a limited amount of network state information was available by each station and the headend. Such limitations result in the fact that each station does not know the latest time at which a SOS will arrive. The "regular flavor" protocols thus assumed that after receiving a SOS on arbitrary channel i, a station must wait a time, denoted as L_{Wait}, before re-tuning can begin at the station. It was found that L_{Wait} should be at least the worst case propagation skew between any two channels. The proposed "deluxe flavor" protocols remove the mandatory fixed waiting time of L_{Wait} and collects the information required for each station to deduce the latest time at which a SOS will arrive at each station (ie. $\max_i t_i^{Start}$).

Armed with the knowledge of $\max_i t_i^{Start}(x)$ for each station and $\max_{i,x} I_i(x)$for the entire network, the minimum required SIG length can then be written as:

$$T_{SIG} = t_{re-tune} + \max_{i,x} I_i(x) \tag{5.25}$$

The issue that the deluxe flavor protocol addresses is how to determine the maximum arrival time of a SOS at a station using a hardware configuration that is active on only one channel at a time. It will be seen that in the process of collecting the information required to find $\max_i t_i^{Start}(x)$, the headend will also have the information required to find $\max_{i,x} I_i(x)$.

The argument begins as follows. Assume a physical ordering to the stations in which each station knows its position number which is independent of the channel it is inserted on. This is illustrated in Figure 5.17. At the beginning of each new data cycle (ie. at the end of each SIG), the headend will poll the stations for their current channel assignment. Of course the collection of this information requires a full end-to-end propagation delay and thus the time between SIGs must be greater than or equal to the end-to-end delay which may limit the channel hopping agility of

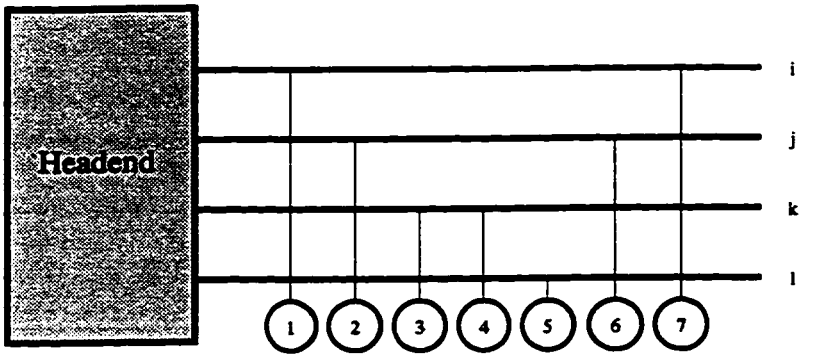


Figure 5.17: Physical Ordering to the stations.

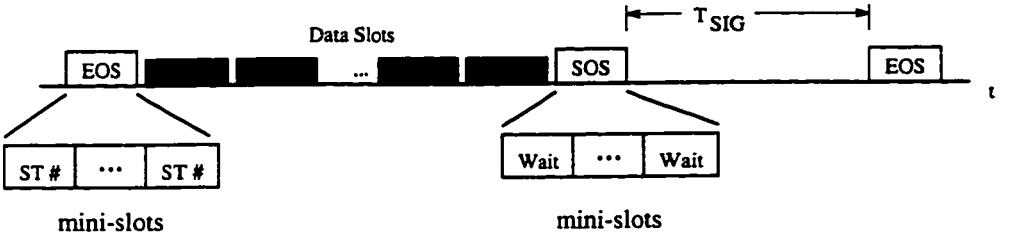


Figure 5.18: Frame format for the deluxe protocol.

the system as the network becomes large.

Figure 5.18 illustrates a proposed frame format for the deluxe protocol. The End Of SIG (EOS) will contain N minislots, where N is the total number of stations in the system. Any station inserted into channel i will write its station number into the first available minislot on that channel. The headend will receive the EOS and the station distribution information that it contains after a full end-to-end propagation delay. Using this information, the headend will fill out the next SOS to contain the required time that *each* station must wait after receiving the SOS on its current channel. For simplicity it is assumed that the SOS also contains N minislots, some of which may be unused. The ordering of station wait times in the SOS will be the same as the order that the EOS was filled in. Thus if station i fills in the EOS in minislot j, its assigned L_{WAIT} value will appear in the SOS at minislot position j.

The headend will contain an algorithm which digests the EOS information, fills out the SOS mini-slots and derives a value for T_{SIG}. Both the SOS mini-slots and T_{SIG} quantities are found from the EOS. Figure 5.19 illustrates a sample EOS for a set of

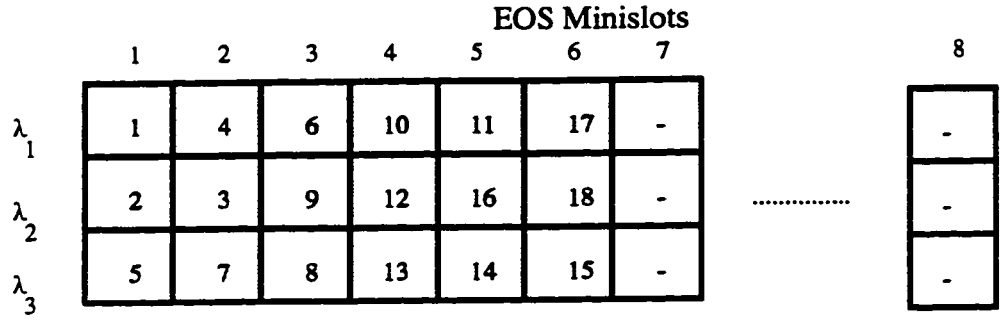


Figure 5.19: Sample EOS for a three channel system with 18 stations.

three channels and eighteen stations. First observe that $\max_{i,x} I_i(x)$ can easily be found from the EOS by noting that the last column of the EOS with a valid entry indicates the maximum number of stations on any channel. Thus the maximum inserted delay can be calculated as,

$$\max_{i,x} I_i(x) = maxcolnum \times AD \tag{5.26}$$

where AD is the delay that an active station introduces. Knowing this the minimum T_{SIG} required to satisfy the current station distribution can be written as

$$T_{SIG} = t_{re-tune} + maxcolnum \times AD. \tag{5.27}$$

The delay station x, which is currently tuned to channel y, must incur before its re-tuning activity may start is calculated as follows,

$$delay_x = \max_{i,j<x} num_i(j) - num_y(x), \tag{5.28}$$

where $num_i(j)$ is the number of stations on channel i prior to station j. The SOS for all channels may thus be calculated via the following algorithm.

1. for EOS element x,y ($EOS_{x,y}$)
   $STATION = EOS_{x,y}$
   $MAXDELAY$ = Maximum column number (j) over all $EOS_{i,j} < STATION$
   $STDELAY = j - 1$
   $SOS_{i,j} = MAXDELAY - STDELAY$

2. Repeat for each valid EOS entry.

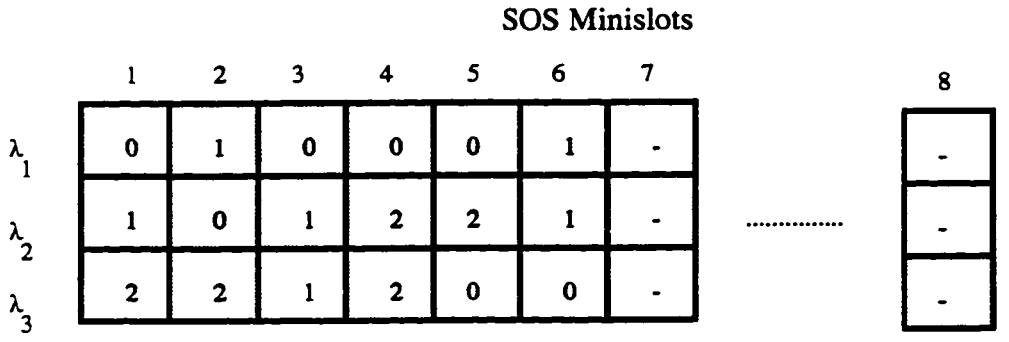


Figure 5.20: Sample SOS for a three channel system with 18 stations.

Figure 5.20 illustrates the resulting SOS generated after the application of the above algorithm to the EOS of Figure 5.19. In this example each station is assigned a integer number. This integer converted to a wait time within the station by multiplying the number by the delay an active station introduces to a channel (defined earlier as AD). Further reductions in the required length of a SIG could be achieved if the headend knew which stations planned to tune to and from each channel. It is felt, however, that the overhead required in collecting, processing, and distributing this information would be too large to make such a protocol feasible in a real network and is only discussed here as a possible option.

5.3.7 Home Channel Reconfiguration

Finally, it may occur that, as with the passive protocols, during normal network operation traffic loading may become skewed. That is, much more traffic may be generated from or destined to a subset of channels. Such occasions may arise when there is a concentration of servers or busy stations on one or more channels. To optimize network performance it is desirable to evenly distribute the load over all channels. This is possible through a reconfiguration phase in which the headend instructs user stations to change home channels. Such a command can easily be placed in the start of SIG or be a broadcast packet placed in a data slot. Since each station is only equipped with a slow tunable receiver this reconfiguration should be initiated only rarely. Once the commands to reconfigure have been issued, normal

network activity would have to be suspended until such time when the re-tuning at all user stations can be guaranteed complete.

5.4 Performance Analysis

In the following sections the performance of the ring, folded bus, and dual bus networks based upon the active coupling technologies discussed in this chapter is examined. Approaches will include both analytical and simulation techniques. Capacity and delay performance indices are derived for the various systems followed by an exploration of the optical link budgets for the systems. Results give an indication of the relative strengths and weaknesses of the designs.

5.4.1 Capacity

Maximum throughput or capacity can be interpreted as the total peak user data rates as a fraction of the total.

Active Analytical Models

In this section, we discuss the capacity performance of the active system protocols. First, we consider the conditional channel capacity, given that a fixed number of busy stations, N, are inserted into a wavelength and actively transmitting. T is defined as the slot transmission time. Overheads due to SIG operation are taken into account. However, when transmitter tuning across multiple channels is taken into account, exact capacity models are very complex. This thesis will present simulation results which show that under uniform loading with randomly chosen destinations, the conditional capacity is an excellent approximation to the system capacity when the number of stations is moderate to large. In addition, the conditional capacity represents certain types of traffic behavior where strong communities of interest form among the stations.

At first, we will ignore the overhead due to the SIG mechanism. Assume that the transmitting stations on a particular channel are numbered from 1 to N. In the

folded-bus and ring, each channel cycles through $\sum_{q=1}^{N} N_q$ packet transmissions before a global quota reset occurs. Also under *all* the protocols specified, an overhead of τ_L seconds is incurred before active transmission resumes. Here, $\tau_L = \tau + I_{MAX}$ is defined to be the total propagation latency of any of the channels, including insertion delays. Thus in the absence of SIGs, the per channel bus capacity is

$$C_{Fbus(NoSIG)} = \frac{\sum_{q=1}^{N} N_q}{\sum_{q=1}^{N} N_q + \tau_L/T}, \tag{5.29}$$

where T is the packet transmission time. The SIG mechanism reduces the capacity further because of additional overheads. As a result, a time period of only $t_{SIG} - T_{SIG}$ is available for transmission in each SIG cycle. The adjusted capacity is given by

$$C_{Fbus} = \frac{\sum_{q=1}^{N} N_q}{\sum_{q=1}^{N} N_q + \tau_L/T}(1 - \frac{T_{SIG}}{t_{SIG}}), \tag{5.30}$$

provided that $t_{SIG} - T_{SIG}$ is divisible by $\sum_{q=1}^{N} N_q T$. When this is not the case, then there may be some overlapping of transmission reset and SIG overheads. As a result, Equation 5.30 is an approximation to the per channel capacity.

In the dual-bus network a similar result can be obtained. In this case we find that

$$C_{Dbus} = \frac{N \cdot \text{BWB\_MOD}}{N \cdot \text{BWB\_MOD} + 1}(1 - \frac{T_{SIG}}{t_{SIG}}), \tag{5.31}$$

where BWB_MOD is the bandwidth balancing modulus and N is the number of stations per channel. Again, since we have ignored possible protocol and SIG interactions, this equation is used as an approximation to the real capacity.

In the ring case, the analysis is not so straightforward due to spatial re-use. However, using a result from an analysis of the Orwell ring [ZVN87], we can make a similar argument which is accurate when the station transmission quotas are large. Note that in Orwell a distributed resetting mechanism is used, whereas we use a centralized version. Following an argument similar to that above, the resulting capacity expression is given by

$$\bar{C}_{RING} = (2 - \frac{4}{M+2})(1 - \frac{T_{SIG}}{t_{SIG}}). \tag{5.32}$$

Simulation Environment

Since Equations 5.30, 5.31, and 5.32 should be viewed as approximations to the real system capacity, there should be a method by which we qualify those approximations. In this thesis computer programs which simulate each of the investigated systems were developed along with a library of discrete event programming tools. The concepts of discrete event simulations and the "c" code libraries are discussed in Appendix C. For the simulations of capacity, a constant backlog of packets are generated at each station. Destinations are picked at random and are evenly distributed among all possible destinations. The system is then run until such time that steady state performance is reached. Statistics are then collected over a period of time representing one million packet transmissions.

Results

In Figures 5.21 to 5.25, capacity results are shown for the active bus and ring networks using SIG flavor A. Comparisons are made with a variety of passive ring and bus designs. In particular, we compare the performance to the BTMSR and BTMSC systems.

In Figure 5.21, the per channel capacity is plotted versus the number of stations per channel. Included in this graph are the analytic results given by Equations 5.30 and 5.32. For comparison purposes, we have used a value of N_q which is high enough so that quota resetting consumes a negligible fraction of system capacity. In the higher capacity bus and ring curves, the maximum possible capacity is plotted by using a very large value of t_{SIG}. In the lower curves, we assumed that t_{SIG} was equal to 20 network propagation delay times. It can be seen that in all cases the analytic models give an excellent prediction of capacity when compared with simulation. For the parameters used, the graph shows the increase in capacity afforded by the ring where spatial re-use is possible. When the number of stations is too small, this advantage obviously disappears. In the curves we also assumed a network latency of 50 slots which includes the delay due to active station insertions. In the simulations, each slot took 1 μs to transmit. Assuming a 1 *Gbps* transmission rate, each slot would

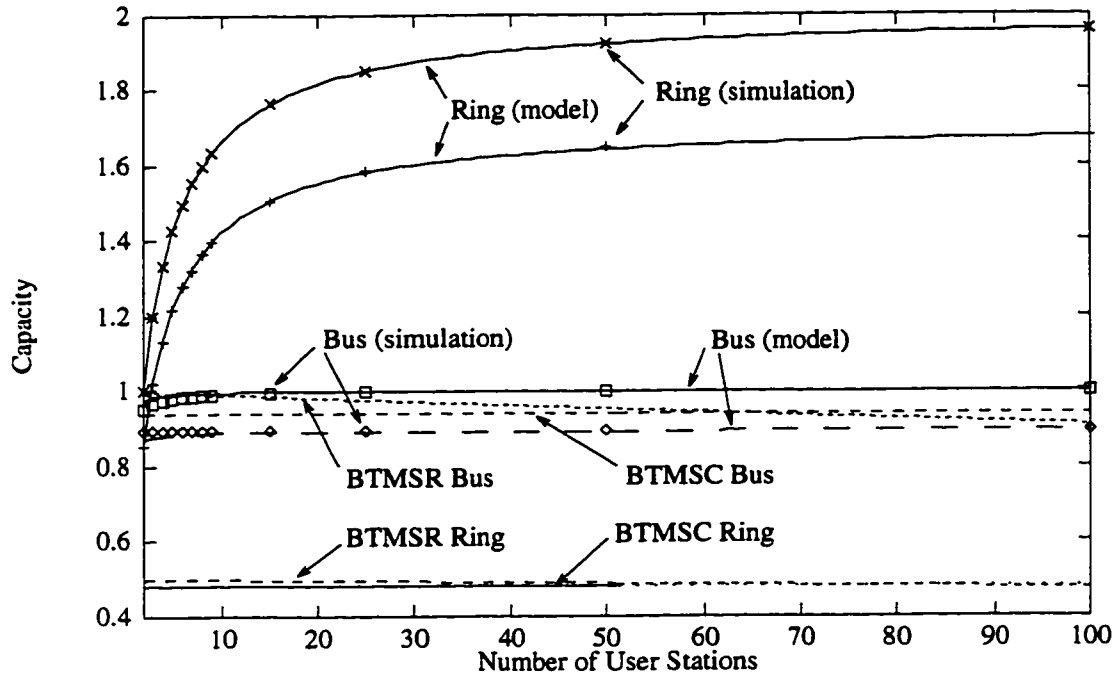


Figure 5.21: Capacity vs. Number of Stations for Active Ring and Folded Bus Systems Using SIG Flavor A.

be 1000 bits long and the end-to-end network length would be approximately 10 km. Furthermore, a coupler re-tuning time of 20 μs was used.

