

**Computer Science Department Technical Report
University of California
Los Angeles, CA 90024-1596**

**DESIGN AND ANALYSIS OF WAVELENGTH DIVISION MULTIPLE
ACCESS LIGHTWAVE PACKET NETWORKS**

Jonathan C. Lu

**September 1992
CSD-920042**

UNIVERSITY OF CALIFORNIA
Los Angeles

**Design and Analysis of
Wavelength Division Multiple Access
Lightwave Packet Networks**

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Computer Science

by

Jonathan Chunhsien Lu

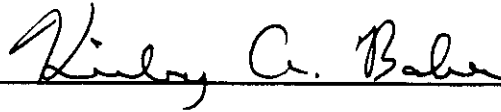
1992

© Copyright by
Jonathan Chunhsien Lu
1992

The dissertation of Jonathan Chunhsien Lu is approved.



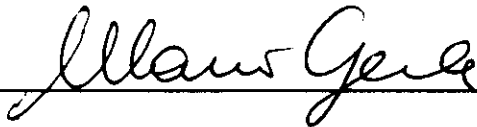
Reza Ahmadi



Kirby Baker



Jack Carlyle



Mario Gerla



Leonard Kleinrock, Committee Chair

University of California, Los Angeles

1992

To my family,
for their everlasting love and support.

TABLE OF CONTENTS

1	Introduction	1
1.1	High Speed Networks and Applications	1
1.2	Wavelength Division Multiple Access (WDMA)	4
1.3	WDMA Architectures and Technologies	7
1.3.1	WDMA Architectures	7
1.3.2	The Power Budget Problem	9
1.3.3	Tunable Lasers and Filters	11
1.4	Previous Related Work	17
1.4.1	Multihop Networks	17
1.4.2	Single-Hop Networks	20
1.4.3	Experimental Systems	24
1.5	Summary of Results	27
2	A Performance Model of WDMA Networks	31
2.1	Introduction	31
2.2	The System Model and the Solution Method	32
2.3	The Uniform Traffic Case	35
2.3.1	Tunable Transmitters and Receivers	35
2.3.2	Tunability on One Side Only (Transmitters or Receivers)	38
2.4	The General Traffic Case	44
2.4.1	The Model	44
2.4.2	An Iterative Solution Procedure	47
2.5	The Hot-Spot Traffic Case	48
2.6	Conclusions	54
3	A WDMA Protocol for LANs with a Passive Star Topology	55
3.1	Introduction	55
3.2	Description of Protocol	56
3.2.1	The Protocol	56
3.2.2	An Illustrative Example	61
3.3	Performance Analysis	63
3.3.1	Model Assumptions	63
3.3.2	The Infinite Population Case	63
3.3.3	The Finite Population Case	68
3.4	Numerical Results	76
3.5	Conclusions	82

4	A WDMA Protocol for MANs with a Dual Bus Topology . . .	86
4.1	Introduction	86
4.2	System Description	87
4.2.1	DQDB Preliminary	88
4.2.2	The WDMA Access Protocol	89
4.3	Performance Analysis	92
4.3.1	Capacity of the System	94
4.3.2	Mean Access Delays for Individual Stations	94
4.4	Numerical Results	96
4.5	Conclusions	99
5	Conclusions and Future Research	101
5.1	WDMA Wide Area Networks	103
5.2	The Case When $W = N$	104
5.3	Protocols to Support Integrated Traffic	105
5.4	High-Speed Circuit/Packet Switch	106
A	107
A.1	Proof That the Denominator $D(z)$ Has W Roots on or Inside the Unit Circle $ z = 1$	107
A.2	Derivations of $f_r(\bar{n})$ and $g_q(i)$	108
	References	110

LIST OF FIGURES

1.1	Loss characteristics and usable bandwidth of single-mode fiber.	3
1.2	Star network. T = transmitter, R = receiver.	8
1.3	Unidirectional bus. T = transmitter, R = receiver.	8
1.4	A multihop lightwave network with a bus topology.	18
1.5	8-NIU logical connectivity graph.	19
2.1	State transition diagram for number of busy wavelength in the system.	35
2.2	State transition diagram for tunable transmitters and receivers.	36
2.3	An upper bound, the W -server loss system.	37
2.4	Throughput versus total load $N\rho$. $N=50$, $W=10$	39
2.5	Throughput versus total load $N\rho$. $N=50$, $W=50$	39
2.6	State transition diagram for tunability on one side only.	40
2.7	Throughput versus total load $N\rho$. $N=50$, $W=10$ and one tunable transmitter.	42
2.8	Throughput versus total load $N\rho$. $N=50$, $W=25$ and one tunable transmitter.	42
2.9	Throughput versus total load $N\rho$. $N=50$, $W=50$ and one tunable transmitter.	43
2.10	State transition diagram for $U^{(i)}$	45
2.11	State transition diagram for $V^{(i)}$	46
2.12	Throughput versus the total load $N\rho$. $N=50$, $W=10$, $r_1 = 1$	49
2.13	Throughput of the hot-spot node versus total load $N\rho$. $N=50$, $W=10$, $b=0.2$	52
2.14	Throughput of the hot-spot node versus total load $N\rho$. $N=50$, $W=10$, $b=0.8$	52
2.15	State transition diagram for H	53
2.16	The relationship between the limiting value S_1 and r_1 . $N=50$, $W=10$	53
3.1	N stations connected by a passive star coupler	57
3.2	Structure of a control slot	58
3.3	A scenario of packet transmissions. ($N=10$, $V=5$, $W=2$, $R=2$.)	62
3.4	Throughput versus offered reservation traffic. $N = \infty$, $W=4$	68
3.5	Throughput versus delay curves. $N = \infty$, $V=10$, $W=4$, $R=10$	69
3.6	An approximate model of the system.	72
3.7	A modified model of Figure 3.6 in the case of $\sigma \leq p$	72
3.8	Throughput versus delay curve. $N=80$, $W=3$, $V=8$, $R=1$, $p=0.6$	77

3.9	Throughput versus delay curve. $N=120, W=4, V=10, R=2, p=0.6.$	78
3.10	Throughput versus delay curves. $N=500, W=4, V=10, R=10, p=0.2.$	80
3.11	Throughput versus V for fixed slot sizes ($V + W$ constant). $N = 500, R=10, \sigma=0.01, p=0.2.$	81
3.12	Delay versus V for fixed slot sizes ($V + W$ constant). $N = 500, R=10, \sigma=0.01, p=0.2.$	81
3.13	Throughput versus propagation delay. $N = 500, W=4, V=10, \sigma=0.01, p=0.2.$	83
3.14	Delay versus propagation delay. $N = 500, W=4, V=10, \sigma=0.01, p=0.2.$	83
3.15	The fraction of success versus the number of data packets transmitted. $N = 500.$	84
4.1	A basic dual bus network.	87
4.2	An example DQDB network.	88
4.3	The structure of a slot on the control channel for $W = 3.$	90
4.4	Total throughput versus traffic load.	97
4.5	Channel efficiency versus traffic load.	97
4.6	Throughput versus delay curves.	98
4.7	Average access delays for individual stations. $W=4.$	100
4.8	Average access delays for individual stations. $W=16.$	100

ACKNOWLEDGMENTS

Many people were involved in completing this dissertation, and here is my chance to express to them my appreciation.

First and foremost is Professor Leonard Kleinrock, my dissertation chairman. Thanks to him for his continued confidence in me. His constant encouragement and insightful guidance have been of immeasurable value to me. It is really more than a pleasure to work with him. I am also thankful to the other committee members consisting of Reza Ahmadi, Kirby Baker, Jack Carlyle, and Mario Gerla, for their invaluable comments and suggestions.

The support by the Advanced Research Projects Agency of the Department of Defense under contract MDA 903-87-C0663, Parallel Systems Laboratory, and contract MDA 972-91-J1011, Advanced Networking and Distributed Systems, for this work is greatly appreciated. I offer my sincere thanks to former colleagues Willard Korfhage, Farid Mehovic, Joy Lin, Bob Felderman, Shioupyn Shen, and Simon Horng, and current group members Chris Ferguson, Donna Nielson, Brian Tung, and Rajeev Gupta for making the environment inspiring and stimulating. Special thanks to Lily Chien and Brenda Ramsey for their administrative assistance.

To my wife Yi-wen and my son Brian, I dedicate this dissertation. They were always beside me with their love and understanding when my research was going nowhere. My last and greatest appreciation goes to my parents for their everlasting love and support in pursuing my dream.

VITA

- 1961 Born, Taipei, Taiwan, Republic of China
- 1983 B. S., Electrical Engineering
National Taiwan University
- 1987 M. S. E., Computer Science & Engineering
University of Michigan, Ann Arbor
- 1988–1992 Graduate Student Researcher under the Parallel Systems Laboratory contract, University of California, Los Angeles

PUBLICATIONS

Jonathan C. Lu and Leonard Kleinrock, "An Access Protocol for High-Speed Optical LANs", *20th ACM Computer Science Conference '92*, pp. 287–293, Kansas City, March 1992.

Jonathan C. Lu and Leonard Kleinrock, "On the Performance of Wavelength Division Multiple Access Networks", *International Conference on Communications '92*, pp. 1151–1157, Chicago, June 1992.

Jonathan C. Lu and Leonard Kleinrock, "A Wavelength Division Multiple Access Protocol for High-Speed Local Area Networks with a Passive Star Topology", to appear in *Performance Evaluation Journal*, 1992.

Jonathan C. Lu and Leonard Kleinrock, "Performance Analysis of Single-Hop Wavelength Division Multiple Access Networks", to appear (invited paper) in *Journal of High Speed Networks*, 1993.

ABSTRACT OF THE DISSERTATION

**Design and Analysis of
Wavelength Division Multiple Access
Lightwave Packet Networks**

by

Jonathan Chunhsien Lu

Doctor of Philosophy in Computer Science

University of California, Los Angeles, 1992

Professor Leonard Kleinrock, Chair

With the advance of fiber optics technology, it is conceivable that we could build multiple access networks on the order of 30 terabits per second (1 terabit = 1,000 gigabits) by using the low-loss band of the optical fiber spectrum (1200-1600 nm). An obstacle is the severe bottleneck at the speed of the electronic interface, typically at a few Gb/s. Wavelength Division Multiple Access (WDMA) eliminates this bottleneck by operating on multiple channels at different wavelengths, with each channel running at full electronic speed.

We first present a mathematical model which closely approximates WDMA multiuser networks with general receiver/transmitter configurations and arbitrary traffic patterns. We first study the case of a uniform traffic matrix and observe that only a small number of tunable transmitters and receivers per station is needed to produce performance close to the upper bound. We then construct a general traffic model and propose an iterative solution procedure. A case of hot-spot traffic is also studied using this model. We find that adding more resources to the hot-spot node helps improve its performance, but only to a limited extent determined by the traffic imbalance.

We next propose a multiple access protocol for a system with a passive star coupler. A station must reserve a wavelength first, then transmit the data on that wavelength. The performance of the protocol is modeled and analyzed for both the infinite and finite population cases. Numerical results show that low delay and high throughput (larger than the electronic speed of a single station) can be achieved. The analysis also shows that the best performance is obtained when the capacities of the reservation channels and the data channels are balanced.

Considering the limitations of star networks in supporting a large number of users in a large area, we then propose a WDMA protocol for a dual bus network. The proposed protocol is a multichannel generalization of the DQDB protocol. An approximate queueing model is built to analyze the system performance. Numerical experiments show that more wavelengths produce higher throughput, but lower efficiency per wavelength. Also we observe that the protocol achieves better fairness than single channel DQDB.

In the end we point out some directions for future research, and conclude that WDMA is a very promising approach to better exploit the enormous bandwidth of optical fiber.

CHAPTER 1

Introduction

1.1 High Speed Networks and Applications

In the past two decades computer communication networks have developed from an experimental technology into an integral component of efficient and successful computing environments. During the 1970's we witnessed the birth and the evolution of the ARPANET where the link speeds were a mere 50 kilobits per second (kb/s). In the 1980's we saw the proliferation of Local Area Networks (LANs), in which stations pump megabits per second (Mb/s), e.g., 10 Mb/s with Ethernet and 100 Mb/s with Fiber Distributed Data Interface (FDDI), into the communication medium. With computer networking, we have been able to support many sophisticated information processing applications via efficient resource and load sharing. However, there are new applications which will consume very large bandwidths, and will require communication networks with capacities as high as tens or hundreds of gigabits per second (Gb/s). Services such as "fiber to the home" [Shu89] and Broadband Integrated Services Digital Network (BISDN) [BKNS89] are expected to provide multiple multimedia connections of high-definition television (HDTV), digital audio, images, and bursty data traffic. As universities and research organizations continually upgrade their computing facilities, we can foresee their inevitable need

to interconnect supercomputers with hundreds of high-resolution full-motion color-graphics workstations. Medical imaging is another important application because its image-quality requirements are so strict that the images must often be transmitted in uncompressed form. In addition to the above is the strong need of a high-speed backbone network to interconnect LANs, and Metropolitan Area Networks (MANs), to support distributed processing systems. With the above applications, it can be envisioned that each end user will access the network at a sustained rate approximately equal to 1 Gb/s [Cat92], and a multi-gigabit network is necessary to support hundreds of them.

Traditional transmission media such as twisted copper wire and coaxial cable can only provide a bandwidth up to a few megahertz (MHz) and a few gigahertz (GHz) [Sta88], respectively. However, the explosive advance of fiber optics technology in the past decade offers a combination of wide bandwidth, low attenuation, and low noise unmatched by any other transmission medium before. Optic fiber is principally taking over the role of twisted copper wire and coaxial cable. The optic fiber provides us with a potential bandwidth of 30 terahertz ($1 \text{ THz} = 10^3 \text{ GHz}$) in the low-loss optical spectrum between $1.2 \mu\text{m}$ and $1.6 \mu\text{m}$, an attenuation of 0.2 decibels (dB) per kilometer at $1.55 \mu\text{m}$ wavelength (0.5 dB at $1.3 \mu\text{m}$), and a bit error rate of 10^{-12} [Hil89]. Figure 1.1 shows the low-loss region of the optic fiber spectrum. It is conceivable that we could build multiple access networks with a total capacity on the order of 30 terabits per second (Tb/s) by using this low-loss band of the optic fiber spectrum [Hen89]. Since the beginning of this decade, optic fibers have been widely deployed in long-haul transmission systems such as the Trans-Atlantic and Trans-Pacific undersea cable systems and land-based interexchange trunks

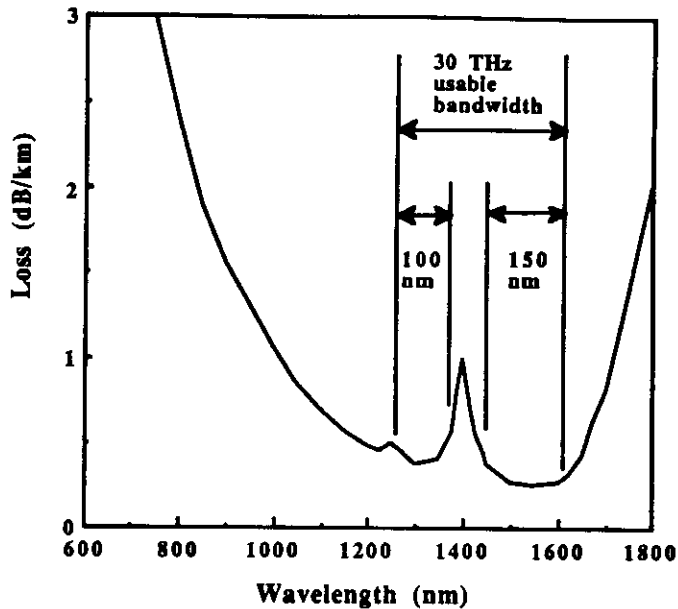


Figure 1.1: Loss characteristics and usable bandwidth of single-mode fiber.

[LZ89]. Twenty-five years ago fiber optic transmission for public telecommunications networks was still viewed as a speculative technology. Today it has become the technology of choice for long distance communication networks.

Although the fiber bandwidth is capable of supporting many terabits per second of throughput, the electronic components at the nodes¹ of the light-wave network, which typically operate at rates (i.e., the rate at which light can be modulated at the laser) of no more than a few Gb/s, can drastically limit the total throughput of the network. In [Gre91], three generations of computer networks can be identified based on the physical-level technology employed. Networks built without use of optic fibers are referred as *first-generation*

¹The words *node* and *station* will be used interchangeably throughout this dissertation.

networks. Several well-known examples include the ARPANET, IBM's SNA, Digital Equipment's DNA, Ethernet, and the IEEE 802.5 Token Ring. *Second-generation* networks still employ traditional architectures but use optic fiber as the transmission medium to enjoy its excellent performance characteristics. Examples of second-generation networks include Expressnet [TBF83], FDDI [Ros86], and Distributed Queue Dual Bus (DQDB) [NBH88]. In these first two generations of networks, however, the maximum throughput of the entire network is limited to the rate that can be supported by the electronics of one of the user stations. In spite of the continuous advances in high-speed electronics and electro-optics technology, this limit on the electronic speed appears to be effectively permanent [AK89]. The *third-generation* networks adopt new approaches which exploit the unique properties of optic fibers (e.g., a capacity four orders of magnitude larger than the peak rate of the electronics) in order to meet the requirements of high-bandwidth applications. The basic problem is to come up with a systems approach that allows the vast bandwidth of the optical medium to be tapped by users, each of which is limited to the maximum electronic speed. In order to accomplish this, traffic corresponding to different user pairs must concurrently be resident in the network at any given moment. If only the traffic of a single user pair can be resident at any moment, then the maximum throughput is limited to the speed of the electronics. Concurrency techniques are therefore the key to tapping the bandwidth of the optical medium.

1.2 Wavelength Division Multiple Access (WDMA)

As discussed in the previous section, third-generation networks must provide multiple concurrent channels, with each channel running at the rate of end

station's electronic interface. In this section we discuss three concurrency techniques, namely Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA), and Wavelength Division Multiple Access (WDMA), wherein multiple channels are addressed according to time slot, waveform, and wavelength, respectively.

TDMA requires that there be many slots per bit time. To achieve this, a special laser is used, which can generate very narrow optical pulses, say of duration 1 to 10 picoseconds (ps), corresponding to bandwidths of 100-1000 GHz; however, the rate of pulse repetition is limited to the much slower electro-optic speed (say, 1 GHz). The narrow pulse is distributed to all the users in each user bit interval. The user will either pass the pulse if the user bit is a logical one, or block the pulse if the user bit is a logical zero. Prior to superimposing the pulses from all the users, various delays are inserted such that none of the pulses overlap in time. In such a manner, the narrow bits associated with each user are time multiplexed together into a TDMA bit frame. The difficulty with this approach lies in generating the various delays needed to keep all pulses properly aligned in the time slots.

We next consider CDMA, wherein each receiver is assigned a unique code word from a set of (pseudo)orthogonal code words. A transmitter wishing to address a receiver appropriately encodes its data using that receiver's code. Since a user's data bit is encoded into a code word consisting of a multiple number of chips (a chip is a narrow pulse), the total capacity of the network is upper bounded by the chip rate. If the chip rate is also limited to the electro-optic speed, then we see that simple CDMA does not offer any more capacity than does the single channel network.

The last scheme in consideration is WDMA, in which multiple channels are created at different wavelengths, with each channel operating at the speed of the electronic interface. Traffic belonging to different user pairs are transmitted in different wavelengths so that they do not interfere with one another. All the signals that are simultaneously transmitted are Wavelength Division Multiplexed (WDM) onto the optical medium where they are combined using couplers and distributed back to each user in the network. The receiver must be able to filter out the desired optical signal and reject the others. We note that this approach allows a multitude of non-interfering transmissions to be simultaneously resident in the network.

The primary drawback of the WDMA approach is the need for wavelength tunability at the transmitters and/or receivers. For each node, either an array of fixed-tuned transmitting lasers is needed, one for each wavelength, or a small number of tunable lasers must be provided. The same is also true of the receivers. The former is prohibitively expensive when a large number of wavelengths are used, while the latter approach requires transmitters/receivers that can be rapidly tuned over a considerable portion of the usable optical band. Both transmitters and receivers with rapid tuning over some fraction of the optical band have been reported in the laboratories [LZ89][KC89]. The WDMA approach holds great promise because it does not require a bit rate faster than the user's electronic interface can operate. Therefore, in the rest of this dissertation we shall concentrate on exploiting WDMA techniques to better utilize the enormous optical bandwidth.

1.3 WDMA Architectures and Technologies

1.3.1 WDMA Architectures

One of the major architectural types of WDMA networks is the *broadcast-and-select* network where all input signals are first combined in a broadcast medium (star coupler or bus) and then distributed to all the outputs. Figures 1.2 and 1.3 show a star and bus network, respectively. There can exist several possibilities, depending on whether the transmitters, the receivers, or both, are made tunable:

1. Tunable transmitters and fixed receivers (at unique wavelengths): In this architecture, a node tunes its transmitter to the appropriate destination receiver wavelength and transmits its data. This system may not require any complicated pre-transmission coordination because it employs fixed receivers. Collision at the destination does exist in such a network, so that a means of contention resolution must be provided. Since only a single wavelength is assigned to each node's receiver, only point-to-point connections are possible and multicast connections cannot be achieved.
2. Fixed transmitters (at unique wavelengths) and tunable receivers: This architecture requires a signaling mechanism to notify the receiver which wavelength to listen to. Multicast connections can be supported by having more than one receiver tuned to the same source wavelength at the same time. *Destination conflict* (the situation when multiple packets are sent to the same destination on different wavelengths at the same time) exists in this architecture, and an arbitration rule must be provided.

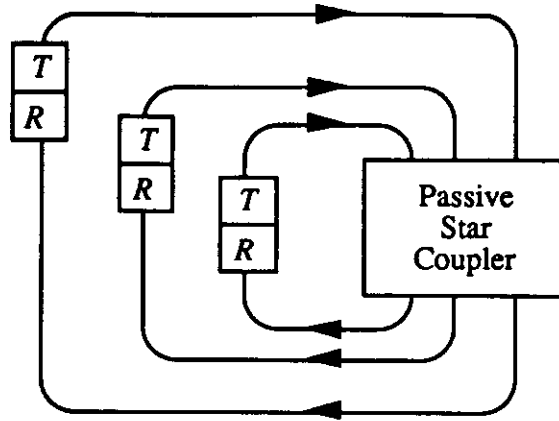


Figure 1.2: Star network. T = transmitter, R = receiver.

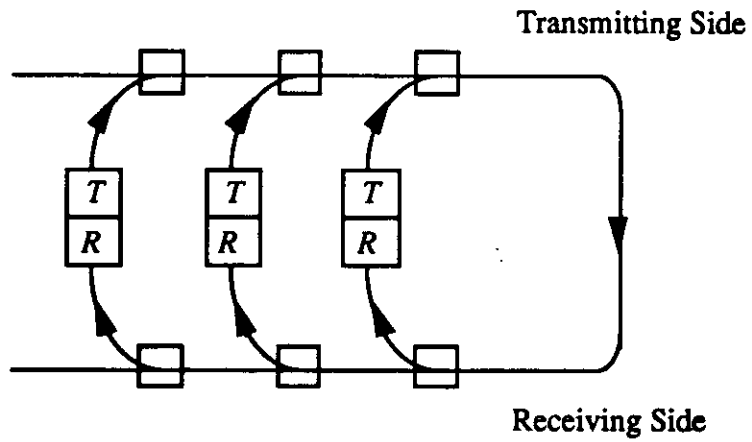


Figure 1.3: Unidirectional bus. T = transmitter, R = receiver.

3. Both transmitters and receivers are tunable: In this architecture the number of wavelengths required can be reduced to a number fewer than the number of stations. This architecture may be the most flexible in accommodating a scalable user population, at the expense, however, of more complicated access control protocols.

An outstanding characteristic of WDMA networks is that, with tunability at the transmitters and/or receivers of each station, almost any arbitrary virtual topology can be created over any given physical topology. Clearly the determination of the optimal virtual topology will depend on the given network load (the traffic matrix), and the number of tunable transmitters and receivers at the stations.

1.3.2 The Power Budget Problem

The traffic capacity of a fiber network, though extremely large, can be limited to less than the bandwidth of the fiber itself due to energy considerations. Let us define E (photons/bit) to be the minimum amount of optical energy required by the receiver to detect a transmitted bit. Consider the star network shown in Figure 1.2. The (passive) star coupler combines signals from all the transmitters (T) simultaneously and broadcasts them to all the receivers (R). Thus, each receiver has access to all the traffic in the network and selects out those messages intended for itself. Suppose that signal power P is available from each of the N transmitters and that the fibers are lossless. The star coupler uniformly distributes these signals to its N output fibers, so that each one delivers to its receiver a contribution P/N from each transmitter. Thus, a given transmitter-receiver pair can exchange data at no more than the power-

limited rate of P/NE bits/sec. The aggregate throughput of the network is therefore P/E . For today's typical semiconductor optical components, $P \approx 1$ mW (1 watt-sec $\approx 10^{19}$ photons) and $E \approx 500$ photons/bit [Per85], and so the power-limited throughput is ≈ 20 Tb/s. Such considerations tend to prevent us from achieving the potential bandwidth (≈ 30 Tb/s) of optic fibers.

The situation is N times worse for networks with bus topologies. Consider the tapped unidirectional bus shown in Figure 1.3. Each user's traffic is first inserted into the transmitting side of the bus by the use of directional couplers, then broadcast to the receivers on the receiving side via a set of splitters. The available signal power at a coupler's output is always considerably smaller than the sum of its input signals. This is because it is impossible to build a power coupler that combines two (or more) uncorrelated inputs and delivers them to its output losslessly [Hen89]. Further analysis shows that the coupler follows the inequality

$$C_1 + C_2 \leq 1$$

where C_1 and C_2 represent the fraction of power coupled to the output from each of the two inputs. This argument can be generalized to the case of an N -station bus. Assume that all transmitters generate equal power, and that the coupling coefficients are adjusted in a "fair" way such that the transmitted power coupled to the network is the same for all stations. In this case only $1/N$ th of the power generated by each transmitter is available at the receiving side. Since this signal must then be distributed among N receivers, it is further reduced by $1/N$, resulting in an overall (transmitter-to-receiver) attenuation of transmitted power by a factor of $1/N^2$. The aggregate throughput of the bus network is therefore P/NE , only $1/N$ of that achieved by the star network.

From the preceding discussion we see that power limitations tend to be more serious than bandwidth limitations for fiber networks, especially those with a bus topology. However, another technology development that helps alleviate the power budget problem is the recent progress on erbium-doped optical amplifiers [Bel91], which can provide extra power to compensate for the tapping and splitting loss mentioned previously. In addition, the passband of these amplifiers is sufficiently broad (tens of nanometers) so that all optical signals within the passband can be amplified simultaneously.

1.3.3 Tunable Lasers and Filters

The two most important technologies that are essential to the WDMA networks are tunable lasers and tunable filters. The parameters of importance are (i) the number of resolvable channels, and (ii) the speed with which the device can switch from one channel to another. Different applications may have different requirements on those two parameters. For example, a tunable filter to select channels carrying home entertainment video may not require as short a tuning time as a filter operating in the internal part of a high-speed switch.

Tunable Lasers

Traditional semiconductor lasers have a wide spectral output, typically about 300 to 600 GHz (1 nm of optical spectrum at a wavelength of 1400 nm corresponds to roughly 150 GHz) [Hil89], which requires a large channel spacing between adjacent channels. This has been much improved by the introduction of lasers with a narrow spectral output, such as the Distributed FeedBack (DFB) lasers [WND83] and Distributed Bragg Reflector (DBR) lasers [MMK87]. A

typical spectral output of such lasers is between 10 and 100 MHz (≈ 0.001 nm), depending on the output power. The spectrum of the modulated laser output is therefore dominated by the signal bandwidth (for a high data rate), and the channel spacing requirement is reduced much further.

Tunable lasers can be classified into four families: thermal tuning, mechanical tuning, injection-current tuning, and acousto-optic tuning [Bra90, LZ89].

- **Thermal tuning:** Thermal tuning lasers are limited to about ± 1 nm of tuning range with a very slow tuning speed, on the order of milliseconds (ms) or longer. This sort of tuning is useful for stabilizing lasers in the laboratory.
- **Mechanical tuning:** This class of lasers changes its transmitting wavelength by rotating a diffraction grating in an external cavity. A tuning range of 55 nm centered at $1.5 \mu\text{m}$ wavelength was reported in [WD83]. The principal drawback of these lasers is their limited tuning speed, on the order of milliseconds, due to the mechanical movements involved for tuning.
- **Injection-current tuning:** There are a variety of semiconductor lasers which are tuned by adjusting the injection current in one or more sections of the laser. These lasers have, by far, achieved the fastest tuning speed (on the order of a few nanoseconds), but the tuning range is limited to 10 nm or so. In these devices, the injected carriers change the effective index of refraction within the optical cavity. In practice, the fraction of change in wavelength is equal to the fraction of change in the effective index n . Since the maximum index change is about 1%, the maximum tuning range

is around 10 – 15 nm [Bra90]. An experiment on a three-section DBR laser [Kob89] demonstrated 20 wavelength tuning while simultaneously amplitude modulating the laser at 200 Mb/s. The tuning time was about 15 ns.

- **Acousto-optic tuning:** Another class of tunable laser is an external-cavity semiconductor laser with an electronically tunable filter within the cavity. Tuning ranges of 7 and 83 nm have been demonstrated for the electro-optic and acousto-optic cases, respectively. The tuning time of the acousto-optic tunable laser was limited by the acoustic velocity in the filter to a few microseconds. The tuning time of the electro-optic tunable laser was not measured.

Table 1.1 summarizes the four major classes of tunable lasers described above. We note that the tuning speed of injection-current tunable lasers is in the order of nanoseconds, which is suitable for gigabit network applications. However, the number of channels it can support is limited. On the other hand, acousto-optic tunable lasers can provide hundreds of channels, but with a slower tuning speed.

Tunable Filters

Based on the mechanism of wavelength filtering, tunable filters can be classified into three categories: passive, active, and tunable optical amplifiers [Bra90, KC89].

- **Passive:** The passive category includes those passive wavelength-selective components which are made tunable by varying some mechanical element

Technology	Thermal Tuning	Mechanical Tuning	Acousto- Optic	Injection- Current
Tuning Range (nm)	2	55	83	10
Number of Channels	10	100	100s	10s
Tuning Speed	ms	ms	μ s	ns

Table 1.1: Comparison of tunable lasers.

of the filter such as mirror position. The well-known Fabry-Perot filter falls into this category. The advantages of such filters are their ready availability and their broad tuning range of hundreds of channels. Their tuning speed, however, is in the order of milliseconds, which is very slow, due to the mechanical movement involved.

- **Active:** In this category, we distinguish two types of filters based upon their mode coupling: either electro-optic, or acousto-optic. In both cases, a periodic perturbation is applied to the incident light, causing the propagating mode of the optical waves within a narrow wavelength range to transform to another mode. The acousto-optic tunable filter has a much broader tuning range than the electro-optic type, namely the entire 1.2-1.6 μm range versus 16 nm. Both types of filters can be tuned reasonably fast. The acousto-optic filter is limited to tuning times of a few microseconds due to the time required to set up the acoustic wave in the interaction region, whereas the electro-optic filter is much faster (typically in the nanosecond range), limited only by the time required to set up the electric field, .
- **Tunable optical amplifier:** The third category of filter is the laser-diode operating as a resonant amplifier which amplifies only those input signals whose frequencies coincide with those of the laser amplifier structure. A capacity of several tens of channels can be achieved, and the switching time is a few nanoseconds since frequency tuning is accomplished by current injection.

Table 1.2 presents a comparison of the major wavelength tunable filter technologies – Fabry-Perot, Acousto-Optics, Electro-Optics, and Semiconductors –

Technology	Fabry-Perot	Acousto-Optics	Electro-Optics	Active Semiconductors
Tuning Range (nm)	50	400	10	1 - 4
Number of Channels	100s	100s	10	10s
Tuning Speed	ms	μ s	ns	ns

Table 1.2: Comparison of tunable filters.

in terms of the relevant system parameters.

Most of the tunable lasers and filters described above were developed very recently and are still in the primitive stage. To fully tap the 30 Tb/s optical bandwidth potential requires tunable lasers and filters which can tune, in a few nanoseconds, to thousands of channels, each operating at a few Gb/s. Major technology breakthroughs are necessary to achieve this. Much more work still needs to be done to increase the tuning speed and the total number of channels, or tuning range, of these devices if even a small fraction of the usable fiber bandwidth is to be utilized for high-speed telecommunication applications.

1.4 Previous Related Work

Several introductory and survey papers on WDMA networks have been published [AK89][Goo89][Bra90][Muk92]. The previous work in this field can be divided into two categories: *Single-hop* networks and *multihop* networks [Muk92]. In a single-hop network, for any source-destination pair, there exists at least one common wavelength on which the source can transmit and the destination can receive. Packets are always transmitted directly from source to destination on the wavelength they share. In this way packets never go through any intermediate node, and therefore, this type of network is called a single-hop network. If, in a network, it is possible that packets may get routed through other intermediate nodes before reaching their destinations, it is called a multihop network.

1.4.1 Multihop Networks

In [Aca87], Acampora proposed a multichannel multihop lightwave network which can achieve the concurrency needed to tap the vast optical bandwidth in a fully distributed lightwave network. Figure 1.4 shows a network with eight users accessing a unidirectional optical bus through Network Interface Units (NIUs) spread along the bus. The physical topology can also take a variety of other forms, e.g., star or tree. Sixteen channels are wavelength division multiplexed on the fiber, but each NIU has only two fixed-wavelength transmitters and two fixed-wavelength receivers, each operating at the maximum electronic speed. Figure 1.5 shows the logical connection pattern of NIUs corresponding to the assignment of transmit and receive wavelengths shown in Figure 1.4. There are two columns of four NIUs each, with a perfect shuffle interconnection pattern

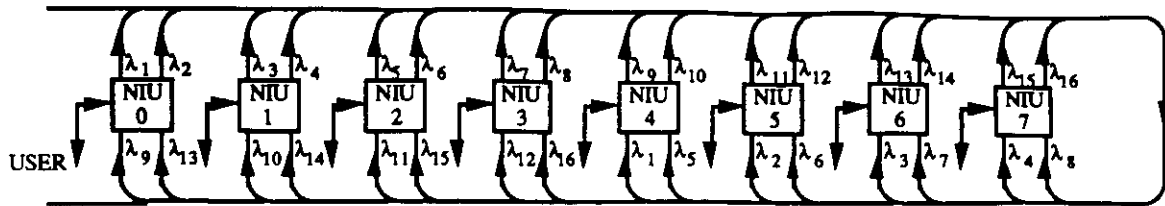


Figure 1.4: A multihop lightwave network with a bus topology.

[Sto71] between the columns (the third column represents the same NIU's as the first column, but is drawn separately to simplify the diagram). Since each NIU can access only a small, fixed set of the WDM channels, packets may need to be routed through intermediate NIUs to reach their destinations. In other words, packets may require multiple hops on different wavelengths to reach their destinations; the user interfaces serve as opto/electro/opto repeaters.

The multihop network was later generalized and called ShuffleNet in [HK88]. In general, the (p, k) ShuffleNet consists of $N = kp^k$ (k and p are positive integers) NIUs arranged in k columns of p^k NIUs each, with the k th column connected to the first. Each NIU is equipped with p fixed-wavelength transmitters and p fixed-wavelength receivers. The connectivity between successive columns is a p -shuffle [Pat81], which is a generalization of the ($p=2$) perfect shuffle and is analogous to the shuffling of p decks of cards.

A ShuffleNet multihop network has a number of important properties and desirable features. First, a typical packet needs only $\sim \log N$ hops before reaching its destination, so the total traffic capacity of the network increases with N roughly by $N/\log N$. It also leads to simple addressing and routing schemes, and it provides the capability of alternate routing in response to congestion and network failures [KS88]. In [EM88], the effect of traffic pattern variability

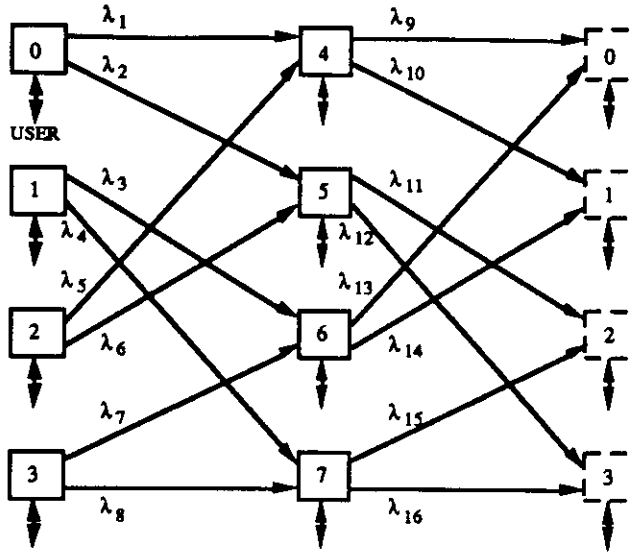


Figure 1.5: 8-NIU logical connectivity graph.

on ShuffleNet is investigated, and it is found that the throughput is reduced by 30 to 50 percent relative to the balanced-load situation, and the amount of throughput reduction generally is not sensitive to the number of stations.

There is no *a priori* reason to be restricted to the shuffle interconnection pattern. Two other papers [BFG90, LA90] propose schemes to optimize the logical connectivity by (slowly) retuning the transmitters and receivers of the stations adaptively to the traffic. Both papers formulate the problem to minimize the mean packet delay, and adopt some heuristics to solve it. For the uniform traffic case, the shuffle pattern is the best solution. For nonuniform traffic, however, different traffic matrices result in different optimal connectivity patterns.

The multihop networks also have some drawbacks. First, each NIU must process transit traffic as well as its own traffic, which may cause large queueing

delays at the NIU. Second, the number of wavelengths required is proportional to the number of nodes, which limits the maximum number of nodes that can be attached to the network. Third, in multihop networks one hop adds a propagation delay from one node to another. Since the node-to-node propagation delay may be very large compared to the packet transmission time in high-speed networks, traveling multihops may incur a very large delay.

1.4.2 Single-Hop Networks

In single-hop networks, traffic is not routed through any intermediate node. As a result, a significant amount of dynamic coordination of wavelengths may be required. Thus, the main challenge in single-hop network design is to develop efficient access protocols for coordinating data transmissions.

A simple mechanism allowing single-hop communications is to use a fixed assignment scheme [GC89] [CG90] where time is divided into cycles consisting of a number of slots. In each slot it is predetermined which pair of users are allowed to communicate over which channel. Thus, no destination collision or conflict exists in such a system. The limitation of this scheme is that it is not adaptive to dynamic traffic requirements. Also the packet delay at light loads can be large. The above work has been extended in [GG92] such that, given a traffic demand matrix, a scheduling algorithm is executed to produce a time slot assignment matrix which results in the minimal cycle length.

A slotted ALOHA and random TDMA protocol were proposed in [GK91], where the tuning range of a node is assumed to be limited. In the slotted ALOHA protocol, a node randomly selects a wavelength that it can transmit on and that the destination can receive on, and transmits the packet on that

wavelength. Collisions may occur at the destination, and retransmissions are required in this case. In the random TDMA protocol, it is assumed that all nodes are equipped with the same random number generator and the same starting seed. For each wavelength, a random number is generated to decide which node has the right to transmit on that wavelength in a slot. All transmissions are conflict-free following the random TDM protocol. The slotted ALOHA protocol results in lower delays for low system loads, while the random TDMA protocol performs better for moderate to high system loads. The observations are consistent with the single channel behavior of these two protocols [Kle76].

A Dynamic-Time WDMA (DT-WDMA) protocol is presented in [CDR90] for a passive star network. N stations are assumed to be attached to the star and $N + 1$ wavelengths are available for communications. One of the wavelengths serves as the control channel, and the other N channels are for actual data traffic, with each node transmitting on a unique channel. Time is divided into slots of fixed size equal to a packet length. The slot on the control channel is further divided into N minislots, one for each node. When a node has a packet to send, it first writes a control packet containing the destination address in its minislot on the control channel, and then transmits the data packet on its unique wavelength at the beginning of next slot. All stations continuously monitor the activities on the control channel. Upon recognizing its address in a minislot, the station tunes its receiver to the corresponding wavelength to receive the data packet from the beginning of next slot. In case multiple packets are addressed to the same destination in a slot (i.e., a destination conflict), the destination node checks the priority fields in the corresponding minislots and tunes to receive the one with the highest priority. The performance of the

protocol is analyzed and the maximum aggregate throughput is about $0.6 * N$ for large N . The drawback of this system is that the number of wavelengths required is one more than the number of stations and the slot size must be large enough to contain N control packets.

Two extensions [CF91][CY91] to the DT-WDMA protocol have been reported. In [CF91] an optical delay is used at each receiver to buffer those packets which would otherwise be lost in a destination conflict. The destination node can retrieve them in later time slots. In [CY91] new data packets upon generation will be announced in the minislots on the control channel and made known to all nodes in the network after a round-trip propagation delay. A distributed algorithm is then executed by all nodes to efficiently schedule conflict-free transmissions. Both these modifications are shown to perform better than DT-WDMA.

In [HKS87, Meh90] a family of single-hop access protocols is proposed for high-speed optical fiber LANs with a passive star coupler. Each station has one tunable transmitter and one tunable receiver, which are each capable of rapid tuning to any of the wavelengths in the system. Again, one channel serves as the control channel and the others are data channels. The transmitting user first sends a control packet over the control channel, containing the destination address and the wavelength to be used for data transmission, and then immediately tunes its transmitter to the chosen data channel and transmits the data packet. The length of a control packet is typically much smaller than the actual data packet length. To receive a packet, an idle receiver is always tuned to the control channel to listen for its address. Upon recognizing its address, it will tune to the transmission wavelength to receive the data packet.

The typical protocol in this family can be denoted by x/y , which indicates that all the users execute the access scheme x over the control channel and the access scheme y over the data channels. Several protocols such as *ALOHA/ALOHA*, *ALOHA/CSMA*, *CSMA/ALOHA*, and *CSMA/M-Server Switch* were proposed and analyzed. However, the analysis only applies to small area networks where the number of users is infinite and the packet transmission time is larger than the end-to-end propagation delay, which is unlikely to be practical in high-speed networks. Also the effect of destination conflicts is not included in the analysis.

All the previous protocols are based on the star coupler topology. However, in a star network, a pair of optic fibers is needed to connect each station to the star coupler, and deploying these fibers for a large number of stations over a large geographical area may be cost-prohibitive. In addition, a large star coupler is usually composed of many stages of individual (2×2) couplers, and large size star couplers are still very expensive because an $N \times N$ star requires $N \log_2 N/2$ (2×2) stars [Bra90]. Therefore, for networks supporting a large number of stations with a large geographical coverage, the linear bus topology is a favored alternative. The major drawback of the linear bus network is the 3-db power loss (due to power splitting) at each tap. However, recent progress on the erbium-doped optical amplifier [Bel91] will help alleviate this problem and make the linear fiber bus more attractive.

Two protocols for linear optical buses, namely AMTRAC [CG88] and multi-channel p_1 -persistent [BM92], have been proposed. Both protocols assume tunable transmitters and fixed receivers at each node. The AMTRAC tries to take the combined advantage of multichannel and train-oriented protocols [TBF83].

When a node has a packet to send, it first tunes its transmitter to the channel its destination is attached to, and executes the attempt-and-defer transmission by sensing the activity on that channel. To ensure fairness a cyclic structure is imposed on the system and only at certain scheduling points in a cycle is a node allowed to start its attempt to transmit. In the multichannel p_i -persistent protocol, time is slotted and node i chooses channel c with probability P_{ic} at the beginning of a slot. After selecting a channel, it then transmits the packet (if any) at the head of buffer for the chosen channel in the next slot arriving on that channel if the slot is free. The proper transmission probabilities, P_{ic} , have been analytically obtained for various fairness criteria in [BM92].

1.4.3 Experimental Systems

The field of *dense* WDM (channel spacing in the order of 1 nm) really began in the experiments of British Telecom Research Laboratory (BTRL) [PS85] and AT&T Bell Labs [OHLJ85] in 1985. In the BTRL experiment the concept of a multiwavelength star network was first introduced. The star size is 8×8 , and mechanically tunable filters are used at each receiver. The closest channel spacing demonstrated was 15 nm, limited by both the laser and filter linewidths. The Bell Labs experiment was the first which demonstrated channel spacing on the order of 1 nm. It was a point-to-point experiment where 10 channels spaced at 1.3nm apart and the operating bit rate was 2 Gb/s per channel, achieving a 1.37 Tb-km/s bandwidth-distance product.

The Bellcore LAMBANET [GKV⁺90] is a multiwavelength network with an 18×16 star coupler. Each node transmits on a unique wavelength, and each node has a wavelength demultiplexer followed by an array of up to 18 fixed

optical receivers. In an experiment 18 wavelengths were used, each running at 1.5 Gb/s with a transmission distance of 57.8 km. This in effect achieves a point-to-point 1.56 Tb·km/s bandwidth-distance product.

The Fast Optical Cross Connect (FOX) [ACG⁺88] was another experimental architecture at Bellcore, with the objective of providing a high-speed cross connect which allows processors to talk to the shared memory modules in a multiprocessor system. The cross connect includes two star networks: one for signals from processors to memories and the other for signals from memories to processors. All processors and memory modules are equipped with tunable transmitters and fixed receivers each at a unique wavelength. To address a memory, the processor tunes to the wavelength of that memory's filter and transmits its request, and vice versa in the opposite direction. Destination collision occurs in this cross connect, and a binary exponential-backoff algorithm was employed to resolve the contention.