In Figure 5.22, capacity is plotted against the number of usable data slots between SIGs, N_{SIG}. For the bus, $N_{SIG} = t_{SIG} - 2(\tau + I_{MAX}) - t_{RE-TUNE}$ and is equal to $t_{SIG} - 4(\tau + I_{MAX}) - t_{RE-TUNE}$ for the ring when time is normalized to slot length. In this case, we assumed 25 stations per bus/ring. Again, the upper bus and ring curves corresponds to the maximum capacity possible and uses a large value of N_q. The other two assume that $N_q = 10$. In the active bus simulations, an interesting phenomenon can be seen. Notice the jagged form of the plot when $N_q = 10$. This behavior is an integer effect and is explained by noting that when SIGs and total transmission quotas are integer multiples, overheads between the two mechanisms may be shared. In this case, the "period" of the maxima is approximately equal to $\sum_{q=1}^{N} N_q T$, or 250 slots. Again it can be seen that the analytic models perform very well. It is also apparent that a fairly high value for the SIG period is required in order

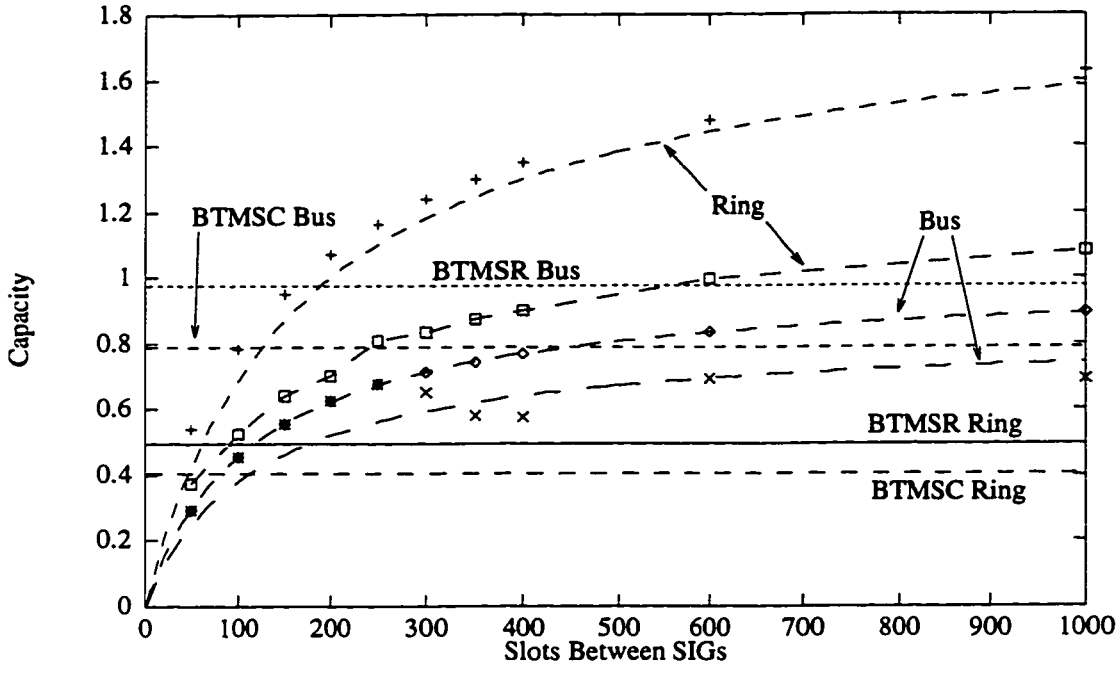


Figure 5.22: Capacity vs. Slots Between SIGs for Active Ring and Folded Bus Systems Using SIG Flavor A.

to ensure efficient operation.

In Figure 5.23, capacity is plotted versus the station transmission quotas using the same system parameters. It can be seen that similar behavior is obtained by all systems. The jaggedness of the BTMSC plots is due to the effect illustrated in Figure 4.16. In Figure 5.24 we also plot comparisons based on the number of slots on the network. This graph illustrates the eventual rolloff of capacity for the active systems using SIG overheads when the size of the network becomes large. Also we have included analytic results which are seen to agree well with the simulations. The parameters used are the same as those indicated for the previous curves. It can be concluded from these results that the active systems using SIG length protocol flavor A are not well suited for use in the metropolitan area due to this rolloff. Finally, in Figure 5.25 we show the degradation in capacity performance as the coupler tuning times are increased. Again, the parameters are the same as the two sets of curves discussed above.

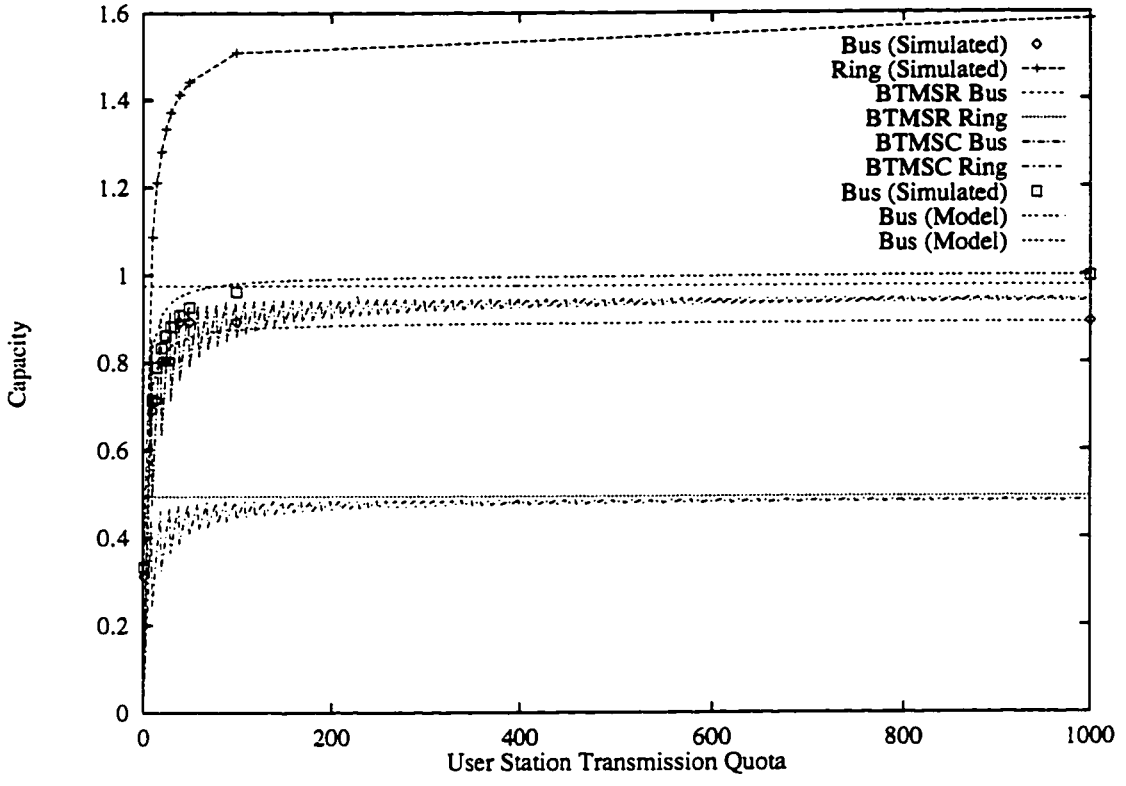


Figure 5.23: Capacity vs. Station Transmission Quota for Active Systems Using SIG Flavor A.

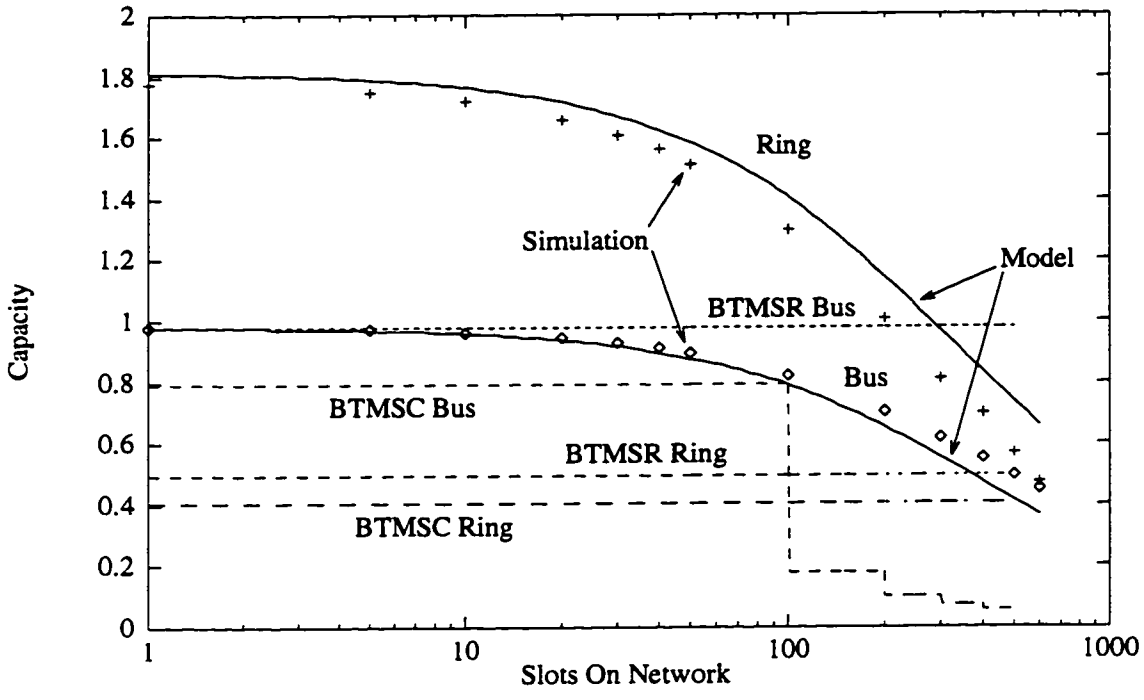


Figure 5.24: Capacity vs. Slots on Network for Active Systems Using SIG Flavor A.

In Figures 5.26 to 5.29, more capacity results are shown based on the derived approximations for the dual bus. In this case we use a total of 100 stations which are equally distributed across $C = 4$ channels ($N = 25$), while arbitrarily choosing one of those channels for collecting statistics. In the curves, each slot is $T = 1\mu sec$, $N_q = 100$, $t_{SIG} = 100\mu sec$, $\tau = 25\mu sec$, $t_{re-tune} = 20\mu sec$, and station insertion delays are 1/5 of a slot duration. Results from SIG length Flavors A, B, and C are all included and comparisons are made with a variety of other designs. In particular, we compare the performance to the BTMSR and BTMSC systems studied earlier.

In Figure 5.26, the per channel capacity is plotted versus the number of stations per channel (N). Included in this graph are the analytic results given by Equation 5.30 and 5.31. Here we assumed the total number of stations was 100 and that the number of channels varied accordingly. It can be seen that in all cases the analytic models give an excellent prediction of capacity when compared with simulation. The curves also show the advantage in capacity which is afforded by Protocols B and C, compared with the original, A. We also find that, when compared to an equivalent

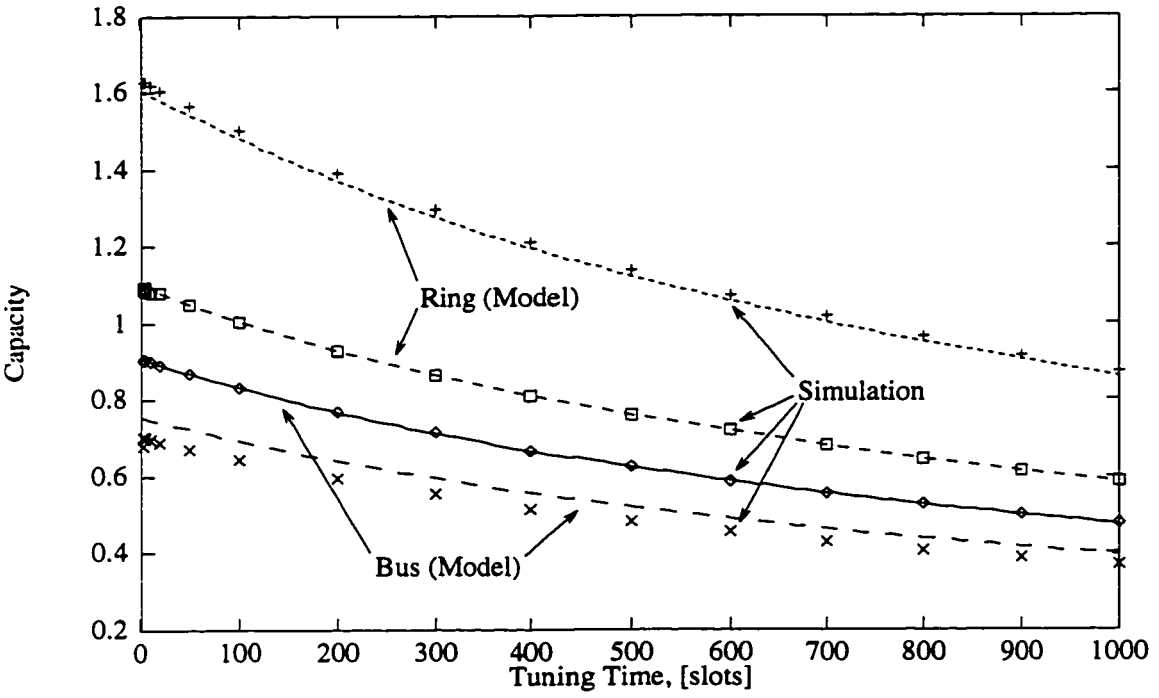


Figure 5.25: Capacity vs. Coupler Tuning Time for Active Systems Using SIG Flavor A.

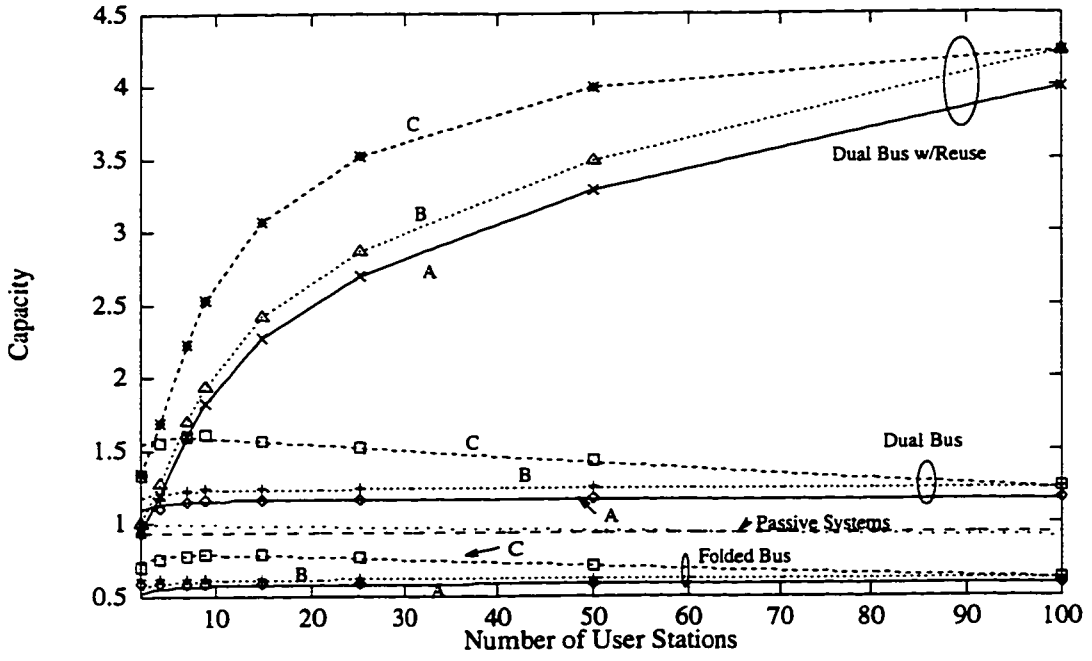


Figure 5.26: Capacity vs. Number of Stations per Channel (N) for Active Folded and Dual Bus Systems.

passively coupled architecture, the active coupled architectures suffer from reduced capacity due to the overheads associated with the SIGs.

In Figure 5.27, capacity is plotted against the number of usable data slots between SIGs. This is equal to $t_{SIG} - T_{SIG}$. Again, it can be seen that the analytic models perform very well. It is also apparent that a fairly high value for the SIG period is required in order to ensure efficient operation. However, the SIG period cannot be made too large, otherwise multi-channel dynamic operation is severely curtailed.