In a further extension of the FOX, a Hybrid Packet Switching System (HYPASS) [AGKV88] was proposed at Bellcore. Two star networks are also used in HYPASS, but one is dedicated to the transport of data from input ports to output ports, and the other is for conveying output port status information back to the input ports. Each input port has one high-speed tunable transmitter and one high-speed tunable receiver. Each output port has one fixed transmitter and one fixed receiver. When an input port has a packet to transmit, it first tunes its receiver to the transmitting wavelength of the desired output port, and listens for status information. When an output port is ready to receive a packet, it transmits a polling signal. Upon hearing the signal, those input ports which have packets addressed to the polling port will transmit. A tree-polling

algorithm is used to resolve the output port contention. Under uniform traffic, the throughput of the HYPASS switch was found to saturate at 0.31 packets per time-slot per port.

IBM's Rainbow [DGL⁺90] research prototype takes the form of a direct-detection, full-duplex circuit-switched MAN backbone in which 32 IBM PS/2's transmitting at 300 Mb/s are interconnected to a 32×32 star coupler. Each station has a fixed transmitter at a unique wavelength, and a Fabry-Perot filter tunable in hundreds of microseconds. An in-band receiver polling algorithm is employed in which the transmitting station continually sends out a connection setup request containing the destination receiver's ID. At the same time, the transmitting station tunes its receiver to the desired destination's transmitter to wait for a setup acknowledgment. An idle receiver continuously scans all channels to see if any transmitter wants to set up a connection with it. After hearing the setup request, the destination station will transmit an acknowledgment, and thereby establish the circuit. This mechanism may not be suitable for packet-switched traffic because of its long setup delay. The Rainbow-I prototype has been demonstrated at Telecom '91 in Geneva, which supports three workstations (which are connected to two FDDI networks 10 km apart by an IBM-owned dark fiber link) transmitting at 270 Mb/s per station. The ultimate goal of the Rainbow project is to realize a 1000-station, 1 Gb/s per station, packet-switched MAN. A recent paper [JRS92] contains information on Rainbow prototypes planned in the future.

1.5 Summary of Results

So far we have given an introduction and a review of WDMA networks as well as a description of previous work in this field. We can see that many WDMA network designs, including both single-hop and multihop networks, have been proposed to tap more out of the optical bandwidth. A number of experimental systems have also been constructed to demonstrate the feasibility and to investigate the technology limitations of WDM communication systems.

To examine the effect of resource contention (of transmitters, receivers, and wavelengths), in Chapter 2, we present a mathematical model which approximates WDMA networks with a general receiver/transmitter configuration and an arbitrary traffic pattern. The model ignores any specific media access protocol by assuming that each station has *perfect* knowledge of all the resources in the system. Packets which cannot be transmitted upon arrival are blocked (i.e., lost) immediately, i.e., no storage buffers are available at the nodes. The average number of packets in transmission in the system is the performance measure used to compare various systems. Our objective is to examine the effect of changing the number of tunable and fixed transmitters and receivers. We first study the case of a uniform traffic matrix and observe that, when the number of wavelengths is fewer than the number of stations, it is better to have both tunable transmitters and tunable receivers, rather than having only either one of them tunable. Furthermore, we find that only a small number of tunable transmitters and receivers per station is needed to produce performance close to the upper bound. We then construct a general traffic model and propose an iterative solution procedure. A case of hot-spot traffic, where a heavy fraction of traffic goes to a specific node, is studied using this model. We find that

adding more resources to the hot-spot node will help improve its performance, but only to a limited extent determined by the traffic imbalance. The match between the model and simulation results are shown to be excellent.

In Chapter 3, we drop the “perfect knowledge” assumption and propose a multiple access protocol for a system consisting of many high-speed bursty traffic stations interconnected via an optical passive star coupler. One channel serves as the control channel, and the others are referred as data channels, as has been the case for several protocols in the literature. Each station has access to all wavelengths with its single tunable transmitter and single tunable receiver. Time is divided into fixed-sized slots. A slot on the control channel is further divided into two subparts: a reservation subpart consisting of a number of minislots each operating as a slotted ALOHA channel, and a tuning subpart consisting of a number of minislots for announcing destination addresses of transmitted packets. When a station has a packet to transmit, it must first contend on those slotted ALOHA channels to reserve a wavelength in a future slot, and then transmits the data on that wavelength when the reserved slot comes by. Broadcast and multicast traffic can also be easily supported with this protocol. We analyze the performance of the protocol for the infinite population case first, and observe the instability of the protocol owing to the slotted ALOHA component. We next set up a Markov chain model to analyze the finite population case. Because the state space is so large, we resort to the Equilibrium Point Analysis (EPA) technique to set up equilibrium point equations and solve for the approximate throughput and mean packet delay. Numerical results show that it is possible to achieve low delay and high throughput (larger than the electronic speed of a single station). We also examine the effect of destina-

tion conflicts and find that they are not a serious concern when the number of wavelengths is much fewer than the number of stations (which is the more likely case in the nearterm). The analysis also shows that the best performance is obtained when the capacities of the reservation channels and the data channels are balanced.

Considering the limitations of star networks in supporting a large number of stations in a large area and conjecturing that networks with dual bus topologies will be very popular in the future because of the dual bus topology in the IEEE 802.6 MAN standard (the DQDB [NBH88]), in Chapter 4 we propose a WDMA protocol for a dual bus network. In this system stations are attached to two fiber buses running in opposite directions. Stations transmit their traffic using the appropriate bus. Again, among all the available wavelengths, one is designated for control purposes and the rest are for data traffic. Time is slotted, and a slot on the control channel is further divided into a number of minislots as large as the number of data channels. The position of a minislot uniquely defines a data channel. Stations execute the DQDB protocol in order to get access to a minislot and the associated wavelength. Destination conflicts are resolved by comparing timestamps of the conflicting packets and the results are announced on the opposite bus. An approximate queueing model is built to analyze the system's performance and its correctness is verified by comparing with simulations results. Numerical experiments show that higher throughput can be achieved with more wavelengths; the efficiency per wavelength, however, drops because with more wavelengths, the chance of a destination conflict increases and more retransmissions are required. We also observe that our protocol achieves better fairness than single channel DQDB because upstream nodes must wait longer

for the acknowledgement to travel back, which somewhat offsets their advantage of accessing idle slots sooner than downstream nodes.

Chapter 5 gives the conclusions and some directions for future research.

CHAPTER 2

A Performance Model of WDMA Networks

2.1 Introduction

In this chapter we present a mathematical model for WDMA networks to examine the effects of resource contention (of transmitters, receivers, and wavelengths) under general traffic patterns. Ramaswami and Pankaj [RP90] compared the three cases of having either tunable transmitters only, or tunable receivers only, or both, assuming each station is equipped with only one transmitter and one receiver. Chlamtac and Ganz [CG89] discussed the design alternatives of WDMA star networks where each station can have multiple transmitters and receivers and some finite buffers. Both of these two previous studies were conducted only for the case of a uniform traffic matrix. In our model we assume that each station has its own hardware configuration and traffic requirement. The model ignores any specific media access protocol by assuming that each station has perfect knowledge of the current status of all the resources in the system. This assumption is reasonably good for the case of a packet switch where the physical distance is small and stations can learn the status of the resources from information broadcast by a centralized controller. The model serves as an upper bound on performance when the system is a network which covers a larger geographical area.

The rest of the chapter is organized as follows: In Section 2.2, we describe the system configuration and assumptions to be used in the mathematical model. Section 2.3 presents the analysis of networks with stations having multiple transmitters and receivers for the uniform traffic case. A general model is constructed in Section 2.4 and an iterative procedure is proposed to solve it for the general traffic case. In Section 2.5, a hot-spot traffic case is then studied using the general model. Section 2.6 summarizes the results.

2.2 The System Model and the Solution Method

The system considered here consists of N stations attached to a broadcast medium (fiber bus or star coupler). The number of wavelengths is equal to W . Node i has t_i transmitters and r_i receivers, each of which may be tunable to any wavelength or which may be tuned to a single fixed wavelength. We assume that a stream of packets arrive to node i following a Poisson process with rate λ_i packets per unit time. Ours is a continuous time model in which packet lengths are exponentially distributed with mean $1/\mu$, the same for all nodes. We choose the average packet length as the time unit of the system by setting $\mu = 1$ throughout the chapter. A packet arriving at node i is addressed to destination node j with probability x_{ij} , $1 \leq i, j \leq N$. Define $\phi_i \triangleq \sum_{j=1}^N \lambda_j x_{ji}$, $1 \leq i \leq N$ as the intensity of generated traffic that is destined for node i . For a packet to be transmitted and successfully received, the three following conditions must all be satisfied simultaneously: (i) there is a free wavelength in the system, (ii) there is a free transmitter, at the source node, which can access that free wavelength, and (iii) there is a free receiver, at the destination node, which can also access that same free wavelength. We assume there is no buffering at any

node. Upon a packet's arrival, it is transmitted immediately if all the three conditions above are true (remember that we have assumed a "perfect" access scheme); otherwise the packet is blocked (i.e. lost) immediately. We assume that each station has complete knowledge of the status (busy or idle) of all the wavelengths, transmitters, and receivers in the system. The throughput of the system, which is defined as the average number of successful packets transmitted per unit time, will be used as the performance measure to compare systems with different configurations and different traffic patterns.

Let the random variable K be the number of busy wavelengths in the system in steady state. Let $p_k \triangleq \text{Prob}[K = k], 0 \leq k \leq W$. Knowing the number of busy wavelengths does not completely describe the state of the system since we also need the current status of the transmitters and receivers of each node. However, we will make the approximation that K is a Markov chain. In this analysis, we will also approximate many of the transition rates of this chain and then provide an exact solution under these approximations. Given that the system is in state $k, 0 \leq k \leq W - 1$, and given a specific free wavelength, we define $\alpha_k^{(i)}$ as the probability that an arriving packet at node i finds at least one of its transmitters free which can access that free wavelength, and $\beta_k^{(j)}$ as the probability that a packet destined for node j arriving at a source node finds, upon its arrival, a free receiver at node j which can access that same free wavelength. We recognize that these two probabilities should properly be computed as a joint probability; we choose to approximate them by assuming independence of the underlying events. Let σ_k denote the transition rate from state k to state $k + 1$ due to the transmission of a new packet. We first note that $\lambda_i x_{ij}$ is the rate of new packets generated by node i and addressed to

node j . The probability that this new packet is successfully transmitted is approximately equal to $\alpha_k^{(i)}\beta_k^{(j)}$. Therefore, under the assumption that all the free wavelengths are equally favored for the transmission of a new packet, σ_k can be calculated as follows:

$$\sigma_k = \sum_{i=1}^N \sum_{j=1}^N \lambda_i x_{ij} \alpha_k^{(i)} \beta_k^{(j)} \quad 0 \leq k \leq W - 1 \quad (2.1)$$

We see that the evolution of K forms a Markov chain which is a birth-death process whose state transition diagram is shown in Figure 2.1. Solving this birth-death process [Kle75], we have

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\sigma_i}{(i+1)\mu} \quad (2.2)$$

where

$$p_0 = \left[1 + \sum_{k=1}^W \prod_{i=0}^{k-1} \frac{\sigma_i}{(i+1)\mu} \right]^{-1} \quad (2.3)$$

The throughput of the system, S , which is also equal to the average number of busy wavelengths in the system, can be calculated by

$$S = \sum_{k=0}^W k p_k \quad (2.4)$$

This, then, is the general setup for our solution. It remains to find σ_k and hence S . This we do in the next two sections.

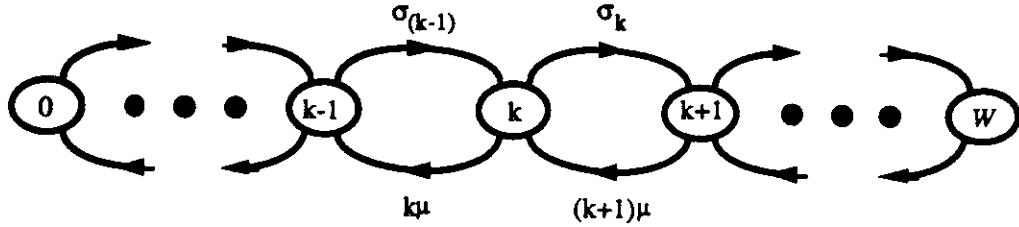


Figure 2.1: State transition diagram for number of busy wavelength in the system.

2.3 The Uniform Traffic Case

In this section we study the uniform traffic case where packets arrive to a station following a Poisson process with rate λ packets per unit time (the same for all stations). A packet will travel from its source station to any of the N stations (including the source itself) with equal probability, i.e., $x_{ij} = \frac{1}{N}, 1 \leq i, j \leq N$.

2.3.1 Tunable Transmitters and Receivers

Here we consider the case where each node is equipped with q ($q \leq W$) tunable transmitters and q tunable receivers, each of which can tune to any of the W wavelengths. The $\alpha_k^{(i)}$'s and $\beta_k^{(i)}$'s are now the same for all stations by symmetry, which we denote by α_k and β_k , respectively. To get the α_k and β_k , we first note, given that the system is in state k , that it is implied that there are also k transmitters and k receivers currently busy in the system. When $k < q$, α_k (β_k) is equal to one because there must be always a free transmitter (receiver) at any node. For the cases $k \geq q$, since there is a total of Nq transmitters (receivers) in the system, we know that the probability that any

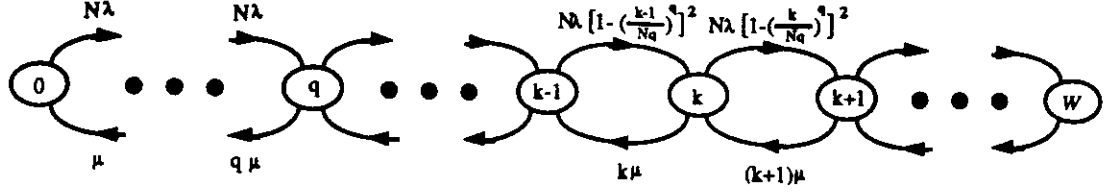


Figure 2.2: State transition diagram for tunable transmitters and receivers.

single transmitter (receiver) is busy equals k/Nq . Therefore, the probability that all the transmitters (receivers) of a given node are busy is approximately equal to $(k/Nq)^q$. One minus this gives us α_k (β_k). Thus, we have the following approximation

$$\alpha_k = \beta_k = \begin{cases} 1 & 0 \leq k \leq q-1 \\ 1 - (\frac{k}{Nq})^q & q \leq k \leq W-1 \end{cases} \quad (2.5)$$

The transition rates σ_k can be calculated using Equations (2.1) and (2.5), and the corresponding state transition diagram is shown in Figure 2.2. Solving this Markov chain, we get

$$p_k = p_0 \frac{(N\rho)^k}{k!} \quad 0 \leq k \leq q$$

$$p_k = p_0 \frac{(N\rho)^k}{k!} \prod_{i=q}^{k-1} \left[1 - \left(\frac{i}{Nq} \right)^q \right]^2 \quad q+1 \leq k \leq W$$

where $\rho = \lambda/\mu$ and

$$p_0 = \left[\sum_{k=0}^q \frac{(N\rho)^k}{k!} + \sum_{k=q+1}^W \frac{(N\rho)^k}{k!} \prod_{i=q}^{k-1} \left(1 - \left(\frac{i}{Nq} \right)^q \right)^2 \right]^{-1}$$

The throughput S can be calculated from Equation (2.4).

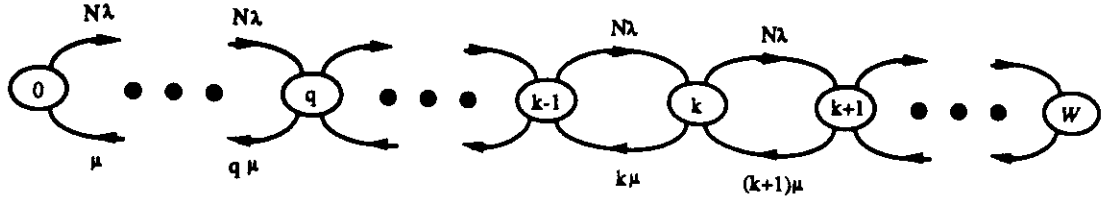


Figure 2.3: An upper bound, the W -server loss system.

An *achievable* upper bound on the throughput can be obtained by assuming all nodes have W tunable transmitters and receivers. In this case, $\alpha_k = \beta_k = 1$, which corresponds to a W -server loss system [Kle75] where each wavelength corresponds to a server. Figure 2.3 shows the corresponding state transition diagram. Solving this, we have

$$p_k = p_0 \frac{(N\rho)^k}{k!} \quad 0 \leq k \leq W$$

where

$$p_0 = \left[\sum_{k=0}^W \frac{(N\rho)^k}{k!} \right]^{-1}$$

The blocking probability of this upper bound system equals

$$p_W = \frac{(N\rho)^W / W!}{\sum_{k=0}^W (N\rho)^k / k!}$$

which is the well-known *Erlang B formula* [Kle75].

In Figures 2.4 and 2.5 we plot the throughput versus the total offered load for $N=50$, $W=10$ and $N=50$, $W=50$, respectively. We show the usual (ideal) upper bound on throughput as equal to the input load up to the point where

the load equals the total system bandwidth; beyond that point, any additional traffic is clearly lost. We see that a small q (much smaller than W) is enough to produce a result close to the upper bound. This is because, in the uniform traffic case, the probability that more than a few packets are going to the same destination at the same time is very small, and only a small number of transmitters and receivers are required at each node.

2.3.2 Tunability on One Side Only (Transmitters or Receivers)

In this section we consider the same uniform traffic case except that each station now has only tunable transmitters *or* receivers, *but not both*. We begin with the case where each node is equipped with one tunable transmitter and f ($f \leq W$) fixed tuned receivers. Each receiver in a station is tuned to a different fixed wavelength and the receivers in the whole system are tuned in a uniform way such that the number of receivers tuned to each wavelength is the same, which equals Nf/W (assumed to be an integer).

By the same arguments as in the previous subsection, α_k can be easily (but approximately) derived from Equation (2.5) by setting $q = 1$.

$$\alpha_k = 1 - \frac{k}{N} \quad 0 \leq k \leq W - 1$$

To get β_k requires a bit of different reasoning. For $k < f$, β_k equals one because the total number of busy receivers in the system is fewer than the number of receivers each station has. To transmit a new packet, the source node can just tune its transmitter to the free wavelength of any idle receiver at the destination. For the case $k \geq f$, recall that all the receivers are tuned in a uniform way over all the wavelengths; therefore, we know that, given that the system is in state

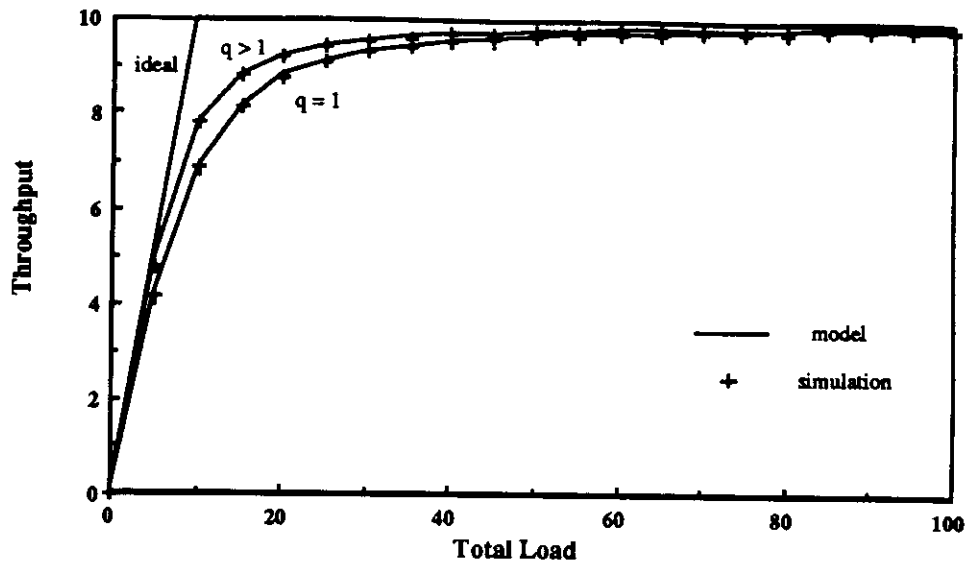


Figure 2.4: Throughput versus total load $N\rho$. $N=50$, $W=10$.

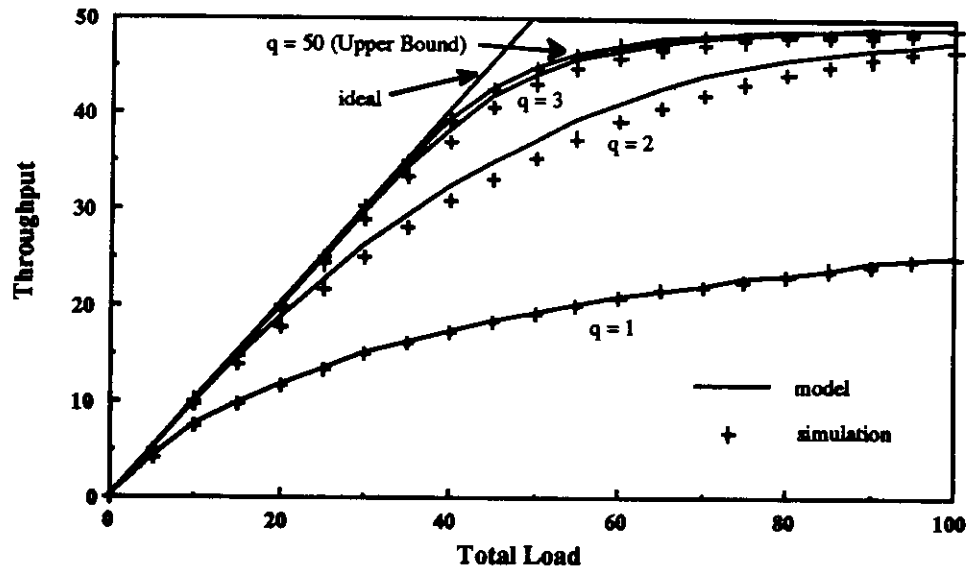


Figure 2.5: Throughput versus total load $N\rho$. $N=50$, $W=50$.