In Figure 5.28 we also plot comparisons based on the number of slots on the network. Increasing numbers of slots on the network corresponds to an increasing end to end propagation delay. Once again, we see that the analytic models give a very accurate indication of system capacity. It is important to note the severe capacity roll-off associated with protocol A, which limits its scalability to larger network sizes. This occurs in both the folded-bus and dual-bus versions. Protocols B and C however, are much more resistant to this effect, as seen in the figure. We also find that the passive coupled architectures show a nearly constant capacity across all network sizes

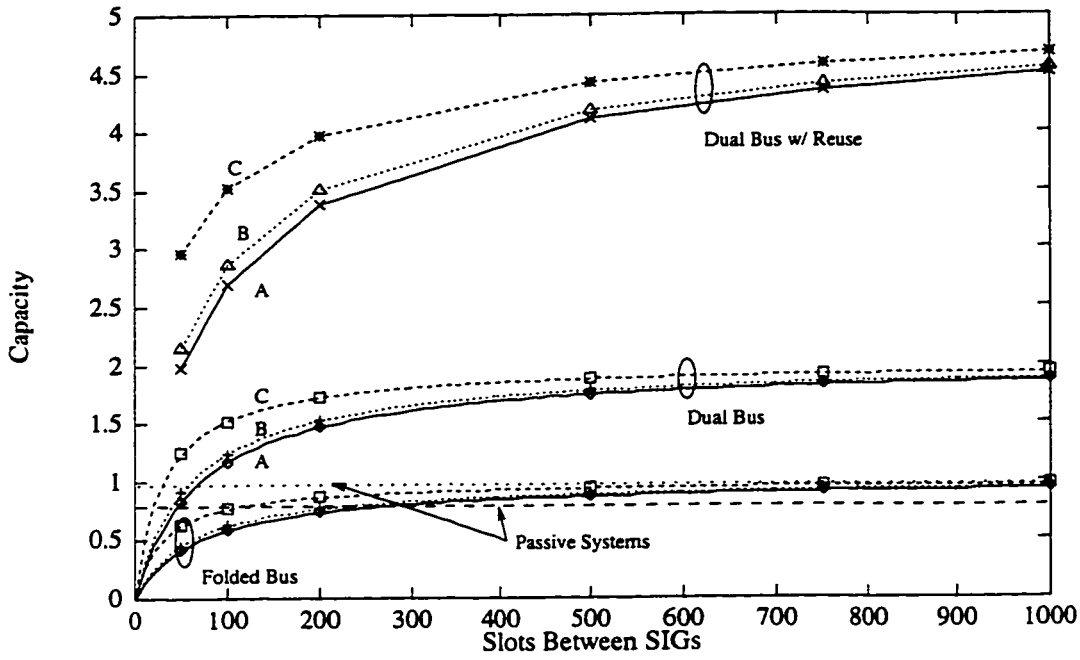


Figure 5.27: Capacity vs. Slots Between SIGs for Active Folded and Dual Bus Systems.

since no SIGs are required in those designs.

Finally, in Figure 5.29 we show the degradation in capacity performance as the coupler tuning times are increased. This effect is similar for all the systems studied. Strong agreement is again seen between the analytic and simulation models. As a matter of interest, we have included in Figures 5.26 through 5.29 the capacity of the dual bus protocol with destination release. Clearly, there is much to gain through the use of active coupling in cases where spatial reuse is possible.

Multi-Channel Results

When stations transmit to destinations randomly chosen across all channels, system capacity is reduced compared to that presented above. This is caused by the additional overhead incurred due to switching between queues at the station. However, one would expect that as the number of stations increases, this effect would be reduced. Figure 5.30 shows a sample of the relation between the conditional and

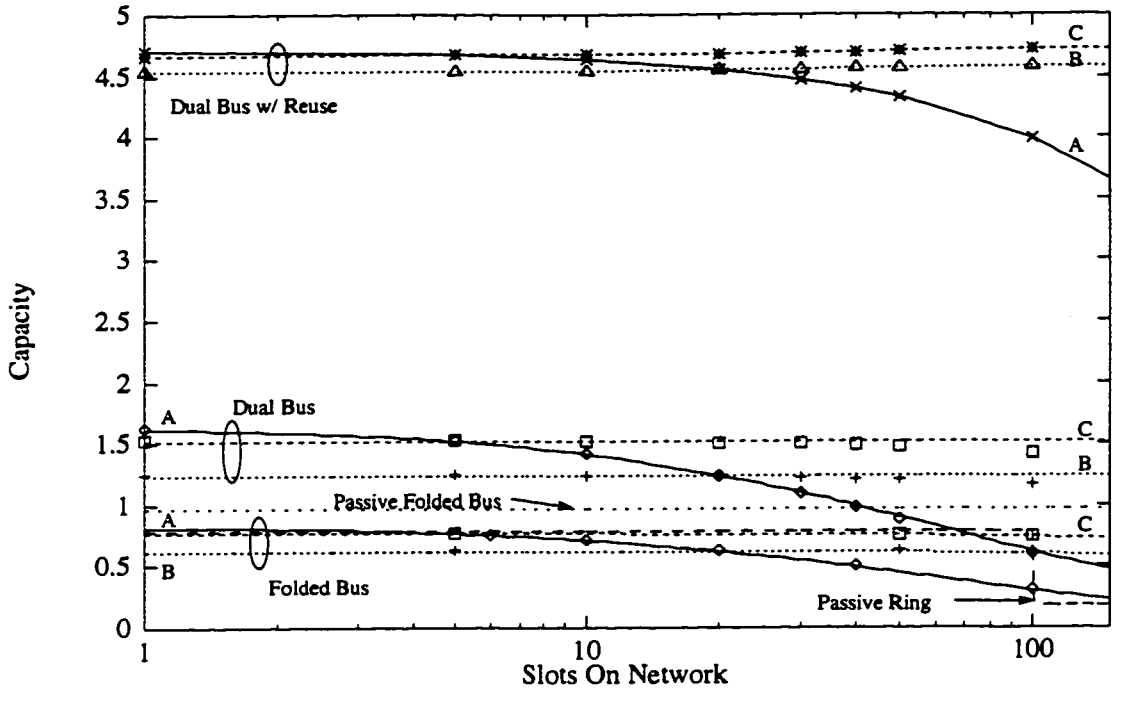


Figure 5.28: Capacity vs. Slots on Network for Active Folded and Dual Bus Systems.

simulated capacities for a 1, 4 and 6 channel system. In the simulations, each station keeps a separate queue for each destination channel. As shown, the capacity quickly approaches the conditional capacity as the number of stations increases. In the 4 channel system, the values are within about 5% when the number of stations per channel exceeds 10. These results suggest that conditional capacity can be used an an accurate approximation for uniform load capacity when the number of stations is moderate to large.

Results Comparisons

The results of this section make a few clear distinctions between the discussed systems. In passive systems, there is a hardware performance tradeoff implying in general that reducing the number of station couplers reduces system capacity. It was also discovered that active systems will typically have inferior capacity performance compared against passive systems of like topology. This effect is pronounced as networks become

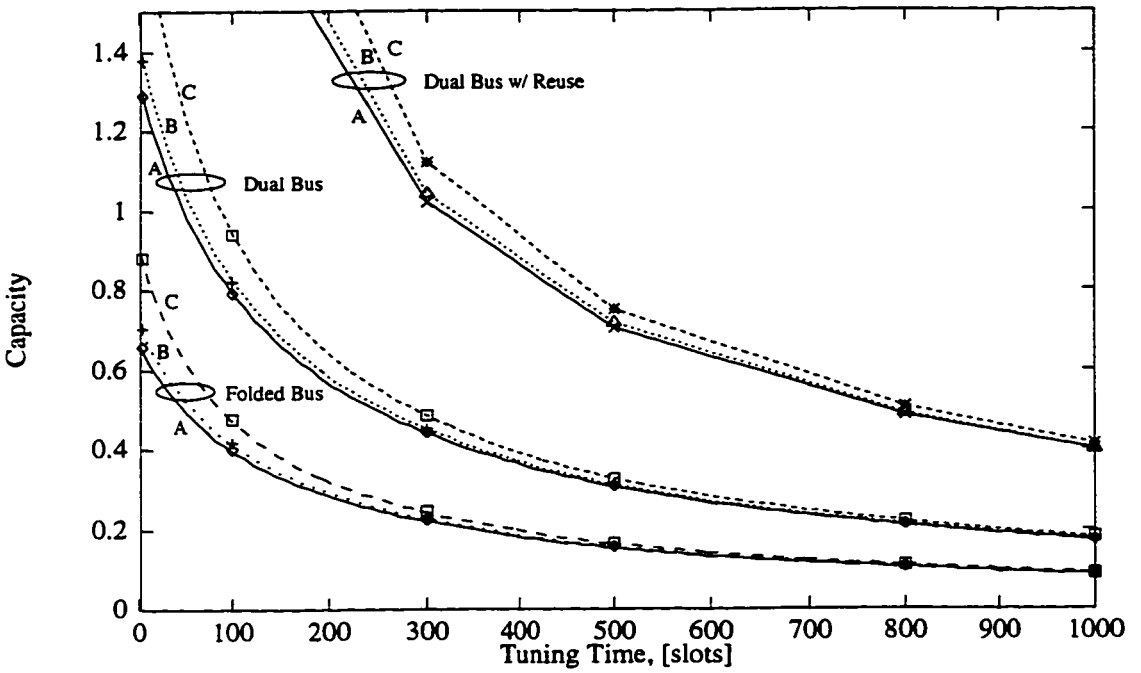


Figure 5.29: Capacity vs. Coupler Tuning Time for Active Folded and Dual Bus Systems.

larger because the overhead due to SIGs become an increasing factor to the active systems. As a result, careful determination of SIG parameters becomes important. Finally, the ability to utilize spatial re-use in active systems proves to be the most efficient way of increasing system capacity and provides a mothod by which active designs surpass the passive ones in capacity performance. For the active dual bus and ring systems this effect is dramatic while hardware requirements remain modest.

5.4.2 Delay

In this section we consider the mean delay experienced by packets generated at the stations. This is an average delay experienced by a packet between its generation at a source station and its reception at a destination. Both simulation and analytic models are given where possible. The objective is to obtain simple closed-form analytic results which can be used to validate the simulation results. Stations are assumed to be uniformly tuned to HOME channels across all C wavelengths. The number of

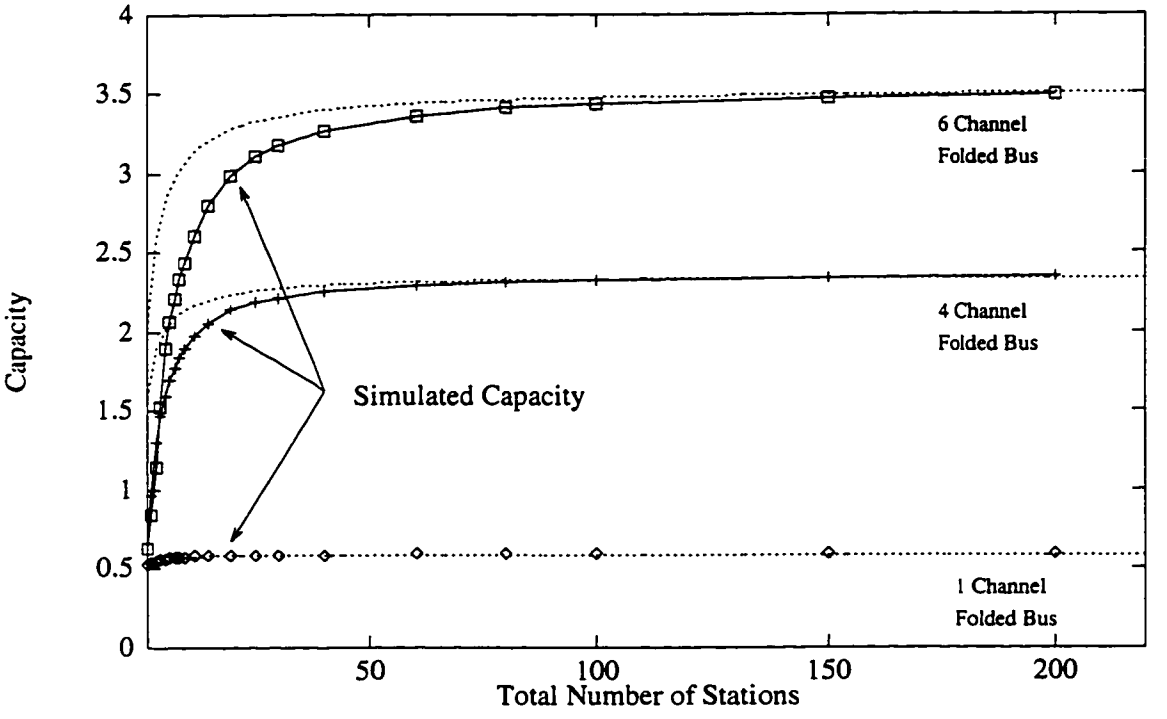


Figure 5.30: Multi-Channel Capacity vs. Total Number of Stations for the Active Folded Bus.

stations per channel is again defined to be N. Consider a particular queue at some average station i which is generating arrivals destined only to a particular channel. We consider the steady-state mean delay experienced by these packets. The buffer is assumed to be infinite and is fed by packet arrivals from a Poisson process at rate $\gamma_i = \gamma$. Packets are fixed in length and occupy one data slot each. Packet destinations are random and chosen independently and uniformly across those stations on the destination channel.

Come-Right-In Model

In the previous chapter an analytical model appropriate for systems in which packets experience gated frame access was presented. When packets are allowed to be serviced at random points within a frame, it is said that the system is a "come-right-in" one. This is the case for the active folded bus protocol.

Consider a queue at station i which is receiving arrivals destined to a specific channel. Packets are assumed to be fixed in length and occupy a single data slot. This system may be modeled by a single input queue and server as in [AFB79]. System time is slotted with a fixed number of data slots available for transmission per frame. Arrivals to the queue are taken to be Poisson with rate λ packets per slot. Departures are taken to occur at the beginning of a slot. Access to these slots follows a "Come Right In" strategy. This system can be viewed as an M/D/1 model with a regularly vacationing server. The average delay experienced by a packet is derived in Appendix B and for the folded bus systems is given by

$$\overline{D} = \frac{\sum_i^M D_i}{M}, \tag{5.33}$$

where D_i is the delay experienced by a packet that arrives during the ith slot interval and M is the total frame time including SIG overhead expressed in slots.

Approximate Analysis

The computation given in Appendix B is very complex. A much simplified approximate model for the folded bus system can also be formulated as follows. This model

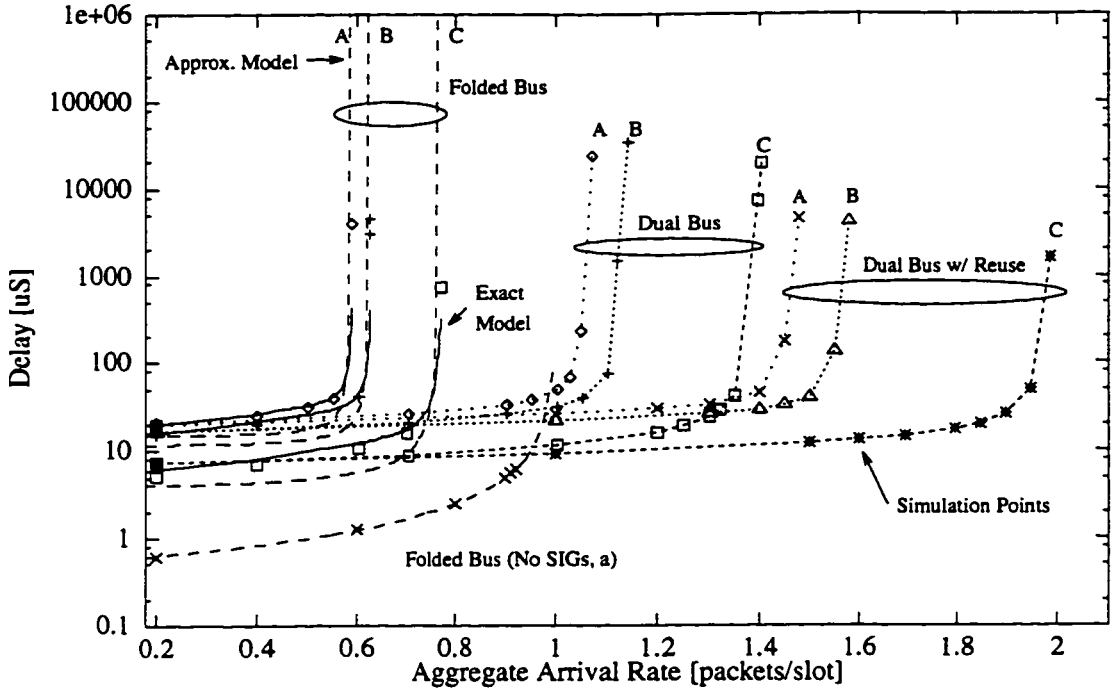


Figure 5.31: Mean Delay Model and Simulation for Active Folded and Dual Bus Systems.

is a version of the simple approximation suggested in [Hay84] for ungated TDMA systems. The idea is based on the observation that under ideal conditions, when SIG and propagation overheads are negligible, the network is again viewed as an M/D/1 queue. Overheads can be compensated for by scaling the service rate in accordance with the capacity of the system. Applying this approximation results in the average transfer delay

$$T_{Fbus} = \frac{(M\lambda_i/2)X_e}{1 - M\lambda_i X_e} + \frac{T_{SIG}^2}{2t_{SIG}} + \frac{1}{2}(1 - \frac{T_{SIG}}{t_{SIG}}) + 1, \tag{5.34}$$

where λ_i is the packet arrival rate per station and X_e is the service rate scale factor. In this case,

$$X_e = \frac{\sum_{q=1}^{M} N_q + \tau_L/T}{(1 - T_{SIG}/t_{SIG})\sum_{q=1}^{M} N_q}, \tag{5.35}$$

where τ_L is the end-to-end propagation delay.