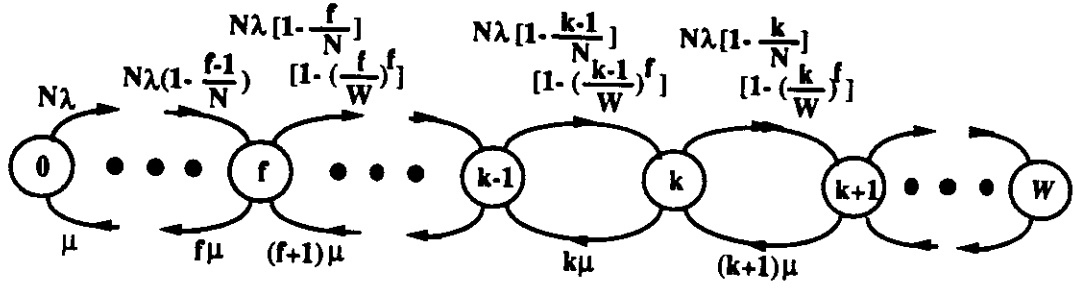


Figure 2.6: State transition diagram for tunability on one side only.

k (i.e., there are currently k busy wavelengths), the probability that the fixed wavelength at an arbitrary receiver at the destination is busy equals k/W . The probability that wavelengths at the receivers of a given node are all busy is approximately $(k/W)^f$, and one minus this gives us β_k as follows:

$$\beta_k = \begin{cases} 1 & 0 \leq k \leq f-1 \\ 1 - (\frac{k}{W})^f & f \leq k \leq W-1 \end{cases}$$

By switching the roles of transmitters and receivers in the discussion above, we can easily obtain the α_k and β_k for the case of multiple fixed transmitters and one tunable receiver per node, which are equal to the β_k and α_k listed above, respectively; the two systems are “duals” of each other. Therefore, the state transition diagrams of those two cases are exactly the same and is shown in Figure 2.6. The exact solution to this approximate Markov chain is

$$p_k = p_0 \frac{(N\rho)^k}{k!} \prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) \quad 1 \leq k \leq f$$

$$p_k = p_0 \frac{(N\rho)^k}{k!} \left[\prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) \right] \left[\prod_{i=f}^{k-1} \left(1 - \left(\frac{i}{W}\right)^f\right) \right] \quad f+1 \leq k \leq W$$

where $\rho = \lambda/\mu$ and

$$p_0 = \left[1 + \sum_{k=1}^f \frac{(N\rho)^k}{k!} \prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) + \sum_{k=f+1}^W \frac{(N\rho)^k}{k!} \left[\prod_{i=0}^{k-1} \left(1 - \frac{i}{N}\right) \right] \left[\prod_{i=f}^{k-1} \left(1 - \left(\frac{i}{W}\right)^f\right) \right] \right]^{-1}$$

The throughput can be calculated from Equation (2.4).

Figure 2.7 shows the case in which the number of wavelengths is small ($W = 10$) compared to the number of nodes ($N = 50$) in the system. We see that there is an interval in the light load range where multiple fixed receivers is better than one tunable receiver because not many wavelengths are in use and a station with multiple receivers can receive more than one packet at a time. However, as the load increases the average number of wavelengths in use increases too, and it is better to have a tunable receiver than multiple fixed receivers because the wavelengths those fixed receivers are tuned to may be all in use (by other stations) and a given station could not receive any packet even though not all of its receivers were busy. Figure 2.8 shows the case where $N = 50$ and $W = 25$ on a different scale. Once again we see the importance of going to a single tunable receiver at heavy load. When the number of wavelengths becomes the same as the number of nodes in the system ($W = N = 50$) as plotted in Figure 2.9, wavelength is no longer the scarce resource and the performance of having tunability on both sides is the same as on one side only. In this case having multiple fixed receivers is always better. Note the excellent match between the results from our approximations and simulations in the figures.

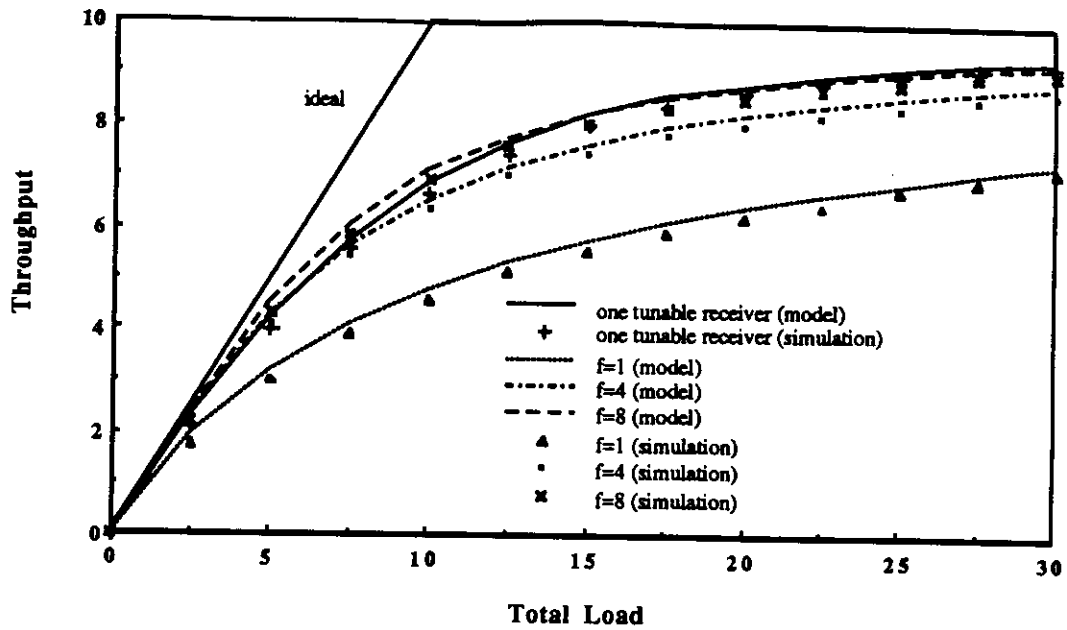


Figure 2.7: Throughput versus total load $N\rho$. $N=50$, $W=10$ and one tunable transmitter.

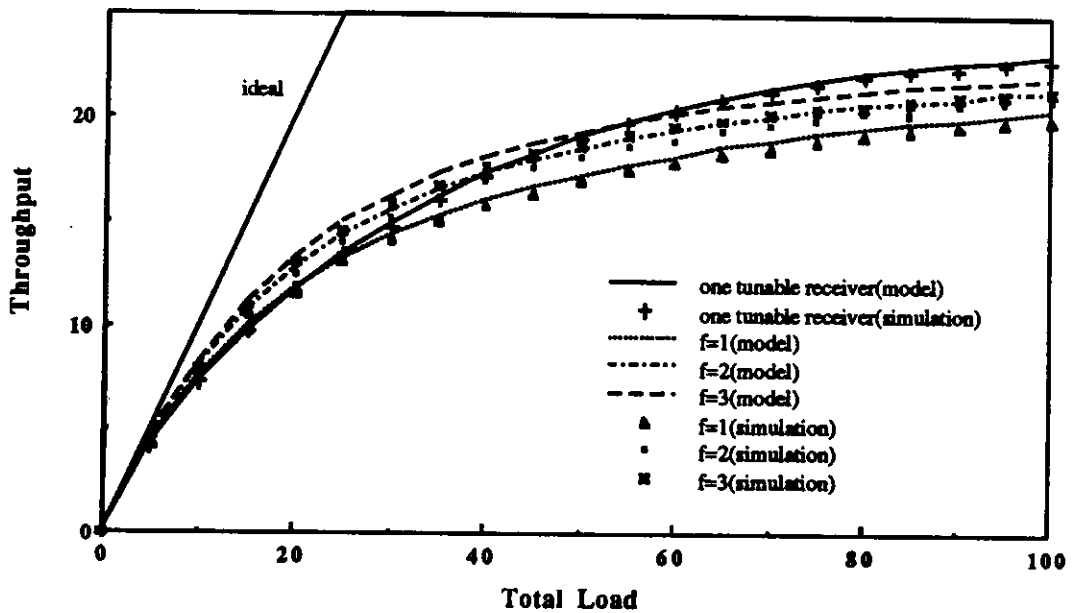


Figure 2.8: Throughput versus total load $N\rho$. $N=50$, $W=25$ and one tunable transmitter.

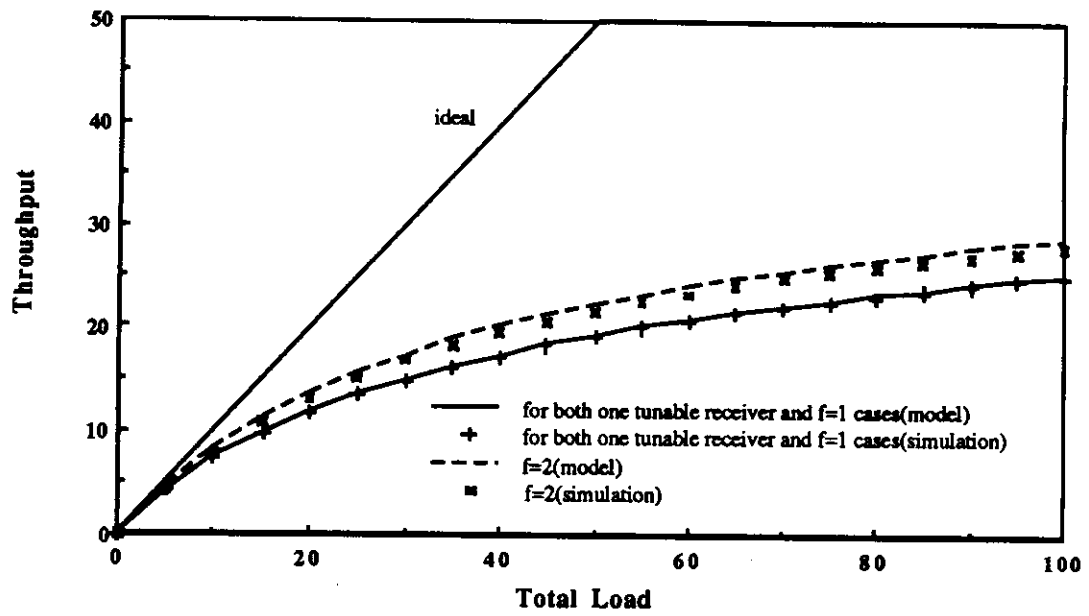


Figure 2.9: Throughput versus total load $N\rho$. $N=50$, $W=50$ and one tunable transmitter.

2.4 The General Traffic Case

2.4.1 The Model

Here we consider the general traffic case. We assume that node i has t_i tunable transmitters and r_i tunable receivers only. Let λ_i^* and ϕ_i^* denote the number of packets successfully transmitted and received by node i per unit time, respectively. Clearly, $S = \sum_{i=1}^N \lambda_i^* = \sum_{i=1}^N \phi_i^*$. $\alpha_k^{(i)}$ can thus be approximated as follows:

$$\alpha_k^{(i)} = \begin{cases} 1 & 0 \leq k \leq t_i - 1 \\ 1 - (k\lambda_i^*/t_i S)^{t_i} & t_i \leq k \leq \min(t_i S/\lambda_i^*, W - 1) \\ 0 & \min(t_i S/\lambda_i^*, W - 1) \leq k \leq W - 1 \end{cases} \quad (2.6)$$

The quantity $(k\lambda_i^*/S)$ is the average number of busy transmitters of node i , given that the system is in state k . $(k\lambda_i^*/t_i S)$ equals the probability that any single transmitter of node i is busy given that the system is in state k . Therefore, $(k\lambda_i^*/t_i S)^{t_i}$ is approximately equal to the probability that all of node i 's transmitters are busy, and one minus that gives us the $\alpha_k^{(i)}$. For those k 's where the value $(k\lambda_i^*/t_i S)$ is greater than one, we set the $\alpha_k^{(i)}$ to zero. The $\beta_k^{(i)}$'s can be derived in a similar way:

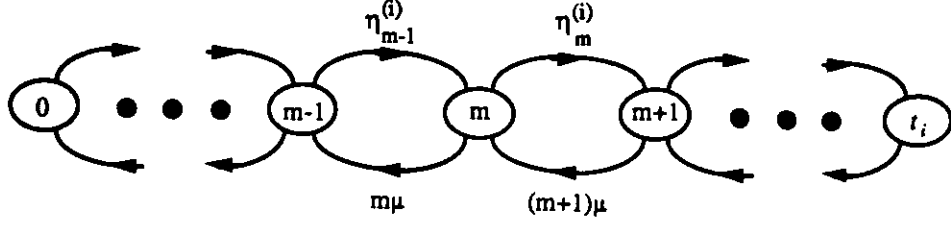


Figure 2.10: State transition diagram for $U^{(i)}$.

$$\beta_k^{(i)} = \begin{cases} 1 & 0 \leq k \leq r_i - 1 \\ 1 - (k\phi_i^*/r_i S)^{r_i} & r_i \leq k \leq \min(r_i S/\phi_i^*, W - 1) \\ 0 & \min(r_i S/\phi_i^*, W - 1) \leq k \leq W - 1 \end{cases} \quad (2.7)$$

Note that when the traffic is uniform, $\lambda_i^*/S = \phi_i^*/S = 1/N, 1 \leq i \leq N$ by symmetry, and Equations (2.6) and (2.7) both reduce to Equation (2.5).

We now derive λ_i^* . Let $U^{(i)}$ denote the number of busy transmitters of node i in steady state with probability mass function (pmf) $u_m^{(i)} \triangleq \text{Prob}[U^{(i)} = m]$. We will approximate $U^{(i)}$ as a Markov process. Define $p_{k|m} \triangleq \text{Prob}[K = k | K \geq m] = p_k / \sum_{j=m}^W p_j$. Let $\eta_m^{(i)}$ be the transition rate for $U^{(i)}$ from state m to state $m + 1$. $\eta_m^{(i)}$ can be approximated as follows:

$$\eta_m^{(i)} = \lambda_i \sum_{j=1}^N x_{ij} \sum_{k=m}^{W-1} \beta_k^{(j)} p_{k|m}$$

The transition rate from state m to $m - 1$ is just $m\mu$, the aggregate rate at which any busy transmitter of node i will finish its transmission first. Figure 2.10 shows the state transition diagram. Solving this, we have

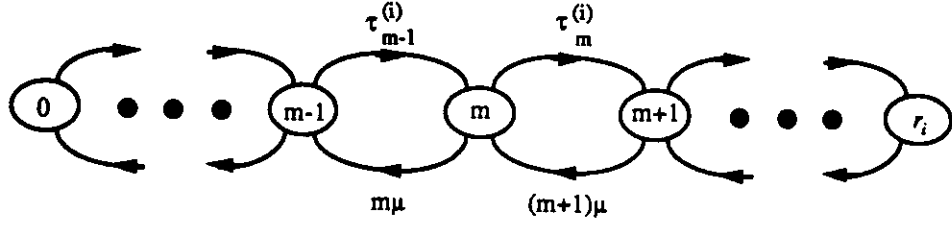


Figure 2.11: State transition diagram for $V^{(i)}$.

$$u_m^{(i)} = u_0^{(i)} \prod_{n=0}^{m-1} \frac{\eta_n^{(i)}}{(n+1)\mu}$$

where

$$u_0^{(i)} = \left[1 + \sum_{m=1}^{t_i} \prod_{n=0}^{m-1} \frac{\eta_n^{(i)}}{(n+1)\mu} \right]^{-1}$$

λ_i^* can be then obtained from

$$\lambda_i^* = \sum_{m=0}^{t_i} m u_m^{(i)} \quad (2.8)$$

The ϕ_i^* can be derived in almost the same way. Let $V^{(i)}$ denote the number of busy receivers of node i in steady state with pmf $v_m^{(i)} \triangleq \text{Prob}[V^{(i)} = m]$. Define $\tau_m^{(i)}$ as the transition rate for $V^{(i)}$ from state m to state $m+1$. $\tau_m^{(i)}$ can be calculated as follows:

$$\tau_m^{(i)} = \sum_{j=1}^N \lambda_j x_{ji} \sum_{k=m}^{W-1} \alpha_k^{(j)} p_{k|m}$$

The transition rate from state m to $m-1$ is just $m\mu$, the aggregate rate at which any busy receiver of node i will finish its reception first. Figure 2.11 shows the state transition diagram. Solving this, we have

$$v_m^{(i)} = v_0^{(i)} \prod_{n=0}^{m-1} \frac{\tau_n^{(i)}}{(n+1)\mu}$$

where

$$v_0^{(i)} = \left[1 + \sum_{m=1}^{r_i} \prod_{n=0}^{m-1} \frac{\tau_n^{(i)}}{(n+1)\mu} \right]^{-1}$$

ϕ_i^* can be obtained from:

$$\phi_i^* = \sum_{m=0}^{r_i} m v_m^{(i)} \quad (2.9)$$

However, we do not really have the p_k 's in the first place to compute those λ_i^* and ϕ_i^* because they depend on each other. In the next subsection, we propose an iterative procedure to solve for these steady state probabilities.

2.4.2 An Iterative Solution Procedure

We define $p_k(n)$, $\lambda_i^*(n)$, $\phi_i^*(n)$, $\alpha_k^{(i)}(n)$, and $\beta_k^{(i)}(n)$ as the values obtained for these quantities at the end of the n th iteration. We start with some initial estimates $p_k(0)$, $\lambda_i^*(0)$, and $\phi_i^*(0)$. One simple initial estimate is to set $p_k(0) = 1/(W+1)$, $0 \leq k \leq W$, $\lambda_i^*(0) = \lambda_i$ and $\phi_i^*(0) = \phi_i$, $1 \leq i \leq N$. The iterative procedure is as follows:

1. Let $n = 1$.
2. Construct $\alpha_k^{(i)}(n)$ and $\beta_k^{(i)}(n)$ from $\lambda_i^*(n-1)$ and $\phi_i^*(n-1)$ using equations (2.6) and (2.7). Solve for $p_k(n)$ from equations (2.1), (2.2), and (2.3).
3. With $p_k(n)$, $\alpha_k^{(i)}(n)$ and $\beta_k^{(i)}(n)$, solve the Markov chains in Figures 2.10 and 2.11 to get $\lambda_i^*(n)$ and $\phi_i^*(n)$.

4. If the difference between $p_k(n), \lambda_i^*(n), \phi_i^*(n)$ and $p_k(n-1), \lambda_i^*(n-1), \phi_i^*(n-1)$, respectively, are less than pre-specified thresholds, then stop. Otherwise, set $n = n + 1$ and go to step 2.

We do not have a proof of the convergence of the procedure above. However, for all the experiments presented in the next section, this procedure converged, and the solutions are very close to the simulation results.

2.5 The Hot-Spot Traffic Case

Here we use the general model just described to study the special case of a “hot-spot” traffic pattern where a large portion of traffic is addressed to a specific node called the hot-spot node. The other $N - 1$ nodes are called “plain” nodes. Without loss of generality, let node 1 be the hot-spot node. We assume all $\lambda_i = \lambda, 1 \leq i \leq N$. From the generated traffic from all the nodes, a fraction of b is assumed to go to the hot-spot node, and the rest goes to the other nodes uniformly, i.e., $x_{i1} = b, x_{ij} = (1 - b)/(N - 1), 1 \leq i \leq N, 2 \leq j \leq N$. Each node has one tunable transmitter and one tunable receiver except node 1, which may have more than one tunable receiver. That is, $t_i = 1, i = 1, \dots, N, r_1 \geq 1$, and $r_j = 1, j = 2, \dots, N$. The effect of various values of b and r_1 on the system performance is investigated below.

Figure 2.12 shows the relationship between the throughput and total load for the case of $N=50, W=10$, and $r_1=1$. We can see that as the bias, b , gets larger, the total throughput of the system is degraded. This is because, while the single receiver of the hot spot node is overloaded, there is not enough traffic generated for exchange among the other nodes.

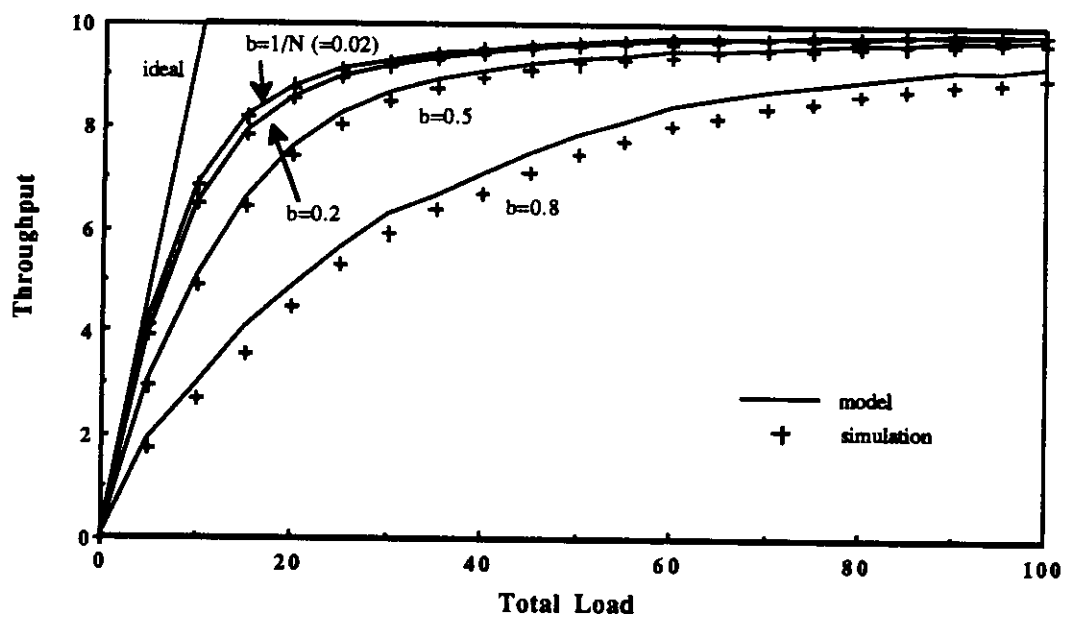


Figure 2.12: Throughput versus the total load $N\rho$. $N=50$, $W=10$, $r_1 = 1$.

Since the receiver of the hot spot node is now the scarce resource, we next study the effect of increasing the number of receivers at the hot-spot node. In Figures 2.13 and 2.14 we plot the received throughput (i.e., ϕ_1^*) of the hot-spot node (node 1 in our example) versus the total load for the cases of $N=50$, $W=10$, $b=0.2$ and $b=0.8$, respectively. We note that, by increasing the number of receivers at the hot-spot node, its throughput can be improved. However, as the load increases, we see that the received throughput of the hot spot node saturates at some value no matter how large a number of receivers it has. This is because when the total load is very heavy, the throughput of the system approaches W , and the received throughput of each node saturates at some value determined by the traffic imbalance. Putting in a lot more receivers at the hot-spot node will not help further increase its received throughput. To compute this saturated throughput for the hot-spot node (node 1), we let the load λ go to infinity. Define H as the number of busy receivers of node 1 in steady state with pmf $\pi_m \triangleq Prob[H = m], 0 \leq m \leq r_1$. The transition rate of H moving out of state m can be calculated as follows: We first note that as λ goes to infinity, there are W packets in transmission in the system all the time. Given $H = m$, we know that there are m packets going to node 1 and $W - m$ to the others. The number of busy receivers of node 1 will increase by one when the transmission of any of the $W - m$ packets addressed to the plain nodes is finished first (with rate $(W - m)\mu$) and the wavelength just freed is immediately grabbed by a new packet addressed to node 1 (packets arrive infinitely fast since $\lambda \rightarrow \infty$), the probability of which we denote by y_m . To compute y_m , we note that there are currently $(N - 1) - (W - m + 1) = (N - W + m)$ plain nodes whose receivers are idle. The probability that the next arriving packet is addressed to the hot-spot node is equal to

$$y_m = \frac{b}{b + (N - W + m) \frac{1-b}{N-1}}$$

Therefore, the rate of H moving from m to $m + 1$ equals $(W - m)\mu y_m$. On the other hand, H will decrease by one if the transmission of any of the m packets to node 1 finishes first (with rate $m\mu$) and the free wavelength is then occupied by a new packet addressed to a plain node, the probability of which is just $(1 - y_{m-1})$ because there are $(N - 1) - (W - m) = (N - W + m - 1)$ plain nodes whose receivers are idle. Thus, the rate of H moving from m to $m - 1$ is $m\mu(1 - y_{m-1})$. The state transition diagram is shown in Figure 2.15. Solving this, we have

$$\pi_m = \pi_0 \binom{W}{m} \prod_{j=0}^{m-1} \frac{y_j}{1 - y_j} \quad 1 \leq m \leq r_1$$

where

$$\pi_0 = \left[1 + \sum_{m=1}^{r_1} \binom{W}{m} \prod_{j=0}^{m-1} \frac{y_j}{1 - y_j} \right]^{-1}$$

The real received throughput of node 1 as $\lambda \rightarrow \infty$, S_1 , can be calculated from

$$S_1 = \sum_{m=0}^{r_1} m\pi_m$$

Figure 2.16 shows S_1 versus the number of receivers of node 1 for the case of $N=50$ and $W=10$. We see that, given an extremely heavy load and a large number of receivers, the hot-spot node can achieve a larger asymptotic throughput as the fraction of traffic addressed to the hot-spot node gets larger.

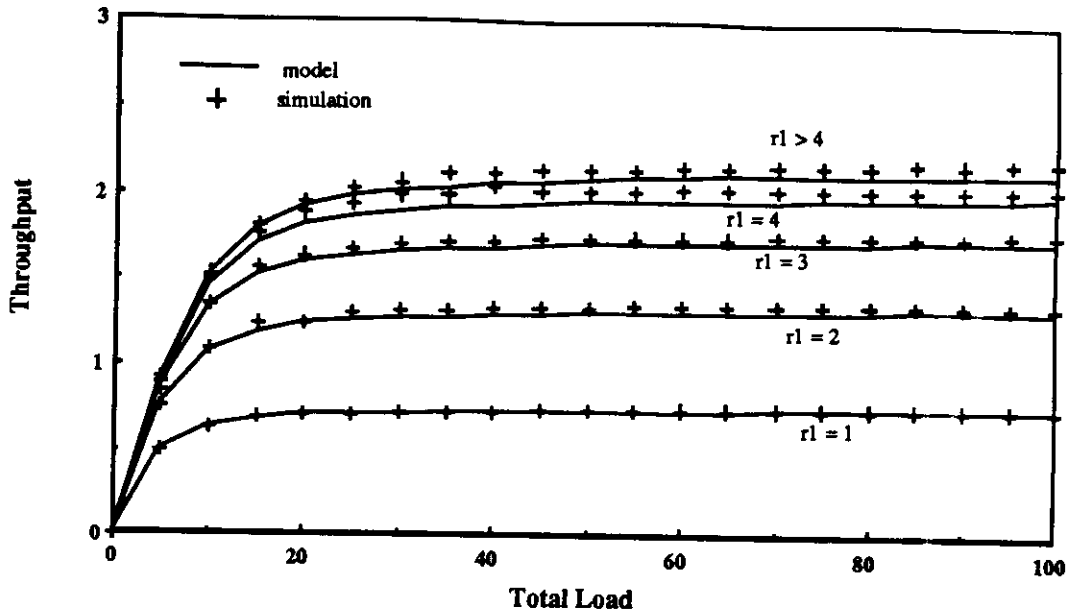


Figure 2.13: Throughput of the hot-spot node versus total load $N\rho$. $N=50$, $W=10$, $b=0.2$.

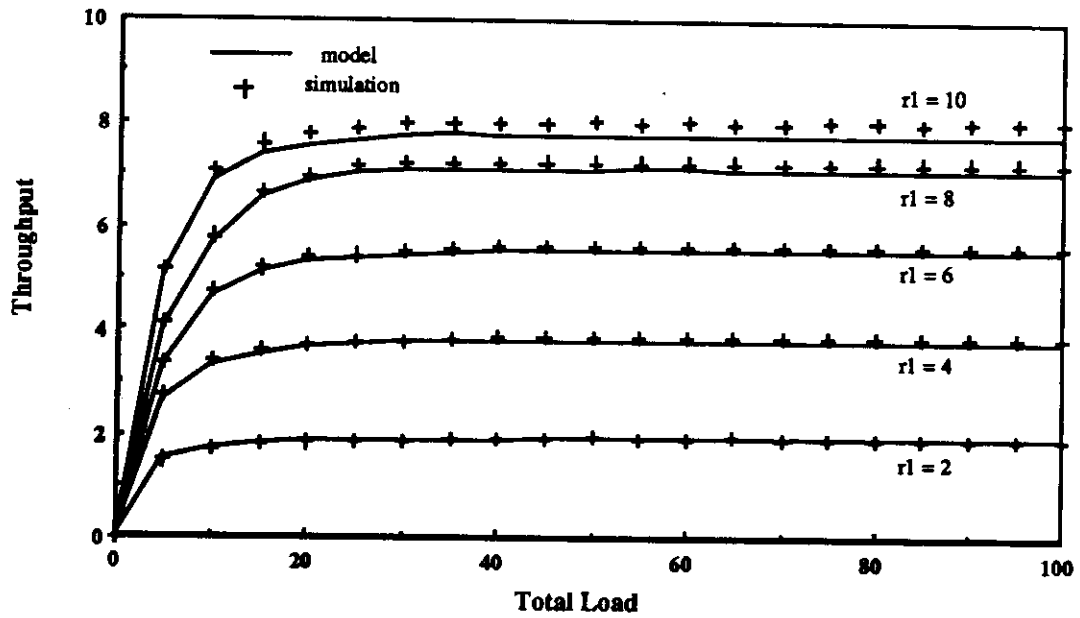


Figure 2.14: Throughput of the hot-spot node versus total load $N\rho$. $N=50$, $W=10$, $b=0.8$.

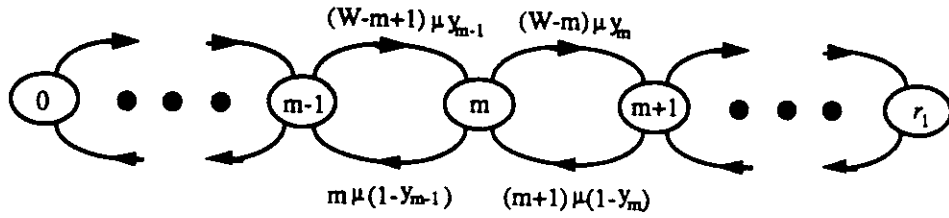


Figure 2.15: State transition diagram for H .

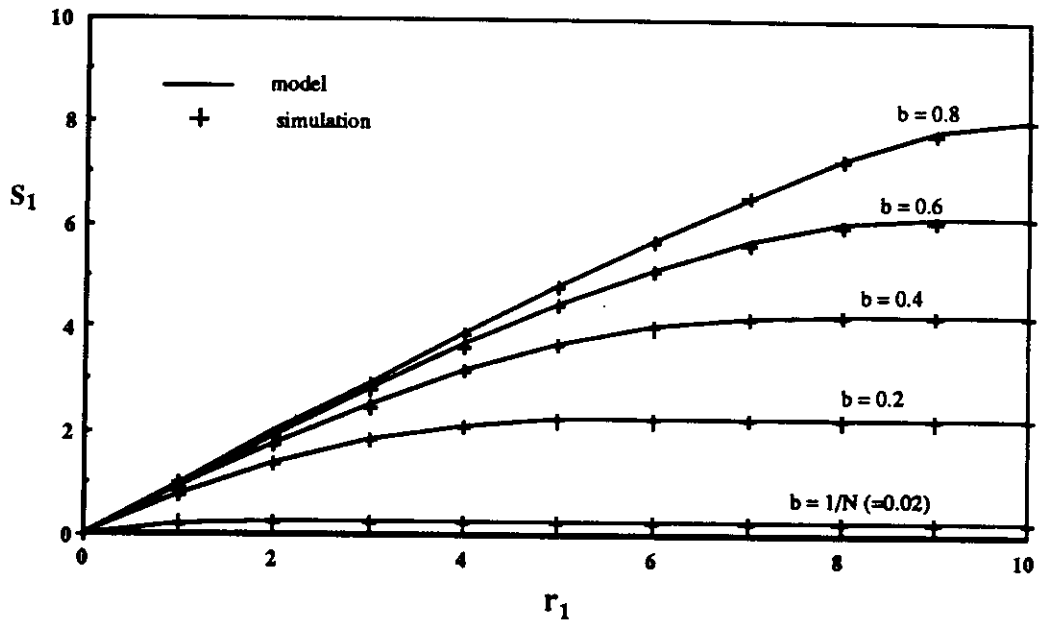


Figure 2.16: The relationship between the limiting value S_1 and r_1 . $N=50$, $W=10$.

2.6 Conclusions

We have presented a mathematical model which approximates WDMA networks. We first built a model to analyze the uniform traffic case. We found that it is better to have both tunable transmitters and tunable receivers than having only one or the other tunable when the number of wavelengths is smaller than the number of nodes (which is most likely the case in the near future [Bra90]). Also a small number of tunable transmitters and receivers at each station is enough to produce performance close to the upper bound. We then constructed a model for systems with general receiver/transmitter configurations and arbitrary traffic patterns. An iterative procedure was proposed to solve the model numerically. We used this model to study a special hot-spot traffic case. We saw that traffic imbalance could degrade the performance of the system. Adding more receivers to the hot-spot node helps improve its performance, but only to a limited extent determined by the traffic imbalance. The match between the results from our approximations and simulations was shown to be excellent.

CHAPTER 3

A WDMA Protocol for LANs with a Passive Star Topology

3.1 Introduction

In this chapter, we drop the assumption of perfect knowledge (i.e., perfect access) and propose a multiple access protocol for a system consisting of many high-speed bursty traffic stations interconnected via an optical passive star coupler as shown in Figure 3.1. There are $(W + 1)$ wavelengths available, w_0, w_1, \dots, w_W to serve N attached stations. The channel at wavelength w_0 serves as the control channel for the exchange of the control traffic, while the other W channels (referred as data channels) are for actual data transmission. Each station is equipped with two laser transmitters: one fixed laser permanently tuned at w_0 and the other laser tunable to any of the wavelengths w_1, \dots, w_W . The output of the two lasers is combined into a 2×1 coupler, the output of which is then connected to one of the inputs of the $N \times N$ star coupler. Signals transmitted at all of the $(W + 1)$ wavelengths are combined at the star coupler and distributed to all of the stations. Each station also has two receivers: one fixed filter permanently tuned at w_0 and the other tunable to any of the wavelengths w_1, \dots, w_W . At the receiver, the input optical signal is split into two portions by using a 1×2 splitter. One portion is fed to the fixed optical

filter which passes only the control wavelength w_0 , while the other portion goes to the tunable filter which is tuned to pass the desired data wavelength.

We describe the details of the protocol in Section 3.2. Section 3.3 presents the analysis of mathematical models of both the infinite and the finite population cases. In Section 3.4, numerical results from both analysis and simulation are given. Section 3.5 provides the conclusion of the chapter.

3.2 Description of Protocol

3.2.1 The Protocol

We assume the existence of a common clock, which may be obtained by distributing a clock to all the stations. The problem of generating the global clock has been addressed in [OS91]. Time is divided into fixed-sized slots. Packets have a fixed length, which is equal to one slot. The propagation delay from any station to the star coupler and then to any other station is assumed to be equal to R slots. Slots on the data channels are called *data slots* and contain the actual data packets. Slots on the control channel are called *control slots* because they carry only control information about the packets and the transmitters. Each control slot consists of a reservation subpart and a tuning subpart. The reservation subpart is divided into V minislots to be used on a contention basis with a slotted ALOHA protocol, and the tuning subpart is divided into W minislots to convey the wavelength tuning information. The structure of a control slot is shown in Figure 3.2 (The field *queue length* in the figure will be defined later).

A station generating a packet will randomly select one of the V reservation

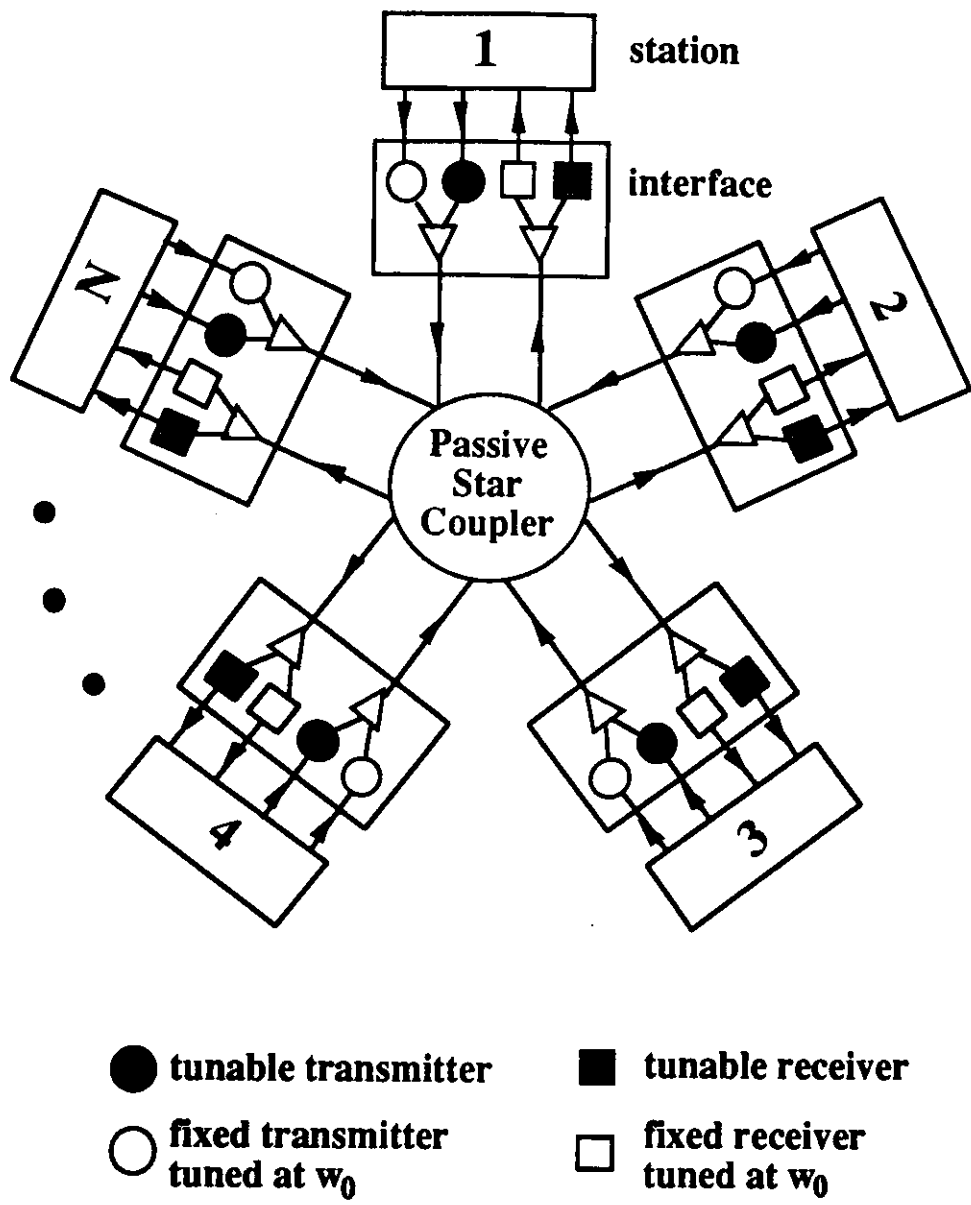


Figure 3.1: N stations connected by a passive star coupler

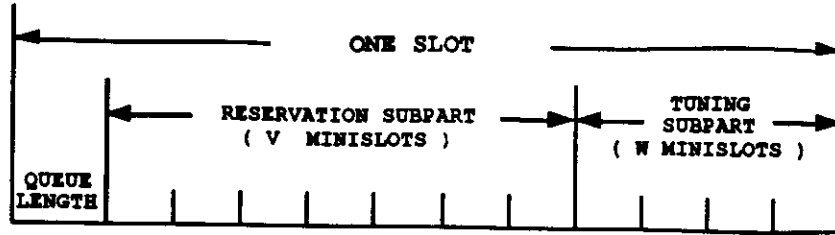


Figure 3.2: Structure of a control slot

minislots in the next control slot and transmit a *reservation* minipacket on the control channel. R slots later the station will hear the result of its reservation. If it is successful, it will be received by all the stations because of the broadcast nature of the control channel. All successful reservations join a common distributed queue of stations waiting to transmit. If there is a collision, the station will randomly select and transmit another reservation minipacket in the next slot with probability p , and with probability $(1 - p)$ it will defer the decision by one slot and randomly select and transmit the reservation in this next slot with probability p , etc.

In the tuning subpart of each control slot, each of the first W stations in the distributed queue will transmit a *tuning* minipacket in an assigned tuning minislot; the minipacket contains the destination address and other relevant information for the data packet to be transmitted in the next slot. In particular, the first station in the queue transmits its minipacket in the first tuning minislot, the second station in the queue transmits in the second minislot, ..., and the W th station transmits in the W th minislot. The position of the tuning minislot uniquely determines the wavelength to use. At the beginning of the next slot, those W stations tune their tunable transmitters to their assigned wavelengths,

with the i th station in the queue using wavelength w_i to transmit its data packet (when there are fewer than W stations in the distributed queue, some data wavelengths will be unused). When the destination sees its address announced in a tuning minislot on the control channel, it tunes its tunable receiver to the corresponding wavelength and receives the data packet at the beginning of the next slot. If two or more packets are addressed to the same destination in a slot, we arbitrarily select the one transmitted on the lowest wavelength number to win the competition (the arbitration can also be made by use of relevant information carried in the tuning minipacket, such as the packet age or priority). The losing stations must start over with the reservation procedure again.

Thus, a station desiring to send a packet must first compete on the ALOHA reservation subchannel to gain access to a minislot on the tuning subchannel. The station then informs its intended receiver to listen (i.e., tune) to a particular wavelength on which the data packet will be transmitted. If a given receiver is informed by more than one station, only one station will be selected, and the others must repeat the entire procedure.

Broadcast traffic can be supported in the following way: when a station has a packet to broadcast to all the other stations, it first competes on the reservation channel. Once it succeeds, then, in the minislot it has reserved on the tuning subchannel, it puts in a special symbol instead of an ordinary destination address. When the other stations see this special symbol on the control channel, they all tune to the corresponding wavelength to receive the broadcast packet arriving in the next slot. That is, broadcast traffic is given higher priority and thus all the other data packets transmitted in the same

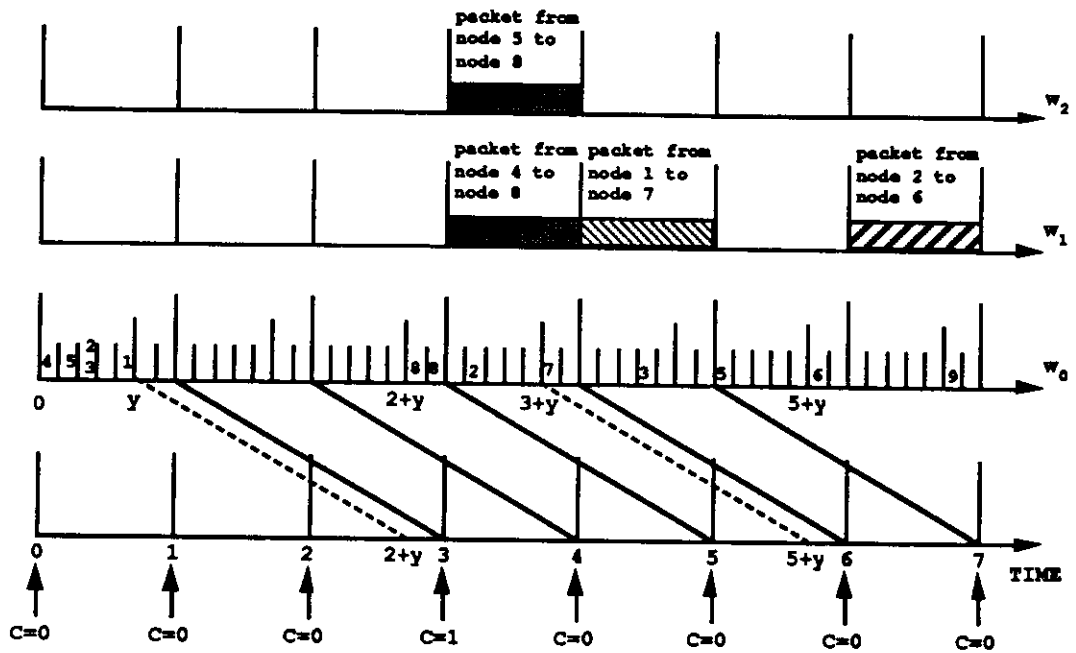
slot are lost. The source stations for these other data packets must repeat the entire protocol as if they had lost in a destination conflict. Multicast traffic can also be supported in a similar way by replacing the special symbol with the multicast group name.

Note that there is a minimum of one round trip delay before a data packet can be transmitted even at light load. A modification of the protocol can be made as follows to improve this: If the current length of the distributed queue, C , is smaller than W and a station decides to transmit a reservation, then the station, instead of transmitting a reservation minipacket, will randomly pick one among those $(W - C)$ free tuning minislots to send out a tuning minipacket, and transmit the data packet on that chosen wavelength at the beginning of next slot. If more than one station chooses the same tuning minislot, a collision occurs and all the stations involved must repeat this procedure provided C is still smaller than W . Otherwise they must send out reservations first and follow the basic protocol. Thus the data packets need not suffer from the round trip delay due to making reservations and the performance at light load is much improved. This modification, however, is not included in the analytic treatment in the following sections.

In order for new stations to join the network, the distributed queue length must be broadcast in each slot on the control channel as shown in Figure 3.2. A station just joining the network first monitors and records activity on the control channel for a period of time equal to the propagation delay (R slots). After R slots it will receive the queue length of the time when it joins the network. With this value and the R -slot record the new station can construct the current value of queue length and proceed with the protocol.