Simulation Environment

The simulation results of this section were generated using methods similar to those used in Section 4.4.2. In this case, delay was measured for those stations with only a single channel of interest and with destinations chosen with equal probability among those stations on the channel.

Results

In Figure 5.31, comparisons are made for mean delay versus throughput (in packets/slot) for the folded and dual bus schemes. Also, the two analytic models are plotted for the folded-bus. The simulation results are shown as discrete points. It can be seen that the models perform very well compared with simulation results. In particular we found that the exact analytic model provides a much closer fit to the simulation points for low to medium loads. Both models accurately predict system capacity. Again, the improved capacity of SIG Protocol Flavors B and C are readily apparent from the graphs.

Multi-Channel Results

When stations generate packets destined randomly across all channels, the situation is much more complex and accurate analytic models are not available. In Figure 5.32, simulation results are shown which compare the protocols under these conditions. In the figure, the relative performance of the protocols is the same as that found above. However, increasing the number of channels causes the delay under light loading to increase as a result of channel queues being blocked while a particular queue is being serviced. This effect is clearly larger as the number of channels increases. This is well illustrated in the figure when we move from 1 channel to 4, 6 and 10 channel systems.

Priority Queue Model

Access to slots in the active ring system occurs quite differently than in the previous folded bus case. This is due to the use of spatial re-use and that access to a slot at a particular station is determined by all station on the network (upstream and down).

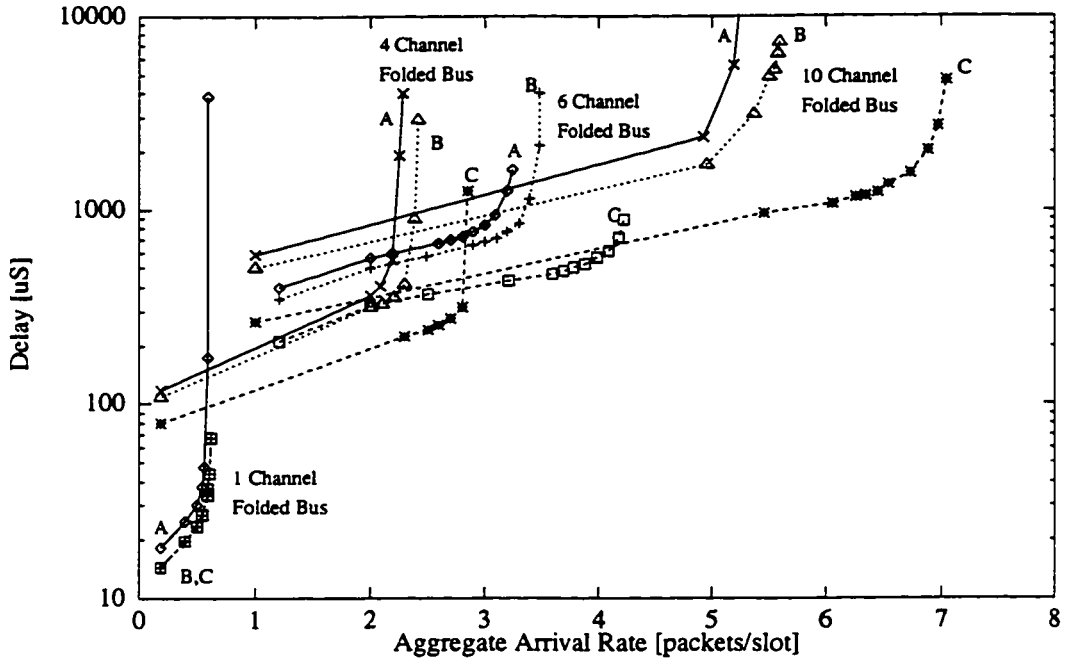


Figure 5.32: Mean Delay Simulations for the Active Folded Bus Network with 1, 4, 6 Wavelength Channels.

Consider a particular queue at some station i which is receiving arrivals destined to a particular channel. We are interested in the steady-state mean delay experienced by these packets. This buffer is assumed to be infinite and is fed by packet arrivals from a Poisson process at rate $\gamma_i = \gamma$. Packets are fixed in length and occupy one data slot. Packet destinations are random and chosen independently and uniformly across those stations on the destination channel.

The mean delay model in the active ring system is formulated as follows. The model is an extended version of that used in [ZVN87] and [Hay84]. Consider a particular WDM ring to be operating under steady-state conditions. In this case, a particular station i must see an idle slot in order to transmit onto the ring it is tuned to. The distribution of such slots is dependent upon transmissions that are made by other stations on the ring and subject to overheads associated with the SIGs. From station i's viewpoint we may thus view the ring as a non-preemptive priority queueing system with three classes of customers. This model is shown in Figure 5.33. The first class of customers correspond to the arrival of the SIG which occupies the

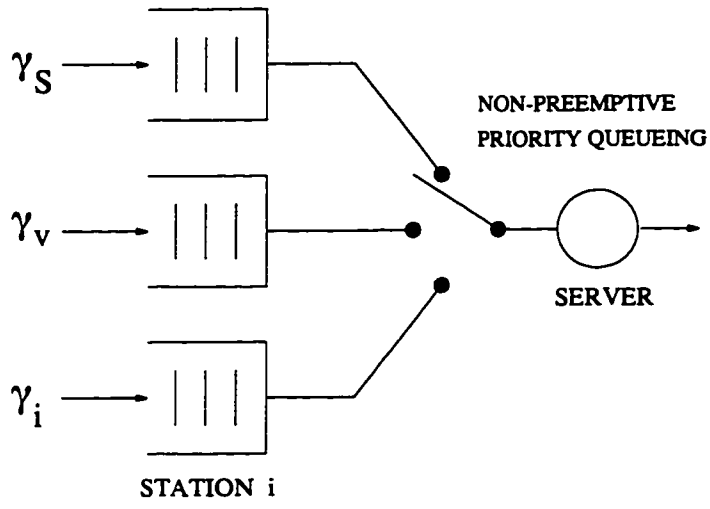


Figure 5.33: Active Ring Queueing Model

server during the re-tuning phases of operation. The mean arrival rate and service time of this class is defined to be γ_S and $\overline{X}_S$ respectively. The second customer class consists of those slots which are used by stations whose transmissions to others must pass station i on the ring. These are aggregated into a single queue fed by a Poisson process at a rate

$$\gamma_v = \sum_{k}^{N} \sum_{j}^{N} \theta_{kj}^{i}, \tag{5.36}$$

where θ_{kj}^{i} is the rate at which packets from station k must pass through station i to be received at station j. Because of the symmetry of the system, it can easily be shown that $\gamma_v = \gamma_i(N-1)/2$. The final traffic class is those customers corresponding to packets arriving to the queue at station i itself. We will assume that this is also a Poisson process with mean γ_i. Lowest priority is given to the traffic arriving to station i.

The delay analysis proceeds by first writing a set of state equations describing the evolution of the number of packets in each of the three queues at the end of each packet departure instant. In [Hay84], an example of this analysis is given and a generating function is derived. Since we are interested in a mean delay result, this can be obtained more directly by squaring the state equations and taking their

expectations. After some manipulation, the mean waiting time (in slots) for packets arriving to station i's queue is obtained as

$$W^i_{RING} = \frac{\gamma_S \overline{X_S^2} + \gamma_v + \gamma_i}{2(1-\gamma_S-\gamma_v)(1-\gamma_S \overline{X}_S - \gamma_v - \gamma_i)}. \tag{5.37}$$

In Equation 5.37 we have substituted for the fact that the mean service time for class i and v traffic is one slot. Since the SIG overhead is fixed, we can also write that $\overline{X_S^2} = \overline{X}_S^2 = (T_{SIG})^2$ slots and that $\gamma_S = 1/t_{SIG}$.

In the active bus case, a similar analysis may be used. In this case, a two-queue non-preemptive priority queueing model is employed. As before, the first queue models the effects of the SIGs and has arrival rate γ_S and service time X_{SB}. All data packet traffic arriving to the user stations is applied to the second queue with arrival rate γ_b. The resulting mean queueing delay is now given by

$$W^i_{BUS} = \frac{\gamma_S \overline{X_{SB}^2} + \gamma_b}{2(1-\gamma_S)(1-\gamma_S \overline{X}_{SB} - \gamma_b)}. \tag{5.38}$$

In this case, $\overline{X^2}_{SB} = \overline{X}^2_{SB} = T^2_{SIG}$. Also for the bus network, a very good general approximation results by assuming that overheads due to transmission quota resetting are accounted for by multiplying the service time of the stations by the factor

$$\frac{\sum_{q=1}^{M} N_q + \tau_L/T}{\sum_{q=1}^{M} N_q}. \tag{5.39}$$

Results

In Figure 5.34, comparisons are made for mean delay versus throughput (in packets/slot) for the active ring and folded bus using SIG length protocol Flavor A. Also, the analytic approximations are plotted. The simulation results are shown as discrete points. It can be seen that the models perform very well compared with simulation results. In the graphs, three plots are given of the active folded bus and ring. In order from highest to lowest capacity, we assume large values for both t_{SIG} and N_q, a large value for N_q and a t_{SIG} of 20 end-to-end propagation times, and $N_q = 10$ with $t_{SIG} = 20$ propagation times. In the latter case, a model approximation for the

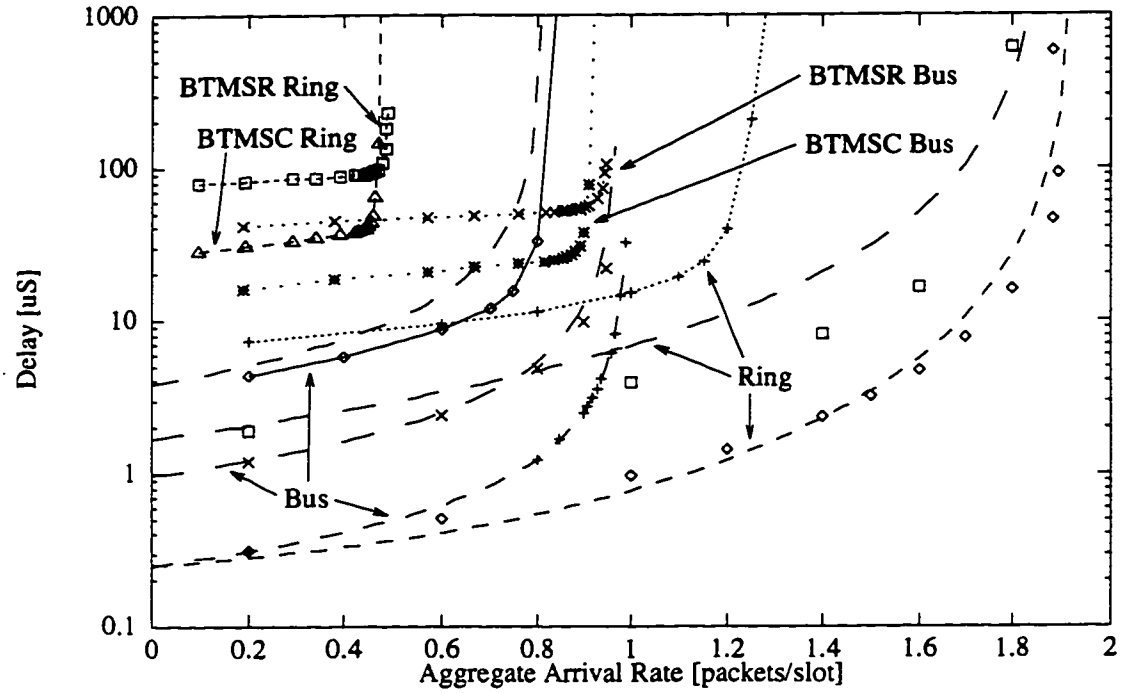


Figure 5.34: Mean Delay Models and Simulation for the Active Ring System.

ring was not formulated. Again, the improved capacity of the active ring is readily apparent from the graphs. Comparisons are also made with passive systems. Notice the better low load delays and higher capacity exhibited by the active systems.

Results Comparisons

The results of this section highlight the benefits of active station designs over passive designs other than in capacity terms. That is, at low loads the more dynamic access nature available using active coupling can greatly reduce delay. The results also indicate the relative systems capacity performance. This shows once again that systems employing spatial re-use such as the active ring and the active dual bus are economically effective approaches to increased system capacity.

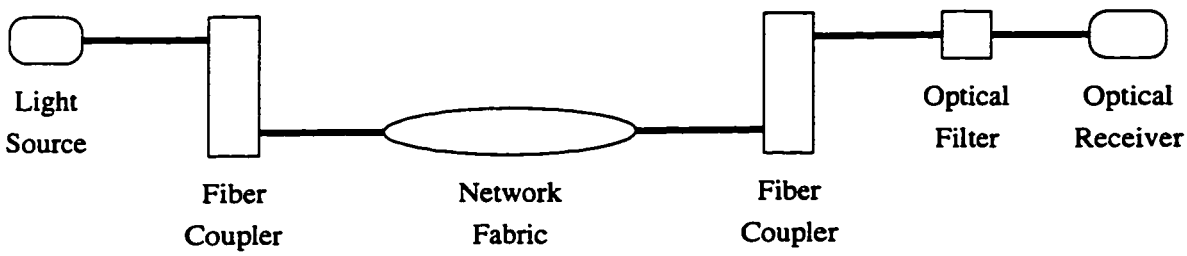


Figure 5.35: Components of the link-budget.

5.4.3 Power Budget

A network's optical power or link budget generally refers to the number of user stations that can be placed on the network so that there is still enough optical energy to drive every receiver in the system. In principle this easily calculated by summing the optical losses through a path from the first transmitter to the final receiver on the network. Figure 5.35 shows a block diagram of a path from optical source to receiver including various components which affect the link-budget calculation.

In this section, some calculations are performed which indicate the loss characteristics that the active couplers must have in order to give similar power budget performance as in a passively coupled system. To get a rough idea, we will assume that the total power loss (in dB) at a station due to a single active coupler is I_A. A family of curves will be presented which considers a range of I_A values. These will be compared against a similar power budget comparison for a single FBT coupled ring and a dual FBT folded bus. The analysis assumes the following quantities are defined in decibels. P_{TX} is the optical power transmitted, P_{RX} is the optical power required at a receiver for detection at an acceptable bit-error rate, I_{fbt} is the insertion loss of each FBT coupler, I_{fiber} is the fiber loss, I_{HE} is the splitting loss at the headend, I_s is loss due to a fiber splice and I_{filt} is loss due to filtering at the receiver. In the passive systems, the coupling coefficient of the FBT taps is the same for all couplers and optimized to give the best performance. The results will thus give an idea of the best performance possible. In practice a specific FBT coupler design would have to be fixed beforehand.

In the active bus case, under worst-case situations, signals in the folded bus configuration may have to travel through $2N$ active couplers before reception. The power available at the input of the headend receiver, P_{RX}, is thus given by

$$P_{RX} = P_{TX} - (2N+2)I_A - (4N+2)I_s - I_{fiber} - I_{HE}. \tag{5.40}$$

The optical receiver sensitivity, P_{RX} is found from [Gre93]

$$P_{RX} = 10\log(\frac{hcER}{\lambda} \cdot 10^3), \tag{5.41}$$

where h is Planck's constant, c is the speed of light in free space, R is the bit rate, E is the receiver sensitivity in photons per bit and λ is the received wavelength. As shown earlier, by equating Equations 5.40 and 5.41 we can find an upper bound on the maximum bit rate at which the network can be operated,

$$R_{bus} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}-(2N+2)I_A+(4N+2)I_s-I_{fiber}-I_{HE}}{10})}. \tag{5.42}$$

Similarly, for the dual bus configuration,

$$R_{bus} = \frac{\lambda}{hcE10^3} 10^{(\frac{P_{TX}-(N+2)I_A+(2N+2)I_s-I_{fiber}-I_{HE}}{10})}. \tag{5.43}$$

Figure 5.36 shows the network bit rates given by the above calculations for a 10 channel system characterized by a particular I_{HE}. Five different values of I_A are shown for the active bus and are indicated in parentheses. We have also included comparisons of the best that the passively coupled networks can do using a similar calculation. The values of λ, P_{TX}, and E were 1.5μm, 1mW, and 750 photons/bit as assumed in [IK92].

Immediately evident from Figure 5.36 is the benefit, other than reduced user station cost, of reducing the loss of the active station taps. Also it can be seen that very low values of coupler loss would be required to achieve a reasonable number of stations and performance similar to the passive case. When all losses are considered, current acousto-optic devices are unable to match this performance [Che90]. In the

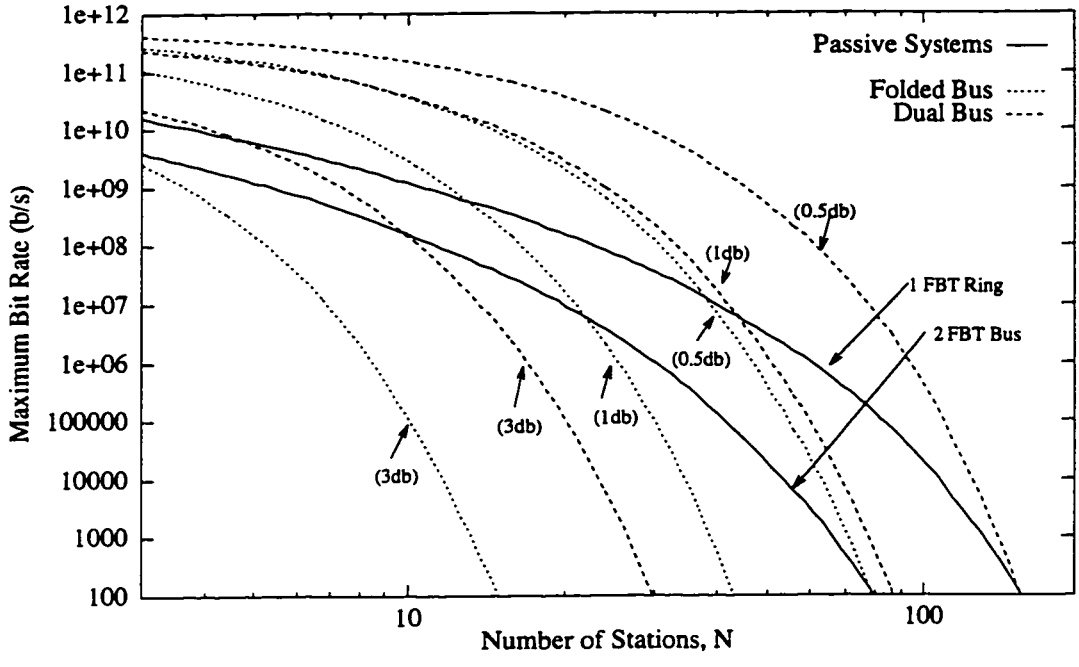


Figure 5.36: Optical Power Budget for the Active Systems.

near term, wideband optical amplification would clearly be required in both the passive and active systems. As a result, active coupling should be viewed as a method of simplifying system protocols to a level comparable with single channel systems.

5.5 Discussion

Three different station designs utilizing active couplers were presented in this chapter. The incentive of these designs was to explore the implications of the use of the active couplers within a station. A main consequence is that a greater number of HOME channel transceivers are required at each station because signals have to be retransmitted when they are received at an active interface. Furthermore, it was shown that care must be taken when the active couplers are tuned in order to avoid re-tuning collisions. However, because of the active station interface, signals are regenerated at each station making synchronization easier. The architecture of the headend was also discussed.

The protocols required for networks employing actively coupled stations were also are discussed. Although some special care must be taken in the protocols to avoid re-tuning collisions. A number of advantages also exist with the designs which may be exploited by way of the media access protocols. When compared with passive designs, the active networks are advantageous in terms of their media access simplicity and ability to implement spatial reuse where appropriate. However, it was shown that a significant amount of complexity can be invested in the SIG interval calculations when optimal performance is desired. In order to simplify system operation, several upper bounds for SIG lengths were derived requiring varying amounts of headend processing. Regardless, these SIG methods make it possible to export some very popular single channel media access protocols to WDM.

This chapter also presented several analytical delay and capacity models for the active network designs and protocols. Performance was examined under a number of conditions including the effect of network size on performance. It was found that in comparisons, performance of passive systems is degraded as the number of users or channels is increased. In active systems, the number of users is not a significant factor, however due to the overheads associated with SIGs, network span can be a limiting factor in terms of capacity performance. Delay result comparisions indicate that the passive system protocols suffer from a increased low load delay due to the gated nature of the media access protocol. This effect is not apparant in a single channel active system. However, the number of channels does affect low load delay. Finally, power budgets were presented for the designs and comparisons indicate that passive coupled networks still require fewer fiber amplifiers for an equivalent network size.

Chapter 6

Conclusions

The work described in this thesis studies the design and analysis of multi-channel photonic ring and bus networks. Several designs were proposed based upon two coupling methods; active and passive. In each case, dynamic, packet switched operation is achieved via appropriate media access protocols. This work has been published in part in [GT95, GT93, GT94, TGB95, GT96b], and has been submitted for publication in [GT96a]. In each design proposal, the goal was to minimize the user station hardware requirements and study the effect of this minimization on the required media access protocol and on system performance. The major contribution of this thesis is the introduction of active wavelength-selective couplers into the design of user stations for packet-switched local and metropolitan area networks. Methods which allow the use of traditional single channel protocols in a multi-channel environment were developed for use with the active couplers. In passively coupled designs, it was shown that the required number of couplers can be reduced to the minimum of one per user station assuming the use of an appropriate media access protocol.

In particular, several designs and protocols were proposed and studied. Passive systems based upon the folded bus and ring topologies were examined based upon either the BTMSC or BTMSR media access protocols. In the first case, user stations require at least two couplers in order to reserve data slots via a mini-slot mechanism. In the later, the minimum number of couplers is reduced to one via the use of a request/confirmation media access protocol. Active systems, possibly based on the

use of acousto-optic devices, were proposed for the folded bus, dual bus, and ring topologies. It was shown that if the tuning of these active devices occurred at an inappropriate time, packets in transit through the coupler would be destroyed. To avoid this "re-tuning collision" a number of Station Insertion Gap (SIG) flavors were proposed, each with different amounts of processing requirements. The use of the SIG mechanism then allowed the export of some common single channel protocols, such as DQDB, to a multi-channel WDM environment. Several of the designs were able to implement spatial re-use in the protocol.

Performance analysis of capacity and delay show that the performance of passive systems is degraded as the number of users or channels is increased. In active systems, the number of users is not a significant factor, however due to the overheads associated with SIGs, network span can be a limiting factor in terms of capacity performance. The method in which SIG lengths are determined does contribute greatly to the scalability of such systems. Delay results indicate that the passive system protocols suffer from a increased low load delay due to the gated nature of the media access protocol. Finally, power budgets analysis indicate that passive coupled networks still require fewer fiber amplifiers than actively coupled networks for an equivalent network size.

From this work, it can be concluded in general that there is a significant performance trade-off associated with the reduction of station hardware in passive systems due to the applicability of various media access methods. In the designs using BTMSC, it was shown that performance is superior to the BTMSR designs when there are only a handful of channels and large number of user stations. However, in cases where there are a large number of channels and a small number of user stations on each channel, the BTMSR systems are more appropriate. Systems based upon the use of active couplers, where and when available, present the ability to export some common single channel protocols. Performance is typically comparable with the original single channel protocol version except for the overheads dedicated to the SIGs. The potential of exploiting spatial re-use in these designs greatly increases the available capacity and thus the designs' attractiveness.

Much of the technology proposed for use in this thesis is still in an early stage

of development. As a result, this thesis does not pretend to present a current, economically viable solution for LANs and MANs. However, it is hoped that studies such as this will provide some justification for other researchers to keep investigating interesting photonic devices such as active couplers.

6.1 Future Work

In the process of this research a number of areas which have received inadequate discussion and treatment either here or in the literature have been identified. These areas may be the subject of future research.

Home Channel Assignment

In this thesis, a simple round robin channel assignment policy was used. This occurs without regard to the traffic distributions on the network. It may be possible to create algorithms which monitor traffic patterns and match the home channel assignments in an appropriate fashion.

Bandwidth Block Allocation

Since a uniform traffic model was assumed, the assignment of block sizes in the passive system protocols was very straight forward and fixed. However, for dynamic traffic patterns a more optimal technique would be required. Studies of algorithms that partition the inter-channel bandwidth could be undertaken. The optimal allocations and appropriate heuristics would be studied.

Protocol Fairness

Although fairness in the passive designs was obvious, it is not so obvious in the active designs. Fairness is a fairly hot topic of debate for DQDB networks in particular. Determining how fairness works out in the multi-channel versions of the protocol would be a worthwhile goal.

Deluxe SIG Flavor Exploration

A deluxe flavor SIG length estimation method was discussed briefly. Although complicated, its investigation in terms of performance might prove worthwhile.

Hybrid Passive/Active Designs

It was shown that the active station designs require a higher level of fiber amplification to be incorporated into the design than in an equivalent passive design. It was suggest in this thesis that a hybrid architecture which used both passive and active couplers could be built with practically the same performance as a fully active station. Further examination of such designs could be performed.

Improved Traffic Models

Better traffic models include the use of a client-server model where traffic loads are focused on specific stations and or channels. Currently, results are based on a uniform load model with Poisson arrivals. Possibly even more realistic results are available through the use of a self-similar traffic model [AZN95, BBC+96]. Finally, traces of real network traffic could be obtained and used to drive the discrete event simulation arrival process. This may allow the tuning of network parameters in cases which are not evident with the simplifying arrival assumptions.

Network Prototypes

A particularly interesting network could be construct based upon either the active or passive station designs. In either case, only a handful of stations should be required to gain valuable insights in terms of performance and construction techniques.

Appendix A

Gated Access Delay

The model consists of a single input queue and a single server as discussed in [JT94, Hay84]. System time is slotted into frames with a number of data slots per frame. Arrivals to the queue are Poisson with rate λ customers per frame, while departures are deterministic at a rate of b customers per frame. Departures occur at the beginning of a frame. The objective is to calculate mathematically the average delay experienced by a customer. The following symbols are used for the analysis.

n_i = Number of enqueued customers at end of frame i.

b = Number of customers removed per frame.

P_j = Probability of j customers enqueued at end of frame.

a_i = Number of customers arriving during frame i.

The analysis begins by embedding a Markov chain at the end of each frame. The state equation for the number of customers in the queue can be written as

$$n_{i+1} = max(n_i - b, 0) + a_{i+1}. \tag{A.1}$$

The probability generating function $N(z)$ of this equation can be written as

$$N(z) = E[z^{n_{i+1}}] = E[z^{max(n_i-b,0)}]E[z^{a_{i+1}}],$$

$$= \left[\sum_{i=1}^{b-1} P_i z^b - \sum_{i=1}^{b-1} P_i z^i - P_0(1-z^b)\right] e^{-\lambda(1-z)},$$
$$= \sum_{i=0}^{b-1} P_i (z^b - z^i) \frac{e^{-\lambda(1-z)}}{z^b - e^{-\lambda(1-z)}}. \quad \text{(A.2)}$$

Rouché's theorem, as described in [Hay84] can be used to solve for $P_i, i \in 0 \rightarrow (b-1)$ by forming a linear set of equations which can be written in matrix form as

$$\begin{bmatrix} z_1^b - 1 & z_1^b - z_1 & \cdots & z_1^b - z_1^{b-1} \\ z_2^b - 1 & z_2^b - z_2 & \cdots & z_2^b - z_2^{b-1} \\ \vdots & \vdots & & \vdots \\ z_{b-1}^b - 1 & z_{b-1}^b - z_{b-1} & \cdots & z_{b-1}^b - z_{b-1}^{b-1} \\ 1 & \frac{b-1}{b} & \cdots & \frac{b-(b-1)}{b} \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_b \\ P_{b-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \quad \text{(A.3)}$$

Where $z_0 \rightarrow z_{b-1}$ are the b roots of the denominator of Equation A.2 with z_0 set equal to the root at 1. With these known values it is possible to calculate any other P_i value by taking the Inverse Fourier Transform of $N(z)$. This can be calculated numerically, as described in [Hay84] by evaluating $N(z)$ at a sufficiently large number of points, K around the unit circle of the complex Z-plane,

$$P_i = \frac{1}{K} \sum_{k=0}^{K-1} N(e^{\frac{-j2\pi k}{K}}) e^{\frac{j2\pi ik}{K}}. \quad \text{(A.4)}$$

Using a similar analysis as [TB92], we derive the delay experienced by a "tagged" customer. The derivation is formulated as shown in Figure A.1.

Define n_0 as the number of customers enqueued at the start of an arbitrary frame. At some time τ after the start of the frame our tagged customer arrives. At the start of the next frame the number of messages enqueued in front of the tagged customer, including the tagged customer, is equal to the number carried over from the previous frame, n_{01} and the number of new customers, a_τ arriving in the interval τ,

$$n_1 = n_{01} + a_\tau + 1. \quad \text{(A.5)}$$

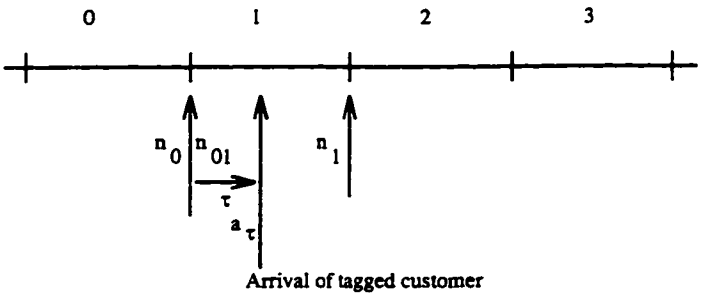


Figure A.1: Delay Model

If we define m as the number of frames required to service n_1 customers then $E(m)$ is the average number of frames. To determine this we need to know $P_m(i)$ the PDF of m. First consider $P_m(i)$ conditioned on n_1 written as $P_m(i|n_1)$,

$$\begin{aligned} P_m(0|n_1) &= \begin{cases} 1, & n_1 = 0 \\ 0, & n_1 > 0 \end{cases} \\ P_m(1|n_1) &= \begin{cases} 1, & 0 < n_1 \leq b \\ 0, & n_1 > b \end{cases} \\ &\vdots \\ P_m(i|n_1) &= \begin{cases} 1, & (i-1)b < n_1 \leq ib \\ 0, & n_1 > ib \end{cases} . \end{aligned} \tag{A.6}$$

After removing the conditioning,

$$\begin{aligned} P_m(i) &= \sum_{j=0}^{\infty} P_m(i|j) P_{n_1}(j), \\ &= \sum_{j=(i-1)b-1}^{ib} P_{n_1}(j). \end{aligned} \tag{A.7}$$

Now we need to determine P_{n_1}. From Equation A.5 we see that n_1 is a function of two independent random variables, therefore the PDF of n_1 can be written as a convolution of the PDFs of these variables:

$$P_{n_1}(j) = \begin{cases} \sum_{i=0}^{j-1} P_{n_{01}}(i) P_{a_\tau}(j-i-1) & , j > 0 \\ 0 & , j = 0 \end{cases} . \quad (A.8)$$

To calculate this we need to determine $P_{a_\tau}(i)$, the probability of i arrivals in the interval before the arrival of the tagged customer, and $P_{n_{01}}(i)$, the probability of i customers remaining in the queue at the beginning of the frame, but after b customers have been removed. $P_{a_\tau}(i)$ is found by first conditioning on τ where

$$P_{a_\tau}(i|\tau) = \frac{(\lambda\tau)^i e^{-\lambda\tau}}{i!}. \quad (A.9)$$

After removing the condition and considerable manipulation, as in [TB92] we arrive at

$$P_{a_\tau}(i) = \frac{1}{\lambda} - \frac{e^{-\lambda}\lambda^{i-1}}{i!}[1 + \sum_{j=1}^{i} \frac{i!}{(i-j)!\lambda^j}]. \quad (A.10)$$

$P_{n_{01}}$ is found by considering that $n_{01} = max(n_0 - b, 0)$ which has a PDF given by

$$P_{n_{01}}(i) = \begin{cases} \sum_{j=0}^{b} P_{n_0}(j) & , i = 0 \\ P_{n_0}(b+i) & , i > 0 \end{cases} . \quad (A.11)$$

And since $P_{n_0}(i)$ is equal to P_i of Equation A.4 we can determine $E(m)$ as

$$\overline{m} = \sum_{i=0}^{\infty} i P_m(i). \quad (A.12)$$

Finally the delay may be written as

$$\overline{D}_{BTMSC} = \frac{T_f}{2} + T_f \times (\overline{m} - 1) + \frac{T_f}{\alpha}. \quad (A.13)$$

where T_f is the length of a frame. The first factor of $\frac{T_f}{2}$ is a result of the assumption that arrivals are uniformly distributed across the frame. Thus packets arrive, on average, in the middle of any given frame. Similarly, the last factor of $\frac{T_f}{\alpha}$ is a product

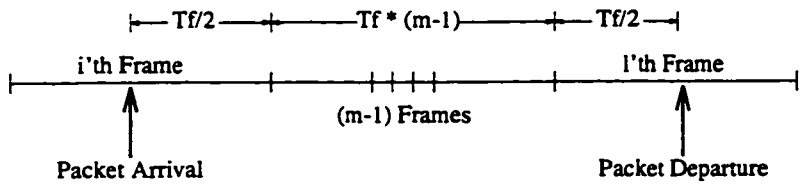


Figure A.2: BTMSC Queuing Delay Model

of an assumption regarding where packets are removed from the queue. For low loads, it is reasonable to approximate the departure of packets close to the start of the frame. Thus the variable α becomes rather large. For high loads, a faithful approximation is to say that packets are removed from the queue in the middle of the frame such that $\alpha = 2$ as illustrated in Figure A.2.