3.2.2 An Illustrative Example

Here we give a simple example where we assume $N=10$, $V=5$, $W=2$, and $R=2$. Let slot t denote the slot from time t to time $t + 1$. Define (s,d) as a packet, where s denotes the source node and d the destination. At time 0, five packets, $(1,7)$, $(2,6)$, $(3,9)$, $(4,8)$, $(5,8)$, are generated. In slot 0, five reservation packets are transmitted with nodes 2 and 3 transmitting in the same minislot. At time $2+y$ (where y is the length of the reservation subpart), nodes 2 and 3 find out that they are involved in a collision, while nodes 4, 5, and 1 realize the success of their reservations and join the distributed queue (in the order 4, 5, 1). In slot 3, node 2 tosses a coin and decides to transmit a reservation minipacket again, while node 3 tosses a coin and decides to defer the decision by one slot. Meanwhile, in the tuning subpart of slot 2, nodes 4 and 5 write their destination addresses (both are 8) into tuning minislots 1 and 2, and, in slot 3, transmit their actual data packets on wavelengths w_1 and w_2 , respectively. In slot 4, node 8 finds out from the control channel that two packets are coming for it and tunes its tunable receiver to w_1 (the lower wavelength) to receive the data packet from node 4. At the same time node 5 realizes that it lost the competition (because of the broadcast nature of the control channel). It tosses a coin, then decides to transmit a reservation minipacket in slot 5 and restarts its reservation procedure again.



C : the length of the distributed queue

Figure 3.3: A scenario of packet transmissions. ($N=10, V=5, W=2, R=2$.)

3.3 Performance Analysis

3.3.1 Model Assumptions

We assume that there are $(W+1)$ wavelenghts available and that the number of stations in the network is N . Each station has a single buffer which is equal to the size of a packet. A new packet arrives at a station with an empty buffer (which we refer to as a “thinking” station) with probability σ at the end of a slot. A packet generated by a station is addressed to any of the other $(N-1)$ stations with equal probability. A source station with a full buffer (i.e., one packet), which we refer to as a “queued station”, will not discard its packet until its successful reception at its destination is recognized; during this time, no new packets may arrive to this station.

3.3.2 The Infinite Population Case

In this subsection we consider the infinite population case where the number of stations, N , is infinitely large. We assume that the total reservation traffic offered to the V ALOHA channels forms a Poisson process with rate G requests per slot. Thus the traffic offered to one ALOHA channel is G/V requests per slot. Specifically, $G = \lim_{\substack{\sigma \rightarrow 0 \\ N \rightarrow \infty}} N\sigma$. The probability that exactly one reservation is transmitted in a single ALOHA channel is $\frac{G}{V}e^{-\frac{G}{V}}$ [Kle76]. Define A_t as the number of successful requests in slot t . Let $A \triangleq \lim_{t \rightarrow \infty} A_t$. The probability mass function (pmf) and z -transform of A are

$$a_j \triangleq \text{Prob}[A = j] = \binom{V}{j} \left(\frac{G}{V}e^{-\frac{G}{V}}\right)^j \left(1 - \frac{G}{V}e^{-\frac{G}{V}}\right)^{V-j} \quad j = 0, \dots, V$$

$$A(z) \triangleq \sum_{j=0}^{\infty} a_j z^j = \left[1 - \frac{G}{V} e^{-\frac{G}{V}} + \frac{Gz}{V} e^{-\frac{Gz}{V}} \right]^V$$

The throughput of the V ALOHA reservation channels is

$$S_{V\text{-ALOHA}} = E[A] = \left. \frac{dA(z)}{dz} \right|_{z=1} = Ge^{-\frac{G}{V}} \quad (3.1)$$

with the maximum throughput, V/e , occurring at $G = V$. Therefore, the capacity of the system will be the minimum of the reservation channel throughput and the data channel throughput, namely, $\min(V/e, W)$.

The packet delay D , defined as the time interval from the packet arrival instant to the successful reception at the destination, can be computed as follows:

$$D = D_r + D_q + (R + 1)$$

where D_r is the reservation delay defined as the interval between the packet arrival instant and the moment the success of the reservation is recognized. D_q is the queueing delay, which is the time period from the instant the success of reservation is recognized until the beginning of the successful transmission of the data packet. $(R + 1)$ slots account for the transmission (1 slot) and propagation delay (R slots). Therefore, the average packet delay is

$$E[D] = E[D_r] + E[D_q] + (R + 1) \quad (3.2)$$

We compute $E[D_r]$ first. Define

$$q_r \triangleq \text{Prob}[\text{a transmitted reservation request is successful}] = e^{-\frac{G}{V}}$$

Therefore, we have

$$\begin{aligned} h_n &\triangleq \text{Prob}[\text{reservation succeeds at the } n\text{th trial}] \\ &= q_r(1 - q_r)^{n-1} \end{aligned}$$

and so

$$\begin{aligned}
 h &\triangleq \text{the average number of reservation requests sent per packet} \\
 &= 1/q_r = e^{\frac{G}{V}}
 \end{aligned}$$

The average time between two consecutive transmissions of the request is $(\frac{1}{p} + R)$; therefore,

$$E[D_r] = (R + 1) + (h - 1)\left(\frac{1}{p} + R\right) \quad (3.3)$$

To compute $E[D_q]$, define C_t as the length of the distributed queue at time t , and $C = \lim_{t \rightarrow \infty} C_t$. Let X be the position of a typical (say tagged) successful reservation among those successful ones in the same slot. Since the probability that any two packets are going to the same destination is zero when the number of stations is infinite, $E[D_q]$ can be computed as follows:

$$E[D_q] = \left[\frac{E[C] + E[X]}{W} \right] \quad (3.4)$$

To get $E[C]$, we first have

$$C_{t+1} = \max(0, C_t + A_{t+1} - W)$$

Assume that steady state exists. Solving this using the technique in [Kle75], we get

$$C(z) = \sum_{j=0}^{\infty} c_j z^j = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{W-i-1} c_j a_j (z^{i+j} - z^W)}{(1 - \frac{G}{V} e^{-\frac{G}{V}} + \frac{Gz}{V} e^{-\frac{G}{V}})^V - z^W} \quad (3.5)$$

where $c_j \triangleq \text{Prob}(C = j)$ is the pmf of C . We denote the denominator of $C(z)$ by $D(z)$. Using Rouché's theorem [Kle75], it can be shown (see appendix A.1) that

W roots of $D(z)$ are on or inside the unit circle $|z| = 1$ provided that $V/e < W$. Those roots must cancel out with the roots of the numerator. Therefore, $C(z)$ becomes

$$C(z) = \frac{B}{(z - z_1)(z - z_2) \cdots (z - z_{(V-W)})}$$

where $z_1, z_2, \dots, z_{(V-W)}$ are the $(V - W)$ roots of $D(z)$ outside the unit circle $|z| = 1$ and B is a constant. The condition $C(1) = 1$ gives us $B = (1 - z_1) \cdots (1 - z_{(V-W)})$. Therefore,

$$C(z) = \frac{(1 - z_1)(1 - z_2) \cdots (1 - z_{(V-W)})}{(z - z_1)(z - z_2) \cdots (z - z_{(V-W)})}$$

and

$$E[C] = \left. \frac{dC(z)}{dz} \right|_{z=1} = - \sum_{i=1}^{V-W} \frac{1}{1 - z_i} \quad (3.6)$$

Next we compute $E[X]$. Let $x_j \triangleq \text{Prob}(X = j)$ be the pmf of X . Define the random variable K to be the total number of successful reservations in the same slot where the tagged successful reservation resides. If the number of successful reservations in a slot is large, then it is more likely that our tagged reservation will reside in this slot. Therefore, the pmf $\text{Prob}(K = k)$ should be proportional to the number of successful reservations in the slot (which is k) as well as to the relative occurrence of such slots (which is a_k) (See chapter 5 in [Kle75]). Since $\sum_{k=1}^V \text{Prob}(K = k)$ must equal to one, we have

$$\text{Prob}(K = k) = \frac{ka_k}{E[A]} \quad k = 1, \dots, V.$$

Since the tagged reservation can be at any position in a group with equal probability, we have

$$\text{Prob}(X = j | K = k) = \frac{1}{k} \quad j = 1, \dots, k.$$

Unconditioning on k , we get

$$\begin{aligned} x_j &= \sum_{k=j}^V \frac{1}{k} \frac{ka_k}{E[A]} \\ &= \frac{1}{E[A]} \sum_{k=j}^V a_k \quad j = 1, \dots, V \end{aligned}$$

and

$$\begin{aligned} E[X] &= \sum_{j=1}^V jx_j = \frac{1}{2} + \frac{E[A^2]}{2E[A]} \\ &= 1 + \frac{1}{2}(V-1)\frac{G}{V}e^{-\frac{G}{V}} \end{aligned}$$

where we have evaluated $E[A^2] = Ge^{-\frac{G}{V}} \left[1 + (V-1)\frac{G}{V}e^{-\frac{G}{V}} \right]$ from the expression for $A(z)$ given earlier. $E[D_q]$ can now be computed according to Eq. (3.4). The average packet delay is finally obtained from Eq. (3.2).

In Figure 3.4, we plot the ALOHA channel capacity and the data channel capacity. For a fixed value of the reservation traffic, we see that the throughput is the minimum of the ALOHA channel capacity and the number of wavelengths. Figure 3.5 shows the throughput-delay curves. These curves give the performance for an infinite population of stations whose collective generation rate of new packets is S . From Figure 3.5 we see that the system is not stable (there are two different values of delay associated with a given throughput) and some dynamic control procedure (e.g., see [Lam74]) will be required to stabilize the system. Note how close the achievable throughput is to the maximum of 4 packets/slot.

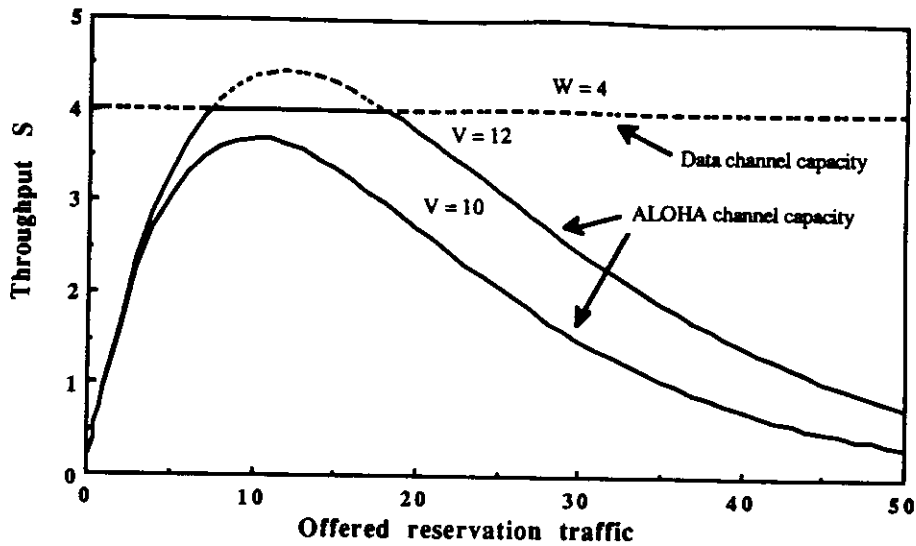


Figure 3.4: Throughput versus offered reservation traffic. $N = \infty, W=4$.

3.3.3 The Finite Population Case

In this subsection we analyze the case when the number of stations in the system is finite. An approximate model of the system is shown in Figure 3.6. In this model, each station can be in one of the following $(3R + 3)$ modes (here R is assumed to be an integer) at any instant: TH, RT, Q, PQ_m, PR_m , and PS_m ($1 \leq m \leq R$). Stations can move from one mode to another mode only at the beginning of each slot.

Stations in each mode move as follows: Stations in the TH (thinking) mode generate a packet with probability σ at the end of a slot. Stations in the Q (queued) mode are currently in the distributed queue. A station that had suffered a collision of its reservation packet and has realized it is in the RT (retransmission) mode and will retransmit the reservation with probability p

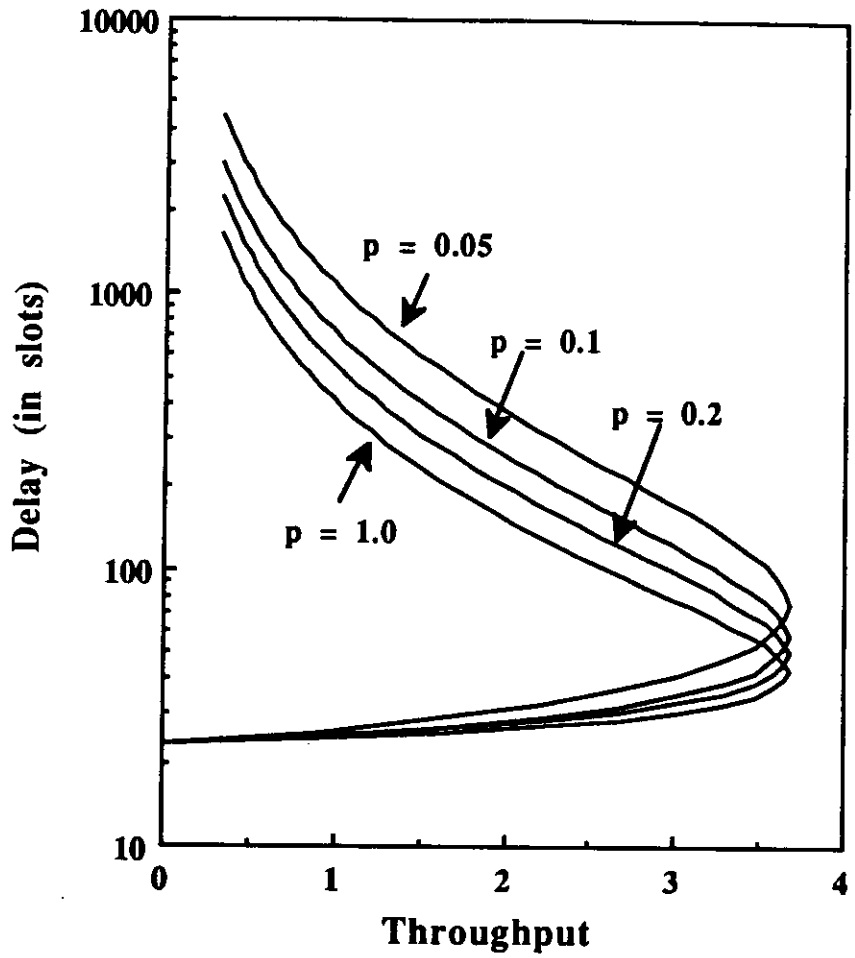


Figure 3.5: Throughput versus delay curves. $N = \infty$, $V=10$, $W=4$, $R=10$.

in the next slot. The PQ_m , PR_m , and PS_m modes, $m = 1, \dots, R$, are unit delay elements and represent the influence of the channel propagation delay. Stations in the PQ_m mode will move into the PQ_{m-1} mode at the next slot with probability 1. Thus, as can be seen in Figure 3.6, stations in the PQ_m mode will enter the Q mode after m slots. The same also applies to the PR_m and PS_m modes.

We now define a state vector for the system. Let n_{RT} be a random variable denoting the number of stations in the RT mode, n_Q that in the Q mode, i_m that in the PR_m mode, j_m that in the PQ_m mode, and k_m that in the PS_m mode, $m = 1, \dots, R$. In the model we will further make a *nonpersistence assumption*: a station, upon entering the RT mode, will randomly reselect a destination for its packet (It is not the case in the real system, but later we will see that this model still predicts the performance very well under this assumption). Define the vector $\mathbf{n} \triangleq (n_{RT}, n_Q, i_1, \dots, i_R, j_1, \dots, j_R, k_1, \dots, k_R)$ as the state vector of the system. Then we can see that the vector \mathbf{n} forms a discrete-time Markov chain with a finite state space.

Unfortunately, since the state space is so large, it is difficult for us to solve this Markov chain. Therefore, we utilize the technique of equilibrium point analysis (EPA) [TI84] to analyze this chain.

3.3.3.1 The Modified Model

To simplify the analysis, we first consider a modification of the model in Figure 3.6 as suggested in [TI84], which combines the two inputs (from the TH mode and the RT mode) of the slotted ALOHA reservation channel. Since we have assumed bursty stations, we shall confine ourselves to the case $\sigma \leq p$. The

modified model is shown in Figure 3.7, where the TH mode in Figure 3.6 has been decomposed into two modes, I and T , and the RT mode in Figure 3.6 has become part of the T mode. A station that has just come out from the PS_1 mode moves into the I and the T modes with probabilities $(1 - \sigma/p)$ and σ/p , respectively. A station in the I mode will move into the T mode at the next slot with probability σ , and a station in the T mode transmits a reservation minipacket (i.e., moves out of the T mode) with probability p . The modified model is equivalent to the original model from the viewpoint of its stochastic behavior, which is interpreted as follows: Let X_1 and Y_1 be random variables representing the time (number of slots) from the moment a station which was originally at the PS_1 and PR_1 modes in Figure 3.6, respectively, moves to the left of the dashed boundary until the instant that station moves to the right of the dashed boundary for the first time (Note that X_1 and Y_1 are simply the time a station spends in the TH and RT modes, respectively). Let X_2 and Y_2 be the corresponding random variables in the modified model (Figure 3.7). It can be shown [TI84] that X_2 and Y_2 have the same pmf's as X_1 and Y_1 , respectively. This means that the stochastic behavior of the parts to the right of the dashed boundary of the two models are exactly the same. Therefore, we can derive any performance characteristics of the model in Figure 3.6 by using the model in Figure 3.7, as long as they are derived from statistics of the part of the model to the right of the dashed boundary.

In the modified model in Figure 3.7, we now let n_T be a random variable representing the number of stations in the T mode. Then, it is apparent that the modified state vector $\mathbf{n} = (n_T, n_Q, i_1, \dots, i_R, j_1, \dots, j_R, k_1, \dots, k_R)$ is also a Markov chain under the nonpersistence assumption.

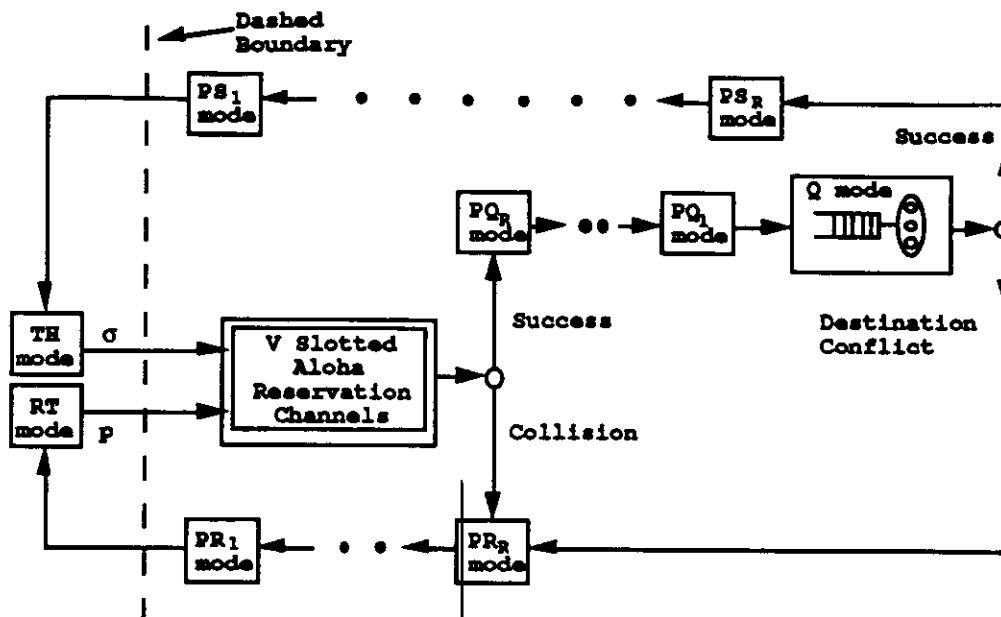


Figure 3.6: An approximate model of the system.

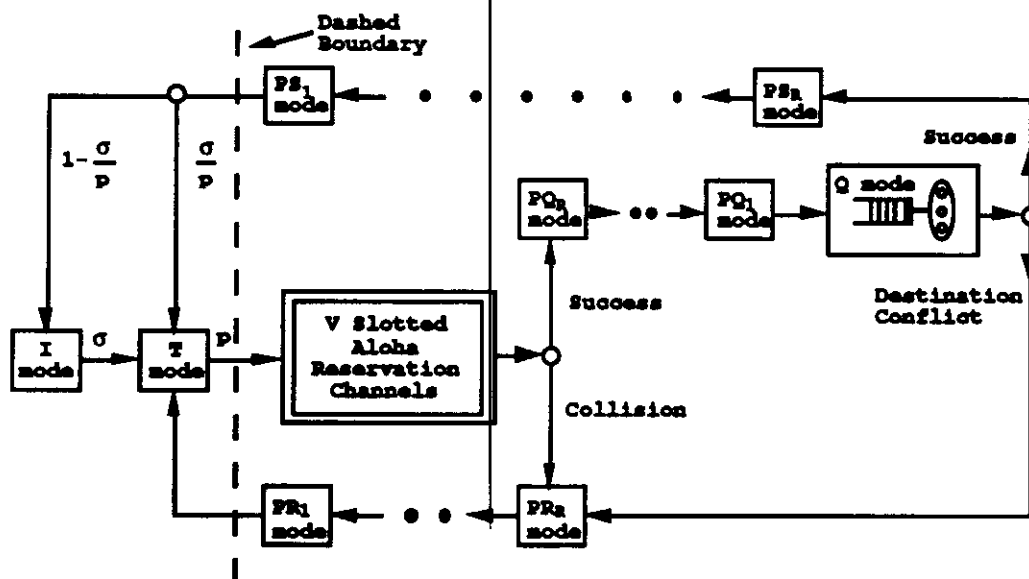


Figure 3.7: A modified model of Figure 3.6 in the case of $\sigma \leq p$.

3.3.3.2 The Equilibrium Point Equation

An equilibrium point is a point in state space such that at that point the expected drift is zero [KL75]. In the EPA method, we assume that the system is always at an equilibrium point. Applying the above condition to all the modes in our model, we get a set of equations called equilibrium point equations, the solution to which gives us one or more equilibrium points.