To best account for all situations, it is then reasonable to assume that packets are removed from the queue at $\frac{T_f}{4}$. Considering this, average delay can be written as

$$\overline{D}_{BTMSC} = \frac{T_f}{2} + T_f \times (\overline{m} - 1) + \frac{T_f}{4}, \tag{A.14}$$

and the average BTMSR delay can be written as

$$\overline{D}_{BTMSR} = \frac{T_f}{2} + \lceil \frac{\tau}{T_f} \rceil \cdot T_f + T_f \cdot (\overline{m} - 1) + \frac{T_f}{4}. \tag{A.15}$$

Appendix B

Come-Right-In Delay

A mapping of this model to the active folded bus protocol is based upon the observation that the protocol uses a frame structure during which K contiguous slots are used for packet transmission, while the remaining slots in the frame are unused to facilitate coupler re-tuning. In the model, we assume a uniform traffic case and that the overhead due to station transmission quotas is negligible. It is also assumed that packets are serviced using a FIFO discipline.

A station which has data for a particular destination channel is lumped with traffic due to other stations on that channel for an aggregate arrival rate, λ. This is illustrated in Figure B.1 where a single queue is fed at a rate λ and serviced by a single server.[1] The frame contains a total of M slots with J of those slots reserved as overhead. The calculated delay shall be valid for a packet generated by a station who has only a single channel of interest. Delay of packets generated at stations with multiple channels of interest will be larger, and dependent on the channel hopping policy used at those stations. For example, at low loads a station with multiple channels of interest, using a policy in which the queue containing the oldest packet is selected, will experience a delay of approximately $M/2$ slot times. On the other hand, a station accessing only a single channel will experience a delay of $\frac{1}{2}\frac{K}{M} + \frac{J}{2}\frac{J}{M}$ at low loads.

[1] A more general treatment of this problem is found in [AFB79].

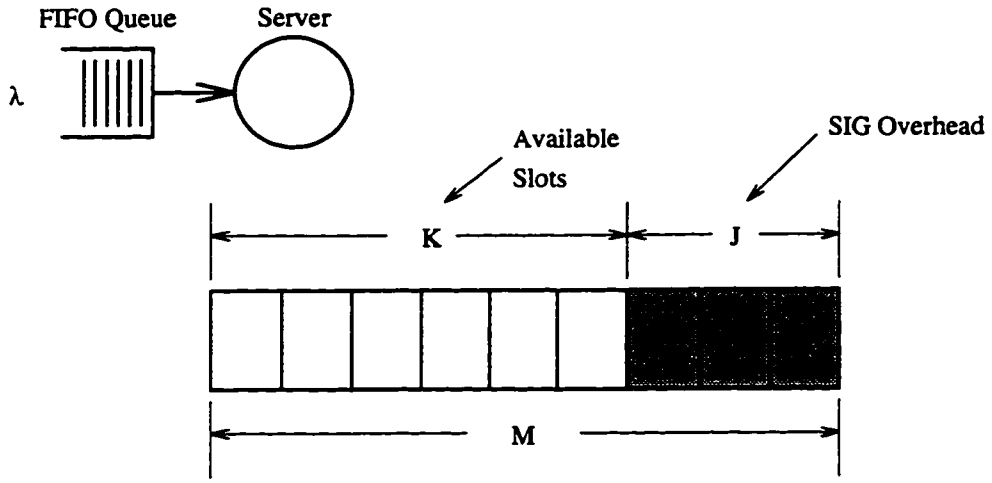


Figure B.1: The "come-right-in" model. M slots in a frame with J of those slots frame overhead.

B.1 Customers in the queue

The following symbols are used for the analysis.

n_i = Number of enqueued packets at the beginning of slot i.

χ_i = Number of arrivals generated in the ith unit of time.

M = Frame length in slots.

J = Number of overhead slots.

The analysis begins by embedding a Markov chain at the beginning of each slot. The state equation for the number of packets in the queue can be written as

$$n_{i+1} = \max(n_i - u_i, 0) + \chi_i, \tag{B.1}$$

where,

$$u_i = \begin{cases} 1, & \text{if slot } i \text{ is available for transmission,} \\ 0, & \text{otherwise.} \end{cases}$$

For simple Poisson arrivals, the generating function of χ_i is written as

$$\chi(z) = e^{-\lambda(1-z)}. \tag{B.2}$$

It is clear that u_i varies in time. As a result $\{n_i\}$ itself has no stationary distribution. Intuitively, this can be realized by visualizing the queue grow in size during the time in which slots are consumed as overhead, and shrink during the time in which slots are available for transmission. The fourth sub-plot of Figure B.2 illustrates this effect by plotting the average queue size at each time slot. It can be shown, however, that

$$n_m = \{n_{mM+1}, n_{mM+2}, \cdots n_{(m+1)M}\}, \tag{B.3}$$

indexed by m, does have a stationary distribution due to the periodic nature of u_i. Since $\{n_i\}$ is a Markov process it follows that n_{mM+i} indexed by m is also a Markov process. Now define the generating function of the limiting distributions of n_{mM+i} for $i = 1, 2 \cdots M$ as

$$\phi_i(z) = \lim_{m \to \infty} E(z^{n_{mM+i}}), \quad i = 1, 2 \cdots M. \tag{B.4}$$

Define $p_{i,0}$ as the steady state probability that the buffer is empty when the ith slot begins. Notice that the set of $\{p_{i,0}\}$ represents M unknowns. It can then be shown that

$$\phi_{i+1}(z) = \begin{cases} \phi_i(z)\chi(z), & \text{if } u_i = 0, \\ (z^{-1}\phi_i(z) + (1 - z^{-1})p_{i,0})\chi(z), & \text{if } u_i = 1. \end{cases} \tag{B.5}$$

Which can then be rewritten in matrix form as

$$\begin{bmatrix} \phi_1(z) \\ \phi_2(z) \\ \phi_3(z) \\ \vdots \\ \phi_M(z) \end{bmatrix} = z^{-1}\chi(z) \begin{bmatrix} 0 & 0 & 0 & \cdots & \epsilon_M \\ \epsilon_1 & 0 & 0 & \cdots & 0 \\ 0 & \epsilon_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & 0 & \cdots & \epsilon_{M-1} & 0 \end{bmatrix} \begin{bmatrix} \phi_1(z) \\ \phi_2(z) \\ \phi_3(z) \\ \vdots \\ \phi_M(z) \end{bmatrix} + (1 - z^{-1})\chi(z) \begin{bmatrix} p_{M,0}\delta_M \\ p_{1,0}\delta_1 \\ p_{2,0}\delta_2 \\ \vdots \\ p_{M-1,0}\delta_{M-1} \end{bmatrix}, \tag{B.6}$$

where,

$$\epsilon_i = \begin{cases} 1, & \text{if slot } i \text{ is dedicated to the source,} \\ z, & \text{otherwise,} \end{cases}$$

and,

$$\delta_i = \begin{cases} 1, & \text{if slot } i \text{ is dedicated to the source,} \\ 0, & \text{otherwise.} \end{cases}$$

The above can be rewritten more compactly as,

$$[\mathbf{I} - z^{-1}\chi(z)\mathbf{A}(z)]\phi(z) = (1 - z^{-1})\chi(z)\mathbf{P}. \tag{B.7}$$

Letting the inverse of the left-hand size matrix of the above be denoted as $B(z)$, it can be shown that the first row of B is,

$$B_1 = \frac{z^{M-J}}{z^{M-J} - \chi(z)} \times \left[1, \prod_{j=2}^{M} z^{-1}\epsilon_j \chi^{M-1}(z), \cdots, \prod_{j=M-1}^{M} z^{-1}\epsilon_j \chi^2(z), z^{-1}\epsilon_M \chi(z)\right]. \tag{B.8}$$

Subsequent rows of B can be computed recursively if desired. The generating function of ϕ_i can now be expressed as,

$$\phi_i(z) = (1 - z^{-1})\chi(z)B_i\mathbf{P}. \tag{B.9}$$

Using Rouché's theorem as described in [Hay84] we can solve for $P_{i,0}, i \in 1 \rightarrow (M - J)$ by forming a linear set of equations which can be written in matrix form as,

$$\begin{bmatrix} 1 & \cdots & 1 & 1 \\ X(\gamma_2^{M-J-1})/\gamma_2^{M-J-1} & \cdots & X(\gamma_2)/\gamma_2 & 1 \\ X(\gamma_3^{M-J-1})/\gamma_3^{M-J-1} & \cdots & X(\gamma_3)/\gamma_3 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ X(\gamma_{M-J}^{M-J-1})/\gamma_{M-J}^{M-J-1} & \cdots & X(\gamma_{M-J})/\gamma_{M-J} & 1 \end{bmatrix} \cdot \begin{bmatrix} p_{1,0} \\ p_{2,0} \\ \vdots \\ p_{M-J,0} \end{bmatrix} = \begin{bmatrix} \rho \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{B.10}$$

Here $\gamma_1 = 1, \gamma_2, \gamma_3, \cdots \gamma_{M_J}$ are the roots of the denominator in equation B.8. The scalar ρ is determined by differentiating $\phi_{i+1}(z)$ with respect to z and letting $z \rightarrow 1$ which is written as,

$$\rho = (M - J) - M \cdot E(x_n). \tag{B.11}$$

With the values of $P_{i,0}$ known, it is then possible to calculate any other value for P by inverting the generating function, $\phi_i(z)$. This can be calculated numerically, as described in [Hay84], by evaluating $\phi_i(z)$ at a sufficiently large number of points around the unit circle of the complex Z-plane,

$$P_{i,j} = \frac{1}{K} \sum_{k=0}^{K-1} \phi_i(e^{\frac{-j2\pi k}{K}}) e^{\frac{j2\pi k}{K}}. \tag{B.12}$$

B.2 Mean Delay

Mean delay is derived by tagging a particular arrived packet in $]i, i+1]$. The number of packets ahead of this tagged packet are composed of packets already in the buffer, n_i, and any new packets arriving ahead of the tagged packet in $]i, i+1]$. Let the number of new packets arriving be denoted as x_i^* whose generating function is,

$$X^*(z) = \frac{e^{\lambda(z-1)} - 1}{\lambda(z-1)}. \tag{B.13}$$

The delay the tagged packet incurs is then the sum of the above factors plus the delay caused by interruption in service due to overhead in the frame. Then, to find the delay, let E_i denote the sum of the above two quantities which may be expressed as $E_i(z) = X^*(z) \times P_i(z)$. Note that this corresponds to a convolution of the pdfs, X_n and $P_{i,n}$. The delay caused by the interruptions can then be derived from E_i and the slot in the frame corresponding to i. For the situation illustrated in Figure B.1, the delay experienced by a tagged packet arriving in slot i can be written as, For $0 \leq i \leq K-1$:

$$D_i = \begin{cases} E_i + 1, & E_i \leq K - (i+1), \\ K - i + J + M((E_M - (K - i)) \bmod K) + (E_M - (K - i)) \div K + 1, & \text{otherwise.} \end{cases} \tag{B.14}$$

For $K \leq i \leq M$:

$$D_i = \begin{cases} M - i + E_i + 1, & E_i \leq K - 1, \\ M - i + M(E_i \bmod K) + E_i \div K + 1, & \text{otherwise.} \end{cases} \tag{B.15}$$

where $K = M - J$. Lastly, since packets are equally likely to arrive in each slot and the slot time is normalized, total mean delay is written as,

$$\overline{D} = \frac{\sum_i^M D_i}{M}. \tag{B.16}$$

Figure B.2 summarizes the intermediate results for the above described procedure. In this example, an aggregate arrival rate is 0.5 packets per slot is used. The entire frame consists of 25 contiguous slots followed by 5 slots of overhead. The first sub-plot gives the pdf for the number of customers in the queue at the beginning of the first slot. Similar results exist for each slot in the frame. The second plot simply illustrates the pdf for the arrival process, χ, used in the calculations. The third sub-plot similarly shows the pdf for the number of packets arriving ahead of a tagged packet. Sub-plots 3 and 1 can then be convolved ($X^*(z) \times P_i(z)$) as shown in the final sub-plot, which finally leads to the calculation of D_i. Lastly, it is a trivial matter to employ equation B.16 to find the average mean delay.

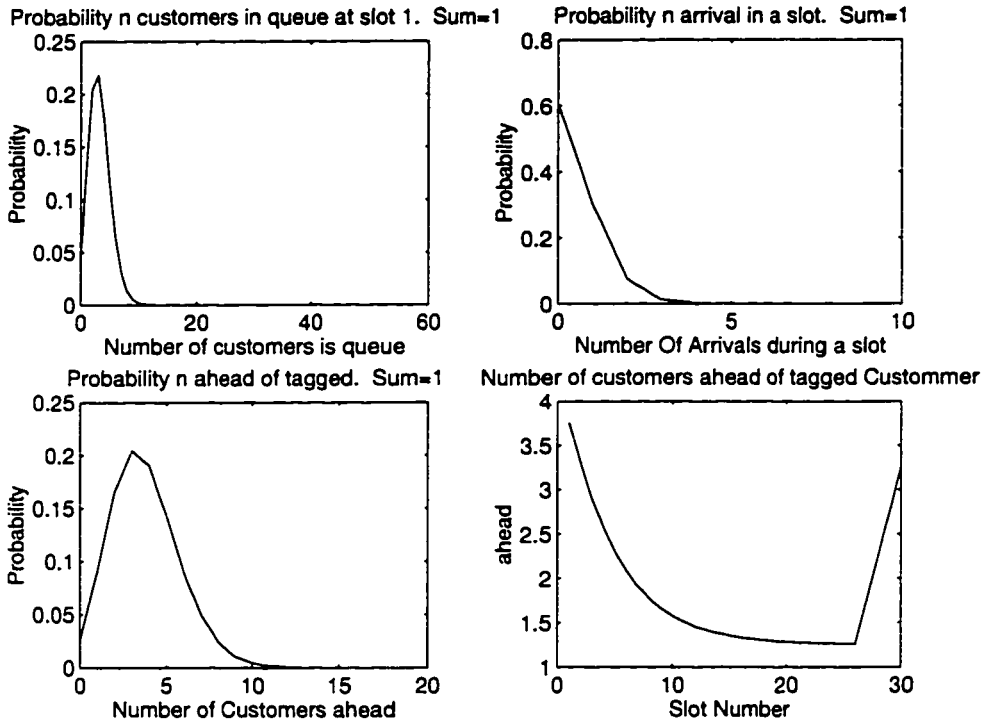


Figure B.2: Some intermediate results for the "come-right-in" model with 30 slots in a frame with 5 of those slots frame overhead. In this case $\lambda = 0.5$ packets per slot.

Appendix C

Discrete-Event Simulation

The execution of a model representing a real, physical system, based upon our understanding of that system, on a digital computer is referred to as computer simulation. As far as a current understanding of the underlying principles of the real physical system can take us, computer simulation emphasizes the concept of "learning by doing". In this thesis, a very specific type of simulation was executed: discrete-event simulation.

The obvious method of simulating systems is to use a continuous time simulator which increments a clock using *time slices*. At each time slice the state of the system is calculated. This approach may be appropriate for many systems, however it can be quite slow and time consuming when the time slices are very small or the simulation needs to be run for a long time.

A discrete-event simulation defines a set of events which occur at specific instances in time. Rather than gradually incrementing a clock, the discrete-event simulator jumps the clock to the time that the next event occurs. Clearly this can speed up the simulation as in many cases we have no interest in the state of the system between events.

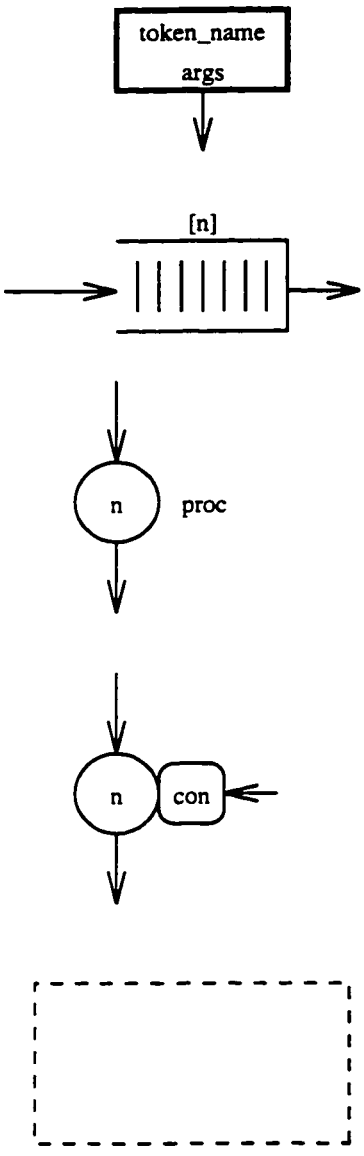
C.1 Why Simulate?

In most physical systems, as in the case of the BTMSC and BTMSR systems, an analytical model of performance will exist. These solutions often take the form of a set of linear equations. However, simulation is indispensable in cases where the solution of the underlying mathematical model is not feasible. This may be because the model is very complex with a large number of variables and interdependencies. Furthermore, when developing a closed-form approach, the model is often twisted for mathematical convenience even when it may not accurately represent the physical system. A computer simulation can more accurately represent a physical system by avoiding the simplifying assumptions often required in mathematical models.

Simulation and closed-form models are complimentary approaches. For example, one might develop a mathematical model for a theoretical physical system. Since the system is theoretical, there will be no way to verify the mathematics for correctness. A simulator can then be built representing the system in order to prove correctness.

C.2 Problem Description

For discrete-event systems the identification of events for simulation can be quite an art. One method is to visualize the flow of *customers* through the system. An event can then be identified at each point in time or space where something occurs to the customer. In most cases there will be several types of customers in the system. Figure C.1 illustrates some primitive building blocks that are used in the construction of a discrete-event simulation. A generation block is used to introduce customers into a system based upon some arrival process. Queues are used at the entrance to servers which have some service time that may be stochastic. A gated server utilizes some control algorithm which may be controlled by the arrival of a separate customer type. Super blocks are used to encapsulate the functionality of a complex object and repeat it as often as necessary without having to redraw the complex object. Using this incomplete set of building blocks fairly complex simulations may be constructed.



Generation Block
args: token_name
arrival process
first order arrival statistic
2nd order " "

Queue
args: n - queue size

Server
args: n - service time
service disipline (eg FIFO)

Gated Server
args: con - control algorithm

Super Block
args: block name

Figure C.1: Primitive Simulation Blocks

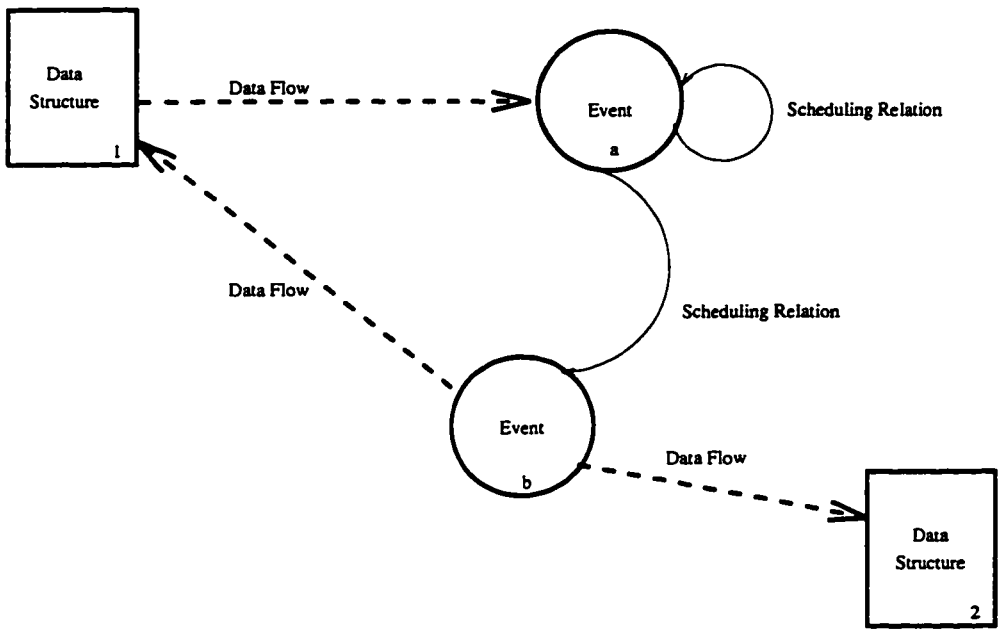


Figure C.2: Primitive Simulation Events

C.3 Events

Once a system has been visualized using the primitive blocks described, it is a straight forward matter of identifying *events*. Simply, each time a customer is generated or encounters a block, an event should be defined. To aid the development of a simulation, a second step of should be engaged. This step clarifies the scheduling relationships between events allowing the transfer of the high level description of the system into computer code. Figure C.2 depicts four basic relations. A data structure block defines a state for a particular customer, server, or control algorithm. Event blocks are the routines which are scheduled to occur at specific times. Date flow and scheduling interdependencies are depicted connecting related blocks.

In Figure C.2 two data structures and two events are shown. It can be seen that event a is responsible for scheduling the time at which both event a and event b will occur. It also can be seen that event a uses data structure 1 without modifying it. Event b is shown not to schedule any events, but modifies both data structures 1 and 2. Note that this example does not do anything useful.

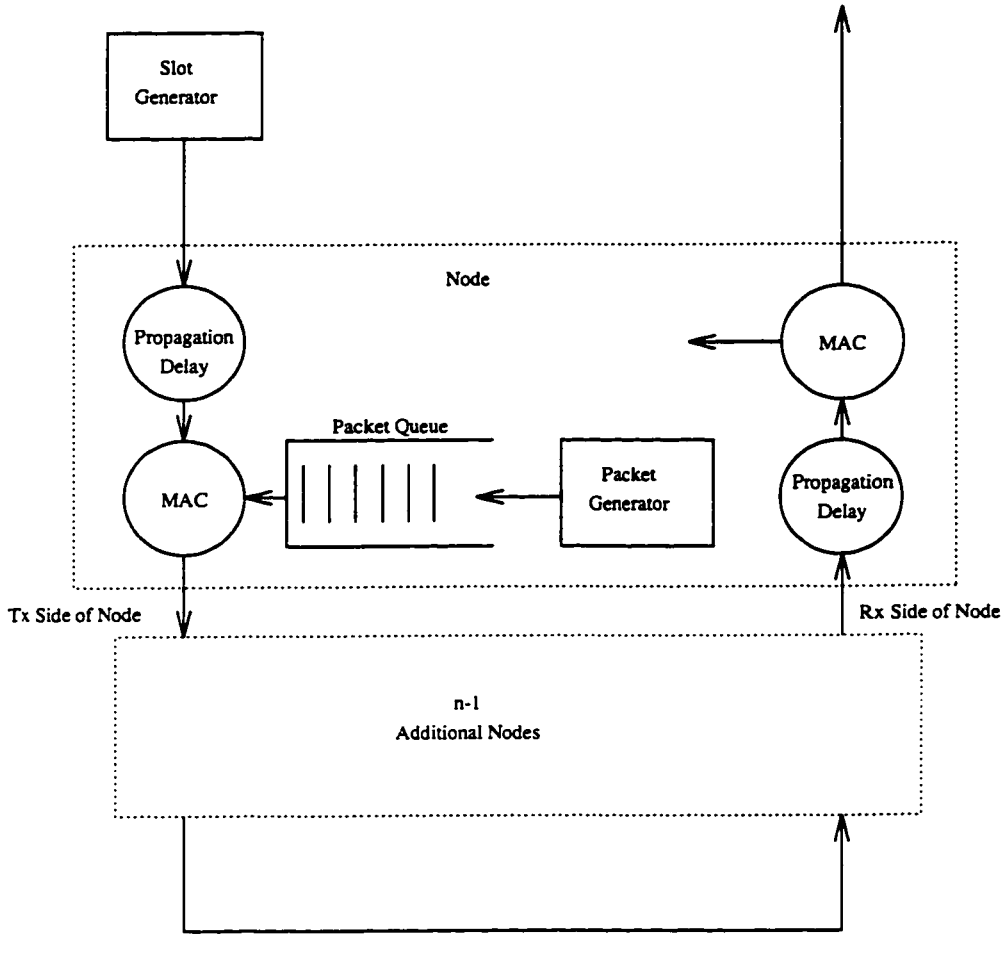


Figure C.3: Morewell Simulation Customer Flow Diagram.

C.4 Examples

Figure C.3 illustrates the customer flow diagram for the simulation of the Morewell passive bus system. Slots are generated which propagate to station, encountering some propagation delay. A second customer type, data packets, are generated at each station and queued until they can be serviced. Service at the station is dependent on a media access protocol. Slots, possibly containing packets propagate to downstream stations on the Tx side of the bus. On the Rx side of the bus, slots are examined and packet customers are removed from the simulation when appropriate. Finally, slot customers are removed from the system at the end of the Rx bus.

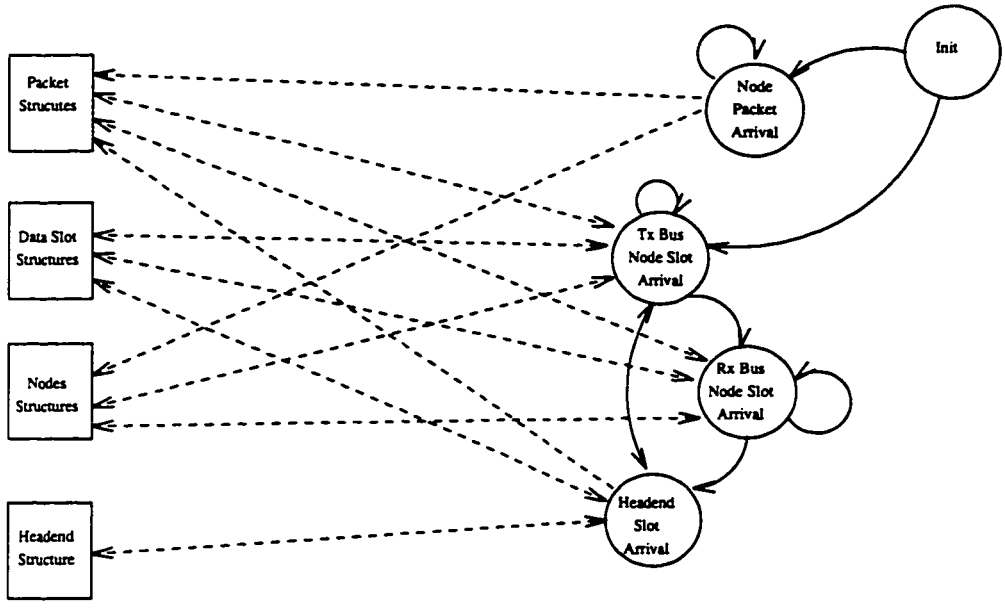


Figure C.4: Morewell Simulation Event Flow Diagram.

Statistics are maintained throughout the system in a number of data structures. The interplay between structures and events is illustrated in Figure C.4. The simulation was then implemented using a "C" code simulation library called SIMLIB.

C.5 The SIMLIB Library

The simlib library consists of a set of simulation tools written in C which facilitate discrete event simulation of computer networks. The purpose of simlib is to provide a "seed" for simulating discrete-event systems. The simlib package has been used for a under-graduate (senior) course in computer networks. For a descriptive overview of the basics of discrete-event simulation, the reader is directed towards [Abu88].

The current version of simlib, grew out of my personal library of simulation tools called GSimLib. T. D. Todd took on the task of writing the documentation and re-writing much of the code to be more "understandable" for under-graduate students that have little experience with the C programming language. Dr. Todd also added the code for the servers.

C.6 Getting Started

A version of the library is available for anonymous ftp at `laser.crl.mcmaster.ca` in `/pub/simlib.c`. In order to use the library, place the file `simlib.c` in your directory and include it at the top of your C program using the statement

```
#include "simlib.c"
```

Simlib will define a number of variables and data structures needed for the simulation including event definitions and an event list. In addition, types are declared for defining servers and FIFO queues. You must initialize SimLib at the beginning of your `main()` routine using the call

```
InitializeSimLib();
```

This sets up the event list and must be done prior to calling any simlib routines. Also, if you are doing multiple simulation runs within the same program, `InitializeSimLib()` should be called prior to starting a new run. Some of the features of the library are discussed below. Simple usage examples are also given.

C.7 Server Declaration and Routines

Servers are constructed by first declaring a pointer to the server and then by calling a library routine to create the server itself. For example, a pointer to a server named "Server1" is first declared using the statement

```
Server *Server1;
```

Then the server is created and initialized using

```
Server1 = CreateServer();
```

The following is a list of all other routines that may be used to access a server.

1. **Server *CreateServer(void)**
   Memory is allocated and the server is initialized. A pointer to the server is returned. During creation, the server state is set to "FREE".

2. **void PutInServer(Server *ServerID, void* ContentPointer)**
   This enters an item into the server. To place something into the server you must pass a pointer to the object through ContentPointer. The server is then marked "BUSY".

3. **void* GetFromServer(Server *ServerID)**
   This removes an object previously placed into service. A pointer to the object is returned and the server is marked "FREE".

4. **ServerState GetServerState(Server *aServerPtr)**
   This returns the current state of a server. A pointer to the server is passed to this routine and it returns an enumerated type being either "FREE" or "BUSY".

5. **double GetServerUtilization(Server *aServerPtr)**
   This routine returns the utilization of the server from the time it was created until the current Clock time.

```
/* Declare something to put into a server */
struct customer {
  float arriveTime;
  float serviceTime;
} aCustomer;

/* Make a server */
Server *Server1;
Server1 = CreateServer();

/* Put the customer into it */
```

```
PutInServer(Server1, (void*)&aCustomer);

/* Sometime later take it out */
struct customer *cPtr;
if(GetServerState(Server1) == BUSY)
        cPtr = (struct customer*)GetFromServer(Server1);

/* Get the Server utilization */
util = GetServerUtilization(Server1);

/* Print the arriveTime */
printf("Customer arrive time = %f\n", cPtr->arriveTime);
```

C.8 FIFO Queue Declaration and Routines

FIFO queues are created in a manner similar to servers. A pointer to the queue is first declared, then a library routine is called to create the queue. As in the server case, arbitrary objects may be stored in the queue and are passed via pointers. Any number of objects may be entered and are removed in first-in-first-out order. In the following example, a FIFOqueue named Queue2 is declared and created.

```
FIFOqueue *Queue2;
Queue2 = CreateFIFOqueue();
```

The following is a list of routines which may be used to manipulate FIFOqueues.

1. **FIFOqueue* CreateQueue(void)**
   This routine creates a queue and returns a pointer to it. The queue is initialized with a size of zero.

2. **void PutInFIFOqueue(FIFOqueue *QueuePtr, void* ContentPointer)**
   This routine enters an object into the queue. Objects entered are added to the back of the queue. This is done by passing a pointer to the object through

ContentPointer. QueuePtr is a pointer to the queue which is being written into.

3. **void* GetFromFIFOqueue(FIFOqueue *QueuePtr)**
   This returns the object currently at the front of the queue pointed to by QueuePtr.

4. **int GetFIFOqueueSize(FIFOqueue *QueuePtr)**
   This routine returns the number of objects currently stored in the FIFOqueue pointed to by QueuePtr.

5. **void *SeeFrontOfFIFOqueue(FIFOqueue *aQueue);**
   This routine returns a void pointer to the contents of the front of the specified FIFOqueue. This permits one to inspect the contents of the front object in the queue without removing it.

```
/* Create some things to put in a queue */
struct customer {
  float arriveTime;
  float serviceTime;
} Customer1, Customer2;

/* Create a queue */
FIFOqueue *ThisQueue;
ThisQueue = CreateFIFOqueue();

/* Put the two customers in it */
PutInFIFOqueue(ThisQueue, (void*)&Customer1);
PutInFIFOqueue(ThisQueue, (void*)&Customer2);

/* Check the size of the queue */
qSize = GetFIFOqueueSize(ThisQueue);
```

```
/* Record the arrive time of the customer at the front of the queue */
/* Leave the customer in the queue */
custPtr = (struct customer*)SeeFrontOfFIFOqueue(ThisQueue);
arTime = custPtr->arriveTime;

/* Take a customer out of the queue */
custPtr = (struct customer*)GetFromFIFOqueue(ThisQueue);
```

C.9 Event List Routines

When Simlib is included in a program, it declares an event list and provides routines for scheduling and executing events. A `Clock` variable is also declared and is initialized to 0.