Let $\bar{\mathbf{n}} = (\bar{n}_T, \bar{n}_Q, \bar{i}_1, \dots, \bar{i}_R, \bar{j}_1, \dots, \bar{j}_R, \bar{k}_1, \dots, \bar{k}_R)$ be an equilibrium point. Let $\delta_I(\bar{\mathbf{n}})$ be the conditional drift (i.e., the conditional expectation of the increase in the number of stations) in mode I in a slot, given that the system is at $\bar{\mathbf{n}}$. Setting $\delta_I(\bar{\mathbf{n}}) = 0$, we then have

$$\delta_I(\bar{\mathbf{n}}) = \bar{k}_1 \left(1 - \frac{\sigma}{p}\right) - \left[N - \bar{n}_T - \bar{n}_Q - \sum_{m=1}^R (\bar{i}_m + \bar{j}_m + \bar{k}_m) \right] \sigma = 0 \quad (3.7)$$

Next, let $X(\bar{\mathbf{n}})$ denote the conditional expectation of the number of stations that move out of mode Q in a slot, given that the system is in state $\bar{\mathbf{n}}$. Evaluating mode Q , we get

$$\delta_Q(\bar{\mathbf{n}}) = X(\bar{\mathbf{n}}) - \bar{j}_1 = 0 \quad (3.8)$$

Next we define two other terms. Let $f_T(\bar{\mathbf{n}})$ denote the conditional expectation of the number of stations that successfully transmit reservation minipackets (thus move from mode T to mode PQ_R) in a slot, given that the system is in state $\bar{\mathbf{n}}$. Let $g_Q(i)$ denote the average number of stations that successfully transmit a data packet (therefore moving from mode Q to mode PS_R) in a slot, given that i stations transmit (i.e. move out of mode Q). It can be derived (see appendix A.2) that

$$f_T(\bar{\mathbf{n}}) = \bar{n}_T p \left(1 - \frac{p}{V}\right)^{\bar{n}_T - 1} \quad (3.9)$$

and

$$g_Q(i) = N \left[1 - \left(1 - \frac{1}{N-1}\right)^{i-1} \left(\frac{N^2 - 2N + i}{N(N-1)}\right) \right] \quad (3.10)$$

Evaluating the conditional expectation of increase for the PR_m , PQ_m , and PS_m ($1 \leq m \leq R$) modes, we obtain the corresponding equations as follows:

$$\bar{i}_1 = \bar{i}_2 = \dots = \bar{i}_R = \bar{n}_T p - f_T(\bar{\mathbf{n}}) + X(\bar{\mathbf{n}}) - g_Q(X(\bar{\mathbf{n}})) \quad (3.11)$$

$$f_T(\bar{\mathbf{n}}) = \bar{j}_R = \dots = \bar{j}_1 \quad (3.12)$$

$$\bar{k}_1 = \dots = \bar{k}_R = g_Q(X(\bar{\mathbf{n}})) \quad (3.13)$$

We did not write down the equation for the T mode since it is linearly dependent on the others. After some manipulation of the equations above, we get the following equations:

$$f_T(\bar{\mathbf{n}}) - X(\bar{\mathbf{n}}) = 0 \quad (3.14)$$

$$g_Q(X(\bar{\mathbf{n}})) \left(1 - \frac{\sigma}{p}\right) - [N - \bar{n}_T - \bar{n}_Q - R(\bar{n}_T p + f_T(\bar{\mathbf{n}}))] \sigma = 0 \quad (3.15)$$

We next model the queueing system in mode Q as a W -server system with a binomial input with mean $f_T(\bar{\mathbf{n}})$ and a fixed service time of one slot for each customer. Define

$$\rho = \frac{f_T(\bar{\mathbf{n}})}{W}$$

which is the utilization of the queueing system. The z-transform of the arrival process is

$$U(z) = \left(1 - \frac{\rho W}{V} + \frac{\rho z W}{V}\right)^V$$

This system was solved in the previous section, and the average number of customers in the system (see Eq. (3.6)) should be equal to \bar{n}_Q , the average number of stations in mode Q . Therefore, we have

$$\bar{n}_Q = - \sum_{i=1}^{V-W} \frac{1}{1 - z_i} \quad (3.16)$$

where $z_i, i = 1, \dots, (V - W)$, are the roots of $U(z) - z^W = 0$ outside the unit circle $|z| = 1$. Also, since $X(\bar{n}) = f_T(\bar{n}) = \rho W$, equations (3.14) and (3.15) become

$$\bar{n}_T p \left(1 - \frac{p}{V}\right)^{\bar{n}_T - 1} - \rho W = 0 \quad (3.17)$$

$$\frac{g_Q(\rho W)}{\sigma} \left(1 - \frac{\sigma}{p}\right) = N - (1 + pR)\bar{n}_T - \bar{n}_Q - \rho RW \quad (3.18)$$

It is clear that the equations (3.16) (3.17) and (3.18) can be solved for ρ . The system is said to be *stable* if only one solution exists. Otherwise, if there is more than one solution, the system is said to be *unstable* [KL75].

3.3.3.3 Throughput and Delay

We now define the throughput $S(\mathbf{n})$ to be the conditional expectation of the number of successfully received data packets in a slot, given that the system is in state \mathbf{n} . Then, it is clear that the throughput at an equilibrium point is expressed as

$$S(\bar{n}) = g_q(X(\bar{n})) = g_q(\rho W) \quad (3.19)$$

The average packet delay, which is the average time, in number of slots, from the moment the packet is generated until the instant the packet is correctly received by the destination, can be approximated from Little's result [Kle75] to give

$$E[D] \cong \frac{N}{S(\bar{n})} - \frac{1}{\sigma} \quad (3.20)$$

3.4 Numerical Results

In this section we use both the analytical model and simulation to investigate the performance of the proposed protocol under various system parameters. In all figures, we note the excellent agreement between the analysis and simulation.

Figure 3.8 shows the throughput versus delay curve for $N = 80$, $W = 3$, $V = 8$, $R = 1$, $p = 0.6$, and σ increasing from 0 to 0.5. Figure 3.9 shows the same performance curve for $N = 120$, $W = 4$, $V = 10$, $R = 2$, $p = 0.6$, and again σ increasing from 0 to 0.5. The 95% confidence interval for the simulation points is also plotted in the figures. Note that in the lower part of the curves the delay increases very slowly with the throughput; thus a high system throughput and low delay can be achieved.

Next we compare both the infinite and finite population analyses for a system consisting of a large number of stations. Figure 3.10 shows the performance curve for $N = 500$, $W = 4$, $V = 10$, $R = 10$, $p = 0.2$, and σ increasing from 0 to 0.2. It is interesting to note the close match among both analyses and the

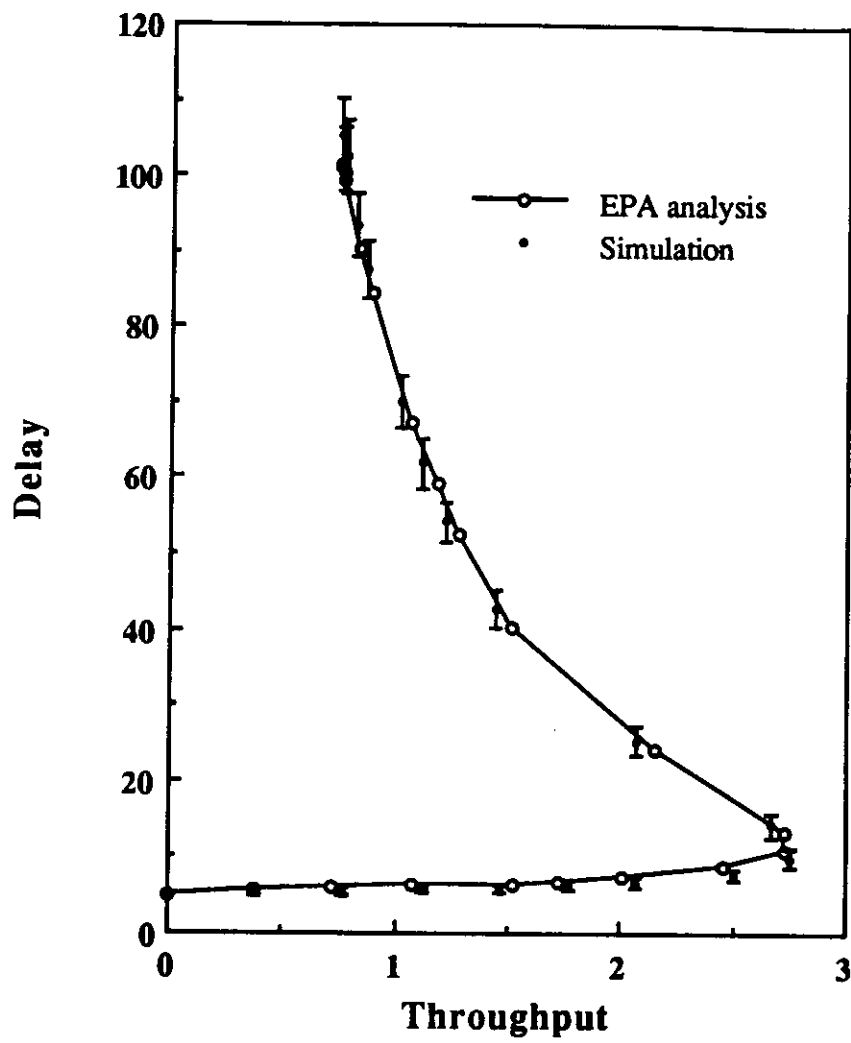


Figure 3.8: Throughput versus delay curve. $N=80$, $W=3$, $V=8$, $R=1$, $p=0.6$.

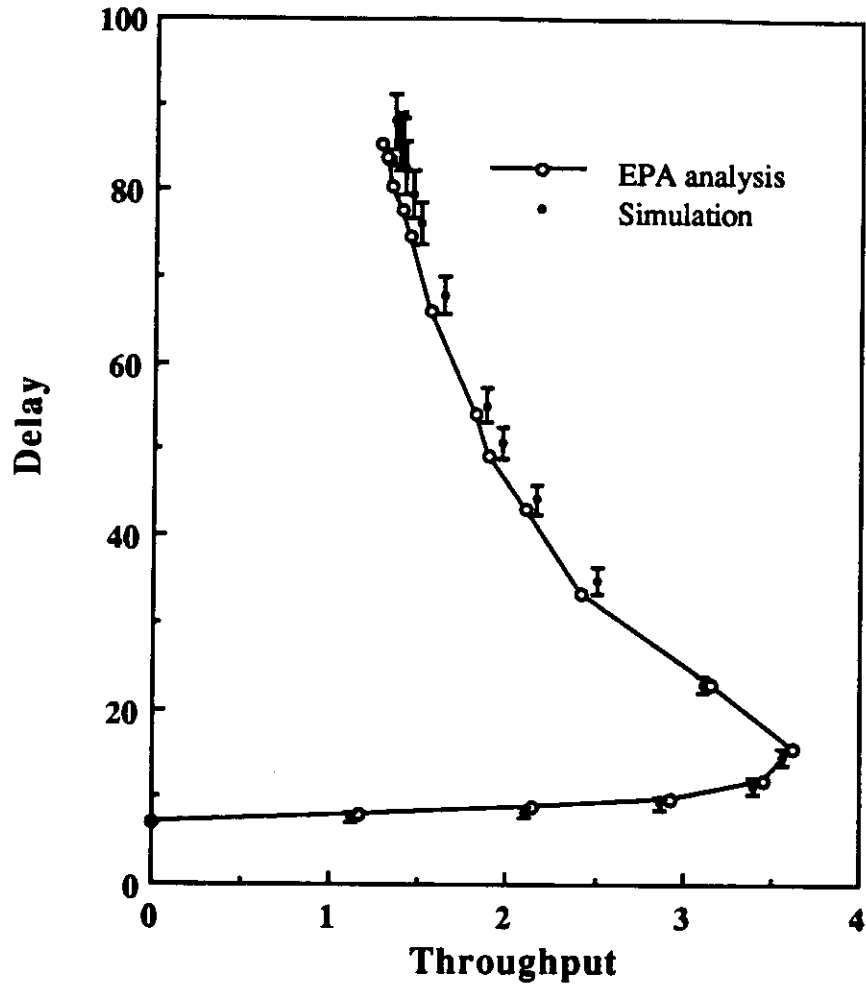


Figure 3.9: Throughput versus delay curve. $N=120$, $W=4$, $V=10$, $R=2$, $p=0.6$.

simulation results.

Figures 3.11 and 3.12 show the effect of varying V and W while keeping their sum constant by fixing the slot size. We see that for different cases, the maximum throughput always occurs at the point where $V \approx eW$. This is not surprising since the capacity of a slotted ALOHA channel is $1/e$. By making $V \approx eW$ the capacities of the V ALOHA channels and the W data channels are balanced. When $V \ll eW$, there is not enough throughput coming out of the ALOHA reservation channels to keep the W data channels busy. When $V \gg eW$, most stations are waiting (in the Q mode) for a wavelength on which to transmit a data packet. When $V \approx eW$, the throughput is roughly equal to the load offered by the stations.

Let D_T and D_Q denote the average time a station spends in modes T and Q , respectively, in a cycle. By Little's result, the throughput is equal to $N/(\frac{1}{\sigma} + D_T + R + D_Q + R)$. Setting both D_T and D_Q to 1, which is the minimum possible value, we obtain an upper bound for the throughput,

$$S_{UB} = \frac{N}{\frac{1}{\sigma} + 2(R + 1)}$$

and a lower bound for the delay,

$$D_{MIN} = \frac{N}{S_{UB}} - \frac{1}{\sigma} = 2(R + 1)$$

which are shown as dashed horizontal lines in Figures 3.11 and 3.12, respectively. The "flat" regions in these figures show how those bounds take effect. Note that when $S_{UB} > \min(V/e, W)$ (the case of $V + W = 14$), the flat region disappears since the throughput is not limited by the station-generated load.

The effect of varying the propagation delay, R , is shown in Figures 3.13 and 3.14. Although the physical distance of the network is usually fixed, R

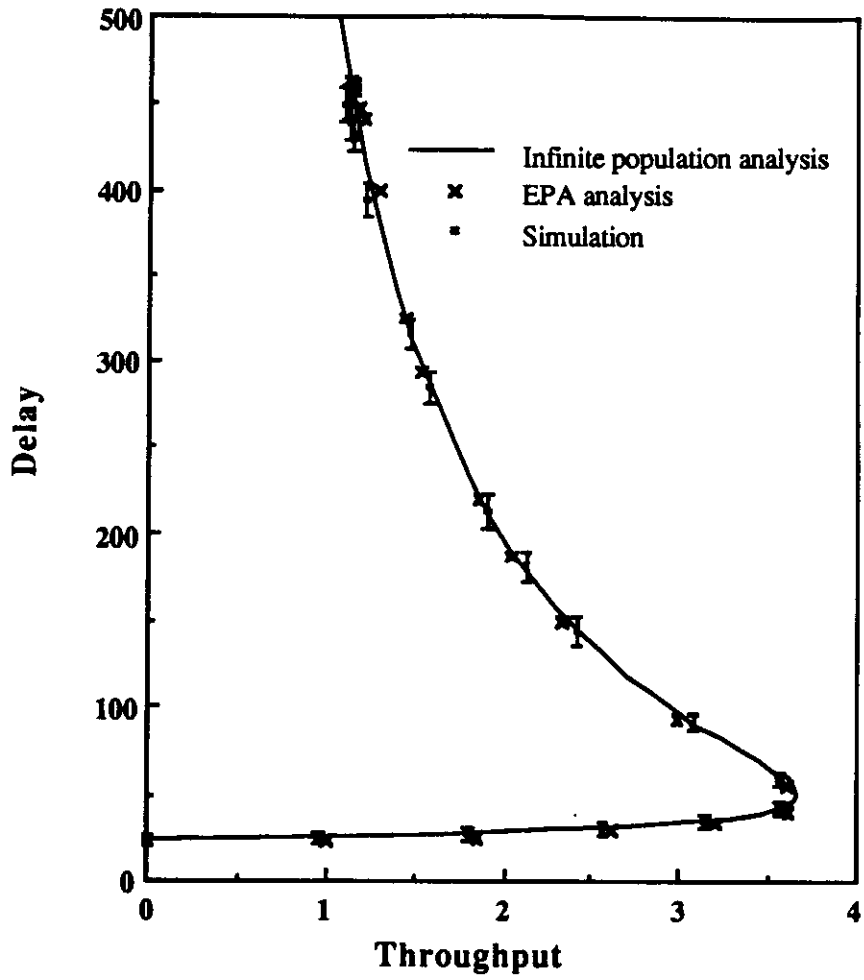


Figure 3.10: Throughput versus delay curves. $N=500$, $W=4$, $V=10$, $R=10$, $p=0.2$.

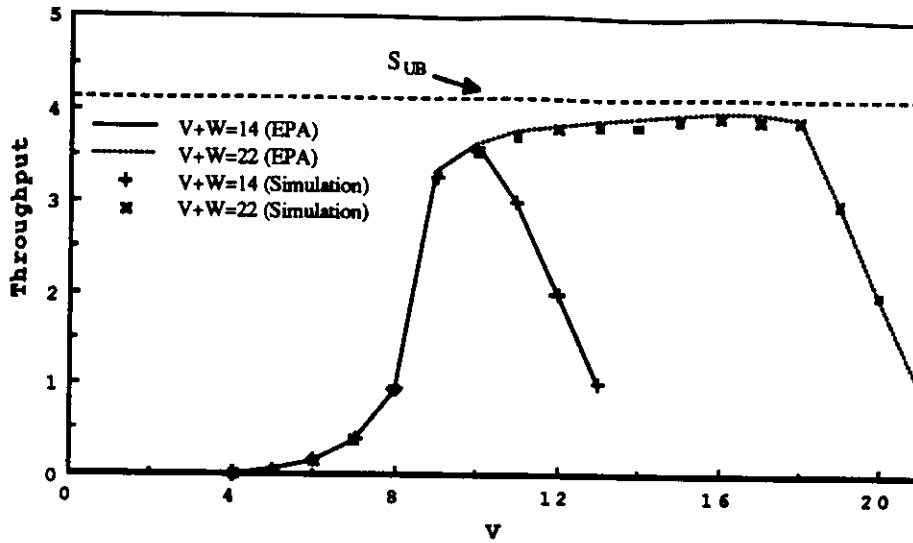


Figure 3.11: Throughput versus V for fixed slot sizes ($V + W$ constant). $N = 500$, $R=10$, $\sigma=0.01$, $p=0.2$.

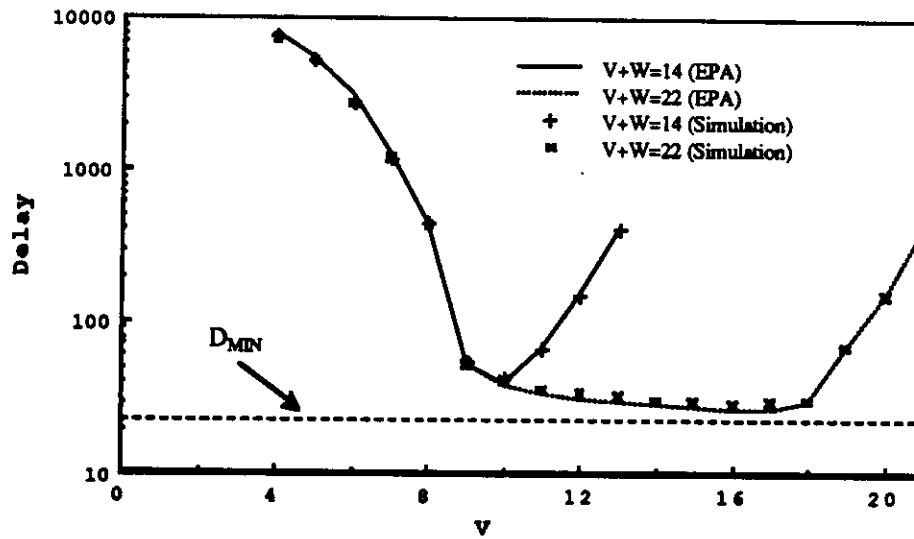


Figure 3.12: Delay versus V for fixed slot sizes ($V + W$ constant). $N = 500$, $R=10$, $\sigma=0.01$, $p=0.2$.

can be varied by varying the packet (and thus the slot) size. We see that when R is small (i.e., large slot size), too much traffic is offered to the V ALOHA reservation channels and the throughput is small. As we increase R , the throughput increases too. It reaches the maximum at $R = 7$, then falls off because there is not enough traffic in each slot when R becomes large. The tail of the throughput curve is bounded by the upper bound S_{UB} . The reason the increase is so sharp around $R = 6$ and 7 is that the system changes from an overloaded system (see Figure 7(d) in [KL75]) to a bistable one and then to a stable one as R increases from 5 to 7. This phenomenon is observed in both the EPA analysis and the simulation. (For the case $R = 6$, the average value of the two solutions obtained from EPA is plotted.)

Figure 3.15 shows the influence of the destination conflicts. Suppose i data packets are transmitted in a slot. The number of packets successfully received by their destinations is $g_q(i)$. We plot the fraction of success, $g_q(i)/i$, versus i assuming $N = 500$. Near-term technology limits W , the number of transmitter/receiver-tunable wavelengths available, to be fewer than about twenty; thus we see that destination conflicts are not currently a serious concern.

3.5 Conclusions

In this chapter a wavelength division multiple access protocol (with W wavelength channels) was proposed to provide a high-capacity optical fiber local area network to a large population of N stations. We assumed $N \geq W$. The stations' traffic was assumed to be bursty as in the case of computer communications. The performance of the protocol was completely analyzed for both the infinite and finite population cases. The numerical results show that an

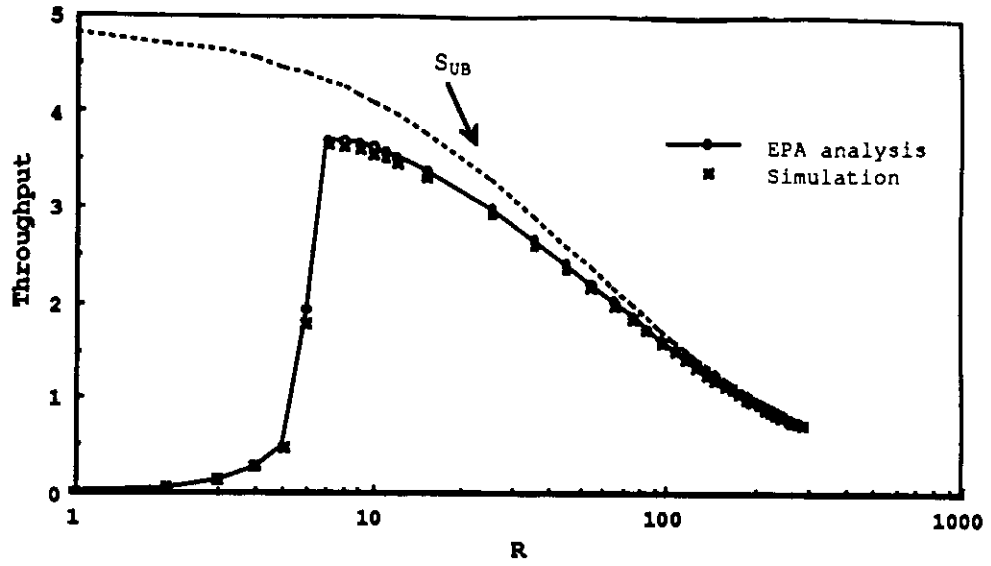


Figure 3.13: Throughput versus propagation delay. $N = 500$, $W=4$, $V=10$, $\sigma=0.01$, $p=0.2$.