To schedule and execute events, first the event routines and event types must be declared, in that order. In the following example, an ArrivalEvent and DepartureEvent are declared. The event routine associated with the ArrivalEvent is `Arrival()` and that with the DepartureEvent is `Departure()`.

```
void Arrival(void*);
void Departure(void*);

EventType  ArrivalEvent =   {"Customer Arrival", Arrival};
EventType  DepartureEvent = {"Customer Departure", Departure};
```

The first two lines declare the event routines associated with the two events. *The argument of every event routine must be declared type* `void*` *and the return type must be declared as type* `void`. The next two statements define the two event types. An initialization must be provided when defining an event. The first initialization argument is a description string which is used for output when the simulation trace is enabled. The second initialization argument must be the name of the event routine defined in the previous lines. For each of the events defined, an event routine with the defined name must be provided later on in the simulation. It will be called whenever

an event of that type is to be executed. For example, a routine named `Arrival()` must be written and will be called whenever an ArrivalEvent is to occur.

In order to schedule an event, the **ScheduleEvent** routine is called. For example, to schedule an **ArrivalEvent** at time 10.0, we would use the statement

```
ScheduleEvent(&ArrivalEvent, 10.0, (void*)NULL);
```

The first argument must be the address of the event to be scheduled. The second argument is the `Clock` time that the event is to occur. The last argument is a pointer which is stored with the event and may be used to pass objects along with the event. When the event occurs, this pointer will be passed as the argument of the called event routine. Note that when passing a pointer to an object, the pointer should be cast to a void pointer. If nothing is to be passed, then pass a NULL pointer using the syntax shown. When an event is scheduled, simlib assigns a unique long int event identifier to the event. This identifier is returned by **ScheduleEvent** and may be saved by the calling program.

The following are the routines which may be used to handle events.

1. **long int ScheduleEvent(eventPointer ePtr, double evTime, void *objPtr)**

   This routine schedules an event pointed to by ePtr to occur at time evTime and stores the pointer objPtr along with the event. When the event occurs, the appropriate event routine is called and this pointer is passed as the argument to the event routine. A unique event identifier is returned when an event is scheduled.

2. **void ExecuteNextEvent(void)**

   This routine removes the next event from the event list, updates the simulation variable `Clock` to the new event time, then calls the appropriate event routine. At that time, a pointer to any data sent with the event will be passed through the argument of the event routine. See the example below.

3. **void StartEventTrace(void)**

   When this routine is called, the event trace function is invoked. Whenever

an event occurs, or is scheduled, SimLib prints on the display. In a working simulation, this will generate a flood of screen output.

4. **void StopEventTrace(void)**
   When called, this routine stops the event trace feature.

The following is an example illustrating the use of these routines.

```
/* Define two event routines and their events */
/* Must return type void and have argument type void* */
void ArriveRoutine(void*);
void DepartRoutine(void*);
EventType PacketArrival = {"Packet Arrival", ArriveRoutine};
EventType PacketDeparture = {"Packet Departure", DepartRoutine};

/* Define a packet */
struct packet {
  int DestinationAddress;
  int SourceAddress;
} Packet1;

/* Sometime later, schedule a packet arrival for Clock = 3.0. */
/* Attach the packet to the event. */
ScheduleEvent(&PacketArrival, 3.0, (void*)&Packet1);

/* At Clock = 3.0 we would find ourselves in the ArriveRoutine */
/* Here is the declaration of this event routine */
void ArriveRoutine(void *dataPtr) {
.
.
.
/* Look at the destination address of the arriving packet */
```

```
/* The passed data must be type cast to the correct type */
struct packet *PktPtr;

PktPtr = (struct packet*)dataPtr;
dAddress = PktPtr->DestinationAddress;

/* Schedule a PacketDeparture */
/* Send the packet along with it */
ScheduleEvent(&PacketDeparture, 10.0, (void*)PktPtr);
.
.
.
}
```

C.10 Random Generators

A couple simple random number generators are provided and are described as follows.

1. **double Random(void)**
   This routine returns a pseudo-random number uniformly distributed over the interval $[0, 1)$. It uses the standard library function `rand()` and the constant `RAND_MAX` for this purpose. It is important that these are available.

2. **double ExponentialGenerator(double rParameter)**
   This returns pseudo-random variates exponentially distributed with a mean of 1/rParameter. Since a Poisson Process has exponentially distributed interarrival times, one may be generated by scheduling events at times defined by calls to `ExponentialGenerator`.

```
/* Define a Poisson Arrival event */
void P_Arrival(void*);
```

```
EventType PoissonArrival = {"Poisson Arrival", P_Arrival};

/* In main(), schedule the first Poisson arrival for Clock = 0 */
ScheduleEvent(&PoissonArrival, 0.0, (void*)NULL);

/* In the P_Arrival() routine, schedule the next Poisson arrival */
/* The time for the next arrival will be the current time plus the next
   exponential interarrival time */
void P_Arrival(void *dummyPtr) {

  ScheduleEvent(&PoissonArrival,
        Clock+ExponentialGenerator(1/MeanInterArrivalTime), (void*)NULL);
}
```

Bibliography

[80285] ANSI/IEEE Standard 802.5-1985. *Token Ring Access Method and Physical Layer Specifications.* IEEE, 1985.

[80290a] IEEE Std 802.4-1990. *Token-Passing Bus.* IEEE, 1990.

[80290b] IEEE 802.6. Distributed dual bus (dqdb) subnetwork of a metropolitan area network (man. Technical report, IEEE 802.6 Proposal and submissions, February 7 1990.

[90285] IEEE Std 902.3-1985. *Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications.* IEEE, 1985. ISBN 0-471-892749-5.

[Abu88] Maurice F. Aburdene. *Computer Simulation of Dynamic Systems.* Wm. C. Brown Publishers, 1988.

[Aca94] Anthony S. Acampora. *An Introduction to Broadband Networks.* Plenum Press, New York, 1994.

[AFB79] R. R. Anderson, G. J. Foschini, and B. Bopinath. A queueing model for a hybrid data multiplexer. *The Bell System Technical Journal*, 58(2):279–300, February 1979.

[AGI88] A. Albanese, M. W. Garrett, and A. Ippoliti. Bellcore metrocore network - a test-bed for metropolitan area network research. In *Globecom*, pages 1229–1234. IEEE, 1988.

[AT95] Nima Ahmadvand and Terence Todd. Dual-hop lans using station wavelength routing. In *ICCCN'95*, pages 622–629. IEEE, September 1995.

[AZN95] Ronald G. Addie, Moshe Zekerman, and Tim Neame. Fractal traffic: Measurements, modelling and performance evaluation. In *Infocom'95*, pages 977–984. IEEE, 1995.

[BBC+96] George Benke, Jim Brandt, Helen Chen, Siamak Dastangoo, and Gregory J. Miller. Performance analysis of atm abr service under self-similar traffic in the presence of background vbr traffic. In *ICCCN'96*. IEEE, 1996.

[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.

[Bel80] A. G. Bell. Selenium and the photophone. *The Electrician*, pages 214, 215, 220, 221, 1880.

[BM93] Subrata Banerjee and Biswanath Mukherjee. Fairnet: A WDM-based multiple channel lightwave network with adaptive and fair scheduling policy. *Journal of Lightwave Technology*, 11(5/6):1104–1112, May/June 1993.

[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.

[BV95] G.B. Brewster and M.K. Vernon. The fairness of dqdb networks with slot reuse. In *Infocom*, pages 1154–1163. IEEE, 1995.

[CG88] I. Chlamtac and A. Ganz. A multibus train communication architecture (AMTRAC) for high-speed fiber optic networks. *IEEE Journal on Selected Areas of Communications*, 6:903–912, July 1988.

[Che90] K. W. Cheung. Acoustooptic tunable filters in narrowband wdm networks: System issues and network applications. *IEEE Journal On Selected Areas In Communications*, 8(6):1015–1025, August 1990.

[Che93] K. W. Cheung. Adaptive-cycle tunable-access (ACTA) protocol: A simple, high-performance protocol for tunable-channel multi-access (TCMA) networks. In *ICC'93*. The Chinese University of Hong Kong, 1993. Hot Topic Session Number: HT-10, HT-11.

[CM92] K. W. Cheung and V. W. Mak. EQEB - a multi-channel extension of the DQDB protocol with tunable channel access. *Globecom '92*, 3:1610–1617, 1992.

[CSR90] S.D. Cusworth, J.M. Senior, and A. Ryley. Wavelength division multiple access on a high-speed optical fibre LAN. *Computer Networks and ISDN Systems*, pages 323–333, 1989/90.

[DSP87] E. Desurvire, J. R. Simpson, and P. C. Parker. High-gian erbium-doped travelling-wave fibre amplifier. *Opt. Lett.*, pages 888–890, 1987.

[ea93a] C. A. Brackett et al. A scalable multiwavelength multihop optical network: A proposal for research on all-optical networks. *Journal of Lightwave Technology*, 11(5/6):736–753, May/June 1993.

[ea93b] S. B. Alexander et al. A precompetitive consortium on wide-band all-optical networks. *Journal of Lightwave Technology*, 11(5/6), May/June 1993.

[FA85] R. M. Falconer and J. L. Adams. Orwell: A protocol for an integrated services local network. *Br. Telecom Tech. Journal*, 3(4):27–35, October 1985.

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

[GBW83] Inder S. Gopal, M. A. Bonuccelli, and C. K. Wong. Scheduling in multibeam satellites with interfering zones. *IEEE Transactions On Communications*, COM-31(8):941–951, August 1983.

[Gow93] John Gowar. *Optical Communications Systems.* Prentice-Hall, 2 edition, 1993.

[GPPB77] A. R. Goodwin, J. F. Peters, M. Pion, and W. O. Bourne. GaAs lasers with consistently low degradation rates at room tempratures. *Appl. Phys. Lett.*, 30(2):110–113, 1977.

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

[GT93] Adrian M. Grah and Terry D. Todd. A simplified photonic bus network design. In *Photonics'93.* IEEE, September 1993.

[GT94] Adrian M. Grah and Terence D. Todd. Simplified photonic bus and ring networks using a headend controller. In *Infocom*, pages 263–271. IEEE, June 1994.

[GT95] Adrian Grah and Terence D. Todd. Photonic ring and bus networks with reduced station hardware. *Journal of High Speed Networks*, 4(1):77–98, 1995.

[GT96a] Adrian Grah and Terence D. Todd. Packet-switched local area networks using wavelength-selective station couplers. *IEEE Transactions On Networking*, 1996. Submitted for publication.

[GT96b] Adrian Grah and Terence D. Todd. Protocols for photonic local area networks using wavelength-selective couplers. In *ICCCN'96*, pages 340–343. IEEE, October 1996.

[Has94] H.S. Hassanein. An effective erasure node algorithm for slot reuse in DQDB. In *Infocom*, pages 1302–1309. IEEE, 1994.

[Hay84] Jermiah F. Hayes. *Modeling and Analysis of Computer Communications Networks.* Plenum Press, New York, 1984.

[HCM90] E.L. Hahne, A. Choudhury, and N. Maxemchuk. Improving the fairness of distributed-queue-dual-bus networks. In *Infocom*, pages 175–184. IEEE, 1990.

[IK92] Mansour I. Irshid and Mohsen Kavehrad. A fully transparent fiber-optic ring architecture for WDM networks. *Journal Of Lightwave Technology*, 10(1):101–108, January 1992.

[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.

[Jan96] Mark Janoska. *Multi-channel optical network architectures with channel controller based media access protocols.* PhD thesis, McMaster University, Hamilton, Ontario, may 1996.

[Jon88] W.B. Jones. *Optical Fiber Communications Systems.* Holt, Rinehart and Winston, 1988.

[Jos96] Sabu Joseph. A hierarchical photonic network architecture using spectral folding (HASF). Master's thesis, McMaster University, January 1996.

[JT93] M. Janoska and T.D. Todd. A simplified optical star network using distributed channel controllers. *GLOBECOM'93*, 1:481–487, December 1993.

[JT94] M. Janoska and T.D. Todd. A simplified optical star network using distributed channel controllers. *IEEE Journal of Lightwave Technology*, 12(11):2011–2022, nov 1994.

[KH66] K. C. Koa and G. A. Hockham. Dielectric fiber surface waveguides for optical frequencies. *Proc. IEEE*, 113(7):1151–1148, 1966.

[Kra91] J. M. Kraushaar. Fiber deployment update-end of year, 1990. Technical report, Federal Commu. Commission, Common Carrier Bureau, Washington, DC 20554, March 1991.

[LF82] J.O. Limb and C. Flores. Description of fasnet - a unidirectional local-area communications network. *Bell Syst. Tech. J.*, 61:1413–1440, 1982.

[LK92] J. Lu and L. Kleinrock. A wavelength division multiple access protocol for high-speed local area networks with a passive star topology. In *unknown*, 1992.

[LT87] L. Luping and T. D. Todd. A simulation study of metropolitan area networks. In *IEEE ICC'87*, pages 963–967, Seattle, Washington, June 1987.

[Mai60] T. H. Maiman. Stimulated optical radiation in ruby. *Nature, Lond.*, pages 493–494, 1960.

[Mil88] S. E. Miller. Overview and summary of progress. In S. E. Miller and I. P. Kaminow, editors, *Optical Fiber Telecommunications II*, pages 1–27. Accademic Press, 1988.

[Muk91a] Biswanath Mukherjee. Architectures and protocols for WDM-based local lightwave networks. CSE-91-32, December 1991.

[Muk91b] Biswanath Mukherjee. The open-ring/active-bus network structure: Access techniques and their heavy-traffic performance. *IEEE Transactions On Communications*, 39(4):474–477, April 1991.

[Muk92] B. Mukherjee. WDM-based local lightwave networks, part 1: Single-hop systems. *IEEE Network Magazine*, 6(3):12–27, May 1992.

[NBH88] R. M. Newman, Z. L. Budrikis, and J. L. Hullett. The QPSX man. *IEEE Communications Magazine*, 26(4):20–28, April 1988.

[NS90] K. Nakagawa and S. Shimada. Optical amplifiers in future optical communications systems. *IEEE LCS*, 1(4):57–62, 1990.

[Pal92] Joseph C. Palais. *Fiber Optic Communications.* Prentice-Hall, Inc., 3rd edition, 1992.

[Ram93] R. Ramaswami. Multiwavelength lightwave networks for computer communication. *IEEE Communications Magazine*, 31(2):78–88, February 1993.

[RCS89] A. Ryley, S.D. Cusworth, and J.M. Senior. Piggybacked token passing access protocol for multichannel optical fibre LANs. *Comput. Comm.*, 12:213–222, 1989.

[Ros89] Floyd E. Ross. An overview of fddi: The fiber distributed data interface. *IEEE Journal On Selected Areas In Communications*, 7(7):1043–1051, September 1989.

[SD93] Krishna M. Sivalingam and Patrick W. Dowd. Performance impact of physical topology on local lightwave networks. Submitted for publication, 1993.

[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.

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

[Sta87] W. Stallings. Local network performance. *IEEE Comm. Mag.*, 25:27–36, 1987.

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

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

[TB92] Terence D. Todd and Allen M. Bignell. Traffic processing algorithms for the signet metropolitan area network. *IEEE Transactions on Communications*, 40(3):568–576, March 1992.

[TC83] C. Tseng and B. Chen. D-Net, a new scheme for high data rate optical local area networks. *IEEE Journal on Selected Areas of Communications*, 1:493–499, April 1983.

[TGB95] Terence D. Todd, Adrian M. Grah, and Oliver Barkovic. Optical local area networks (lans) using wavelength-selective couplers. In *Infocom'95*, pages 916–923. IEEE, April 1995.

[Tyn78] John Tyndall. *Six lectures on light delivered in America in 1872-1873.* New York : D. Appleton, 2nd ed. edition, 1878.

[Wer66] A. Werts. Propagation de la lumière cohérente dans les fibers optiques. *L'Onde Electrique*, 46:967–980, 1966.

[WL90] Tsong-Ho Wu and Richard C. Lau. A class of self-healing ring architectures for sonet network applications. In *Globecom*, pages 403.2.1–403.2.8. IEEE, 1990.

[ZVN87] M. Zafirovic-Vukotic and I.G. Niemegeers. Performance modelling of the orwell basic access mechanism. *Comp. Commun. Rev.*, 17(5), 1987.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6″

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