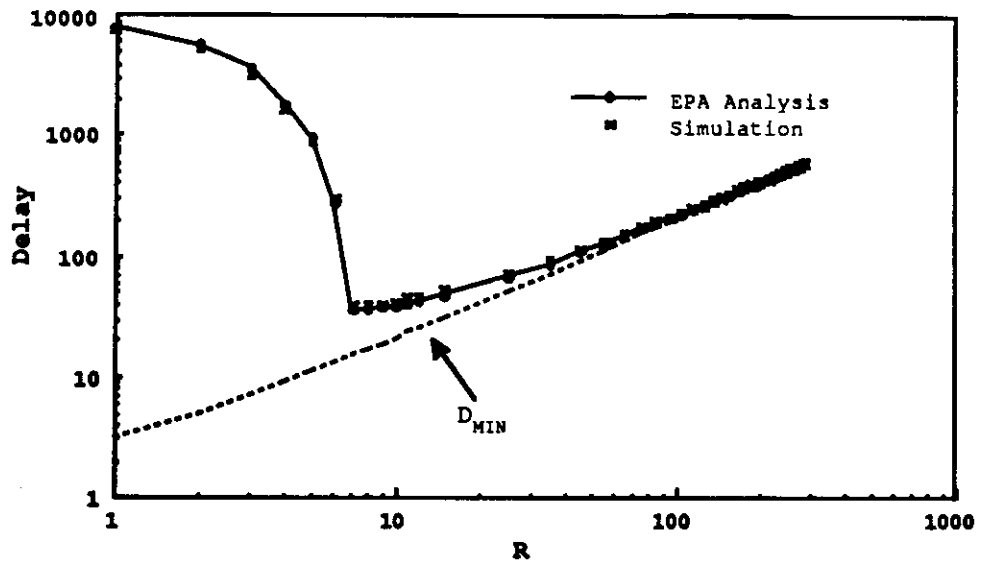


Figure 3.14: Delay versus propagation delay. $N = 500$, $W=4$, $V=10$, $\sigma=0.01$, $p=0.2$.

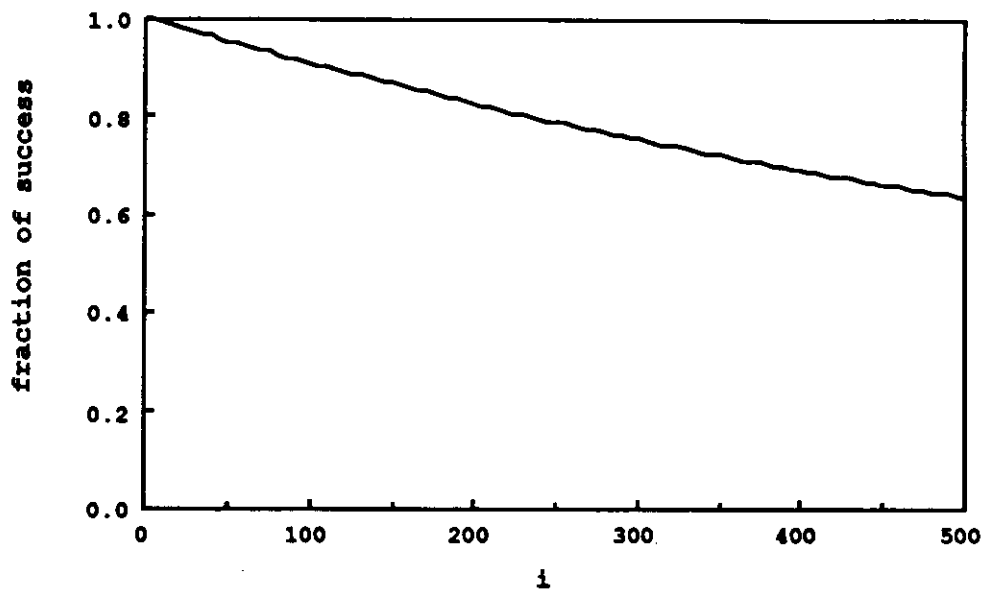


Figure 3.15: The fraction of success versus the number of data packets transmitted. $N = 500$.

aggregate throughput substantially larger than the electronic speed of a single station can be supported. The effects of various system parameters and their optimal selection were also investigated. By comparison with simulation, our analytical approximations were shown to be excellent.

The protocol proposed in this chapter possesses the following advantages: it can support a large number of stations with a few wavelengths (some of the schemes, e.g. [Che90][OS91], require the number of wavelengths to be the same as the number of stations), new stations can join the system (after monitoring the system for a period of one propagation delay) without shutting down the system, and broadcast and multicast traffic can be easily supported. The protocol also has its drawbacks: it requires some stabilizing procedures (as in any system with an ALOHA component), and it cannot support real-time traffic (which will play an important role in future high speed integrated networks). New protocols are still needed to efficiently support the transport of mixed-media traffic.

CHAPTER 4

A WDMA Protocol for MANs with a Dual Bus Topology

4.1 Introduction

In the previous chapter we proposed a WDMA protocol for a star network. Although a star topology is more power-effective than a bus topology, the star network does, as discussed in Section 1.4.2, have a few limitations such as the high cost of the fiber plant deployment for a large number of stations distributed over a large area, and the expensive cost of fabricating large-size star couplers. Therefore, it is often better to run a linear optical fiber bus to connect user stations in a large geographical area such as a metropolitan area, especially with the help of recent progress on the erbium-doped optical amplifiers. In fact, the IEEE has selected the Distributed Queue Dual Bus (DQDB) [NBH88] networks as its 802.6 MAN standard, which assumes a dual bus topology. Thus, networks based on dual bus topology are expected to be very popular in the future, and it is important to develop efficient multichannel access protocols.

In this chapter we propose a WDMA protocol which is a generalization of the basic single-channel DQDB protocol to the multichannel case. In Section 4.2, we first explain the basic operation of the DQDB protocol, then we extend it to the multichannel case. An approximate queueing model is constructed

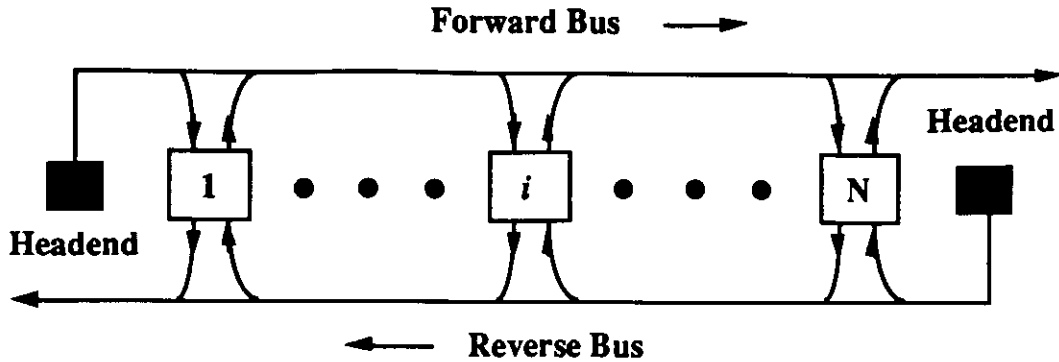


Figure 4.1: A basic dual bus network.

in Section 4.3 to solve for the capacity of the system and mean packet delays for each station. In Section 4.4, numerical results from both simulation and analysis are plotted and compared. Section 4.5 concludes the chapter.

4.2 System Description

A basic dual bus network is shown in Fig. 4.1. There are two unidirectional fiber buses, called the *forward* and the *reverse* bus, running in opposite directions. Stations are connected to both buses. Two headends located at the end of each bus continuously generate streams of fixed-length slots. A station uses the forward bus to transmit traffic for stations (called *downstream* stations) to its right, and the reverse bus for stations (called *upstream* stations) to its left. Several dual bus networks with different access schemes, e.g., Fasnet [LF82] and DQDB [NBH88], have been proposed in the literature. Our WDMA protocol is a generalization of the DQDB protocol to the multiple channel case, so we shall first describe the basic operation of the DQDB protocol in brief in the next subsection.

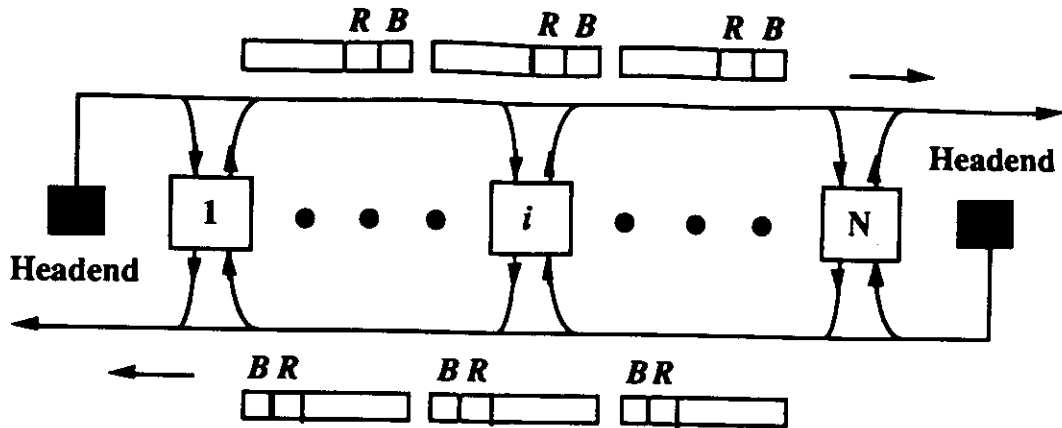


Figure 4.2: An example DQDB network.

4.2.1 DQDB Preliminary

An example DQDB network is drawn in Figure 4.2. In a DQDB network, a 53-byte *segment* is the data unit which consists of a 5-byte header and a 48-byte segment payload. A segment is exactly equal to one slot long. The DQDB network employs a distributed queuing protocol to control the access to the fixed-length slots on the buses. We shall only explain the mechanism to access the forward bus since the access to the other bus is identical, but independent.

DQDB uses two control bits, a busy (B) and a request (R) bit in each slot, to control access to the bus. Each station keeps two counters, a request (RQ) counter and a countdown (CD) counter. When the station has no segments to send, it increases the RQ counter by one for every slot passing by on the reverse bus with the R bit set, and decreases the RQ counter by one for every slot passing by on the forward bus with the B bit unset. In this way the value of the RQ counter at a station approximately equals the number of empty slots that downstream stations need to transmit their data segments. When a station

has a segment to transmit on the forward bus, it will first find a slot on the reverse bus with the R bit unset and set it to one; it then transfers the current value of the RQ counter to its CD counter, and resets the RQ counter to zero. This action loads the CD counter with the number of downstream segments queued ahead of it. This, along with the sending of the R bit on the reverse bus, effectively places the segment in the distributed queue.

The station continues to increase the RQ counter by one for each R bit set on the reverse bus. Now, however, the station will decrease the CD counter (instead of the RQ counter) by one whenever it lets an idle slot pass by on the forward bus. When the CD counter goes to zero, the station waits for the next idle slot and writes its segment into that slot. If the station has more data segments to transmit, it will try to set another R bit on the reverse bus and then start to count down again; otherwise, the station goes back to the idle state.

4.2.2 The WDMA Access Protocol

The system considered here is also a dual bus network where N stations are connected. Each station can transmit and receive on both buses. There are $(W+1)$ wavelengths available, w_0, w_1, \dots, w_W , in the system, where, as in the previous chapter, the channel on wavelength w_0 is dedicated to the exchange of the control information, and the other W channels are used for the transmission of the actual data traffic. For each bus, each station has two lasers: one fixed laser tuned at w_0 and the other tunable laser which can be tuned to any of the wavelengths, w_1, \dots, w_W , in a few nanoseconds. The outputs of those two lasers are merged by a coupler before transmitting into the fiber bus. For each bus,

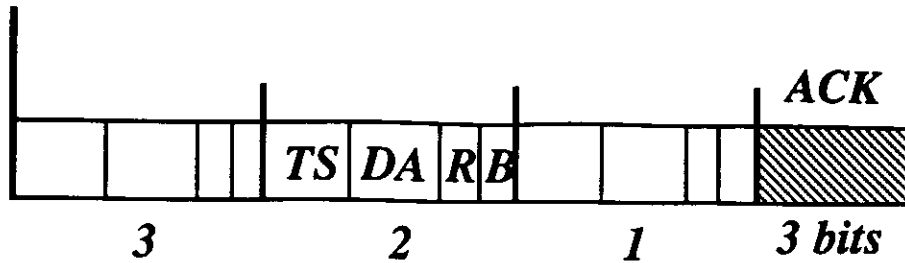


Figure 4.3: The structure of a slot on the control channel for $W = 3$.

each station is also equipped with one filter fixed tuned at w_0 and one tunable filter tunable to any of the wavelengths, w_1, \dots, w_W , in a few nanoseconds. The signal received at the receiving tap is divided into two portions by a splitter before fed to those two filters. The fixed filter is to monitor the activities on the control channel, while the tunable filter will extract the desired data from a wavelength specified on the control channel.

We again assume that the headends continuously generate streams of fixed-length slots. A packet (i.e., a data unit) length is equal to one slot. A slot on the control channel is divided into a W -bit Acknowledgment (ACK) field and W minislots. Each minislot consists of a busy (B) bit, a request (R) bit, a destination address (DA) field, and a timestamp (TS) field. Figure 4.3 shows the structure of a control slot for $W = 3$. Note that the number of minislots in a control slot is exactly equal to the number of data wavelengths, therefore, the position of the minislot uniquely defines a data channel.

Here again we shall only describe the control mechanism for access of the forward bus because of the symmetry. All the control activities occur on the control channel. As in DQDB, each station keeps a record of two counters, the request (RQ) counter and the countdown (CD) counter. When the station

has no data to send, it decreases the RQ counter by one as each idle minislot (instead of a *slot*) passes by on the forward bus, and increases the RQ counter by one as it sees an R bit set in a minislot on the reverse bus. When a new packet arrives at a station, a unique timestamp is assigned to it. The station then sets a request bit in a minislot on the reverse bus, and transfers the value of the RQ counter to the CD counter and resets the RQ counter to zero. The station then decreases the CD counter as each idle minislot passes by on the forward bus. When the CD counter reaches zero, the station waits for the next idle minislot, sets the B bit to one, writes the destination address and timestamp of the packet into the DA and TS fields, respectively. The station then tunes its tunable transmitter to the data wavelength defined by the position of the control minislot it just accessed, and transmits the packet on that wavelength at the beginning of the next slot.

To receive data, each station constantly monitors the control channel. Whenever it sees its address announced in a minislot, it tunes its receiver to the corresponding wavelength to receive the packet at the next slot boundary. In the case of a destination conflict where more than one packet is addressed to the same destination in a slot, the one with the smallest timestamp wins the contention. To notify those source stations of the outcome of their transmissions in a slot, the headend examines the minislots in the same slot as they pass by, computes the results of destination conflicts (if any) according to timestamp ordering, and writes the outcome of the destination conflicts into the acknowledgment field in the next slot launched on the reverse bus. An ACK bit set to one means the failure of the associated transmission. Therefore, after a station transmits a packet on the forward bus, it must wait for some time and examine

the ACK bit on the reverse bus corresponding to its transmission to see if the transmission was successfully received by the destination. If the packet was lost in a destination conflict, then the source station must repeat the reservation procedure all over again.

4.3 Performance Analysis

In this analysis we assume uniform traffic. We further assume that there is infinite buffer space at each station, and new packets are generated only at the moments just before the slot boundaries reach the station. Define p as the probability that a station will generate a new packet in a slot. A new packet goes to any of the other $(N-1)$ stations with the same probability $(\frac{1}{N-1})$. For the system to be stable, we require $Np/2W < 1$. That is, $p < 2W/N$. However, the maximum value of p may be substantially smaller than $2W/N$ because of destination conflicts and the resultant retransmissions. Denote p_{ij} as the carried traffic intensity from node i to node j . We have $p_{ij} = \frac{p}{N-1}, 1 \leq i, j \leq N, i \neq j$. Define p'_{ij} as the total traffic load from i to j including retransmissions. Since the operations of the buses are symmetric, we shall analyze the performance for traffic on the forward bus only.

Since a destination conflict is resolved by the ordering of timestamps, in the long run we expect that all stations involved in a destination conflict will have the same probability to win. Therefore, given j , all the p'_{ij} 's, $i < j$, should be the same. Denote $\gamma_j = p'_{ij}, j = 2, \dots, N, i < j$. We can also interpret γ_j as the probability that a station will transmit a packet to node j on the forward bus in a slot.

Let q_j be the probability that a packet transmission to node j is successful

(i.e., it wins the destination conflict, if any, and it is successfully received by node j). Then we have $\gamma_j = p'_{ij} = \frac{p_{ij}}{q_j} = \frac{p}{(N-1)q_j}$. Now consider a packet (the *tagged* packet) transmitted to node j in a slot. Given that there are k more packets (besides the tagged one) addressed for node j transmitted in the same slot, the tagged will win the destination conflict with the probability $1/(k+1)$. Since, in addition to the node transmitting the tagged packet, there are $(j-2)$ other nodes that also generate traffic for node j on the forward bus, the probability that there are k other packets addressed for node j in the same slot is equal to $\frac{1}{k+1} \binom{j-2}{k} \gamma_j^k (1-\gamma_j)^{j-2-k}$. Summing over all possible values of k , we have

$$q_j = \sum_{k=0}^{\min(j-2, W-1)} \frac{1}{k+1} \binom{j-2}{k} \gamma_j^k (1-\gamma_j)^{j-2-k}$$

where k must be less than or equal to $W-1$ because there can be at most W packets transmitted in a slot. We then obtain γ_j as shown below:

$$\gamma_j = \frac{p/(N-1)}{\sum_{k=0}^{\min(j-2, W-1)} \frac{1}{k+1} \binom{j-2}{k} \gamma_j^k (1-\gamma_j)^{j-2-k}} \quad (4.1)$$

where γ_j can be solved numerically for a given p . For $j \leq W+1$, Equation (4.1) can be transformed to

$$\frac{1}{j-1} [1 - (1-\gamma_j)^{j-1}] = \frac{p}{N-1}$$

Solving the above equation, we obtain γ_j ,

$$\gamma_j = 1 - \sqrt[j-1]{1 - \frac{(j-1)p}{N-1}} \quad (4.2)$$

4.3.1 Capacity of the System

Given the γ_j 's, the total load (including retransmissions) applied by station i , $1 \leq i \leq N - 1$, on the forward bus is defined as $\Gamma_i \triangleq \sum_{j=i+1}^{N-1} \gamma_j$ (note that node N generates no traffic for the forward bus), and the overall load on the forward bus sums up to $\Gamma \triangleq \sum_{i=1}^{N-1} \Gamma_i = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \gamma_j$. The capacity of a single bus is that value of p , say p^* , that sets the overall load, Γ , equal to W . p^* is the maximal rate for new packet generation at each station, and $N \cdot p^*$ is equal to the capacity of the system.

4.3.2 Mean Access Delays for Individual Stations

We first define $T_i =$ mean packet access delay for node i (in number of slots), where the access delay is the time interval from the instant a packet is generated until the beginning of its successful transmission. Since each request can be satisfied by a minislot (which in turn defines a slot on a certain data channel) on the control channel, we shall express the parameters in units of minislots. We assume that the total requests generated by node i and the downstream nodes arrive at node i follow a geometric process with rate $\sum_{j=i}^{N-1} \Gamma_j/W$ requests/minislot. We further assume that the idle minislots which arrive at node i also follow a geometric process with rate equal to one minus the load from the upstream nodes, i.e., $1 - (\sum_{j=1}^{i-1} \Gamma_j/W)$. Note that we have effectively modeled the queue at node i serving local requests and requests from downstream nodes as a Geom/Geom/1 queue. Using the results in [Kle75], the average number of packets (waiting to access minislots) in node i 's buffer can be found equal to

$$\begin{aligned}
N_{WM} &= \frac{\Gamma_i}{\sum_{j=i}^{N-1} \Gamma_j} \cdot \frac{\sum_{j=i}^{N-1} \frac{\Gamma_j}{W} (1 - \sum_{j=i}^{N-1} \frac{\Gamma_j}{W})}{(1 - \sum_{j=1}^{i-1} \frac{\Gamma_j}{W}) - \sum_{j=i}^{N-1} \frac{\Gamma_j}{W}} \\
&= \frac{\Gamma_i}{W} \cdot \frac{1 - \sum_{j=i}^{N-1} \frac{\Gamma_j}{W}}{1 - \sum_{j=1}^{N-1} \frac{\Gamma_j}{W}}
\end{aligned}$$

Let d_i be the round-trip propagation delay (in minislots) between node i and the downstream headend. Node i does not realize the result of the failure of those packets it transmitted and which lost a destination conflict until the acknowledgments come back on the reverse bus. For those packets transmitted by node i that failed in destination conflicts, node i must repeat the reservation procedure to reschedule their retransmissions, and the traffic intensity of retransmission at node i equals to the total load of node i minus the rate that new packets arrive at node i , i.e., $(\Gamma_i - \frac{p(N-i)}{N-1})/W$ (remember that we only consider traffic on the forward bus). Therefore, by Little's result [Kle75], we have that the average number of packets at node i waiting for acknowledgments is

$$N_{WA} = d_i (\Gamma_i - \frac{p(N-i)}{N-1}) / W \quad (4.3)$$

The average number of packets at node i is equal to $N_{WM} + N_{WA}$. From Little's result we then have

$$\begin{aligned}
T_i &= \frac{N_{WM} + N_{WA}}{\frac{p(N-i)}{N-1}} && \text{minislots} \\
&= \frac{1}{W} \cdot \frac{N_{WM} + N_{WA}}{\frac{p(N-i)}{N-1}} && \text{slots}
\end{aligned}$$

4.4 Numerical Results

Here we assume a dual bus network where $N=20$ stations are attached and are numbered 1 to 20 from left to right. The bus length is equal to 10 slots, and stations are placed uniformly along the bus at equal distances. In this case, $d_i = W \cdot 2 \cdot 10 \cdot (N - i)/(N - 1)$ in minislots. We define the utilization factor $\rho \triangleq \frac{Np}{2W}$ as the ratio of the total new packet generation rate over the maximum capacity of the system.

In Figure 4.4 we plot simulation results for the total throughput on one bus versus ρ for different numbers of wavelengths. We see that systems with a larger number of wavelengths can support a larger throughput. The maximum throughput, however, is always less than W because some of the transmissions are lost (and thus wasted) in destination conflicts. Also note that our analysis predicts the system's capacity (maximum throughput) very well. In Figure 4.5 we normalize the throughput to the number of wavelengths, and we see that the maximum efficiency of each wavelength drops as W increases since the chance of destination conflicts becomes larger as more packets can be transmitted in a slot. Figure 4.6 plots the average access delay over all the stations versus the traffic load.

Next we plot mean access delays (for traffic on the forward bus only) for individual stations in Figures 4.7 and 4.8 for $W=4$ and $W=16$, respectively. It is well-known [KC90] that single-channel DQDB favors the upstream nodes while downstream nodes suffer longer access delay. In our protocol, however, stations at both ends experience smaller delays than stations located in the middle. This is because downstream stations get to see the acknowledgments sooner than upstream stations. This effect, combined with the favoring of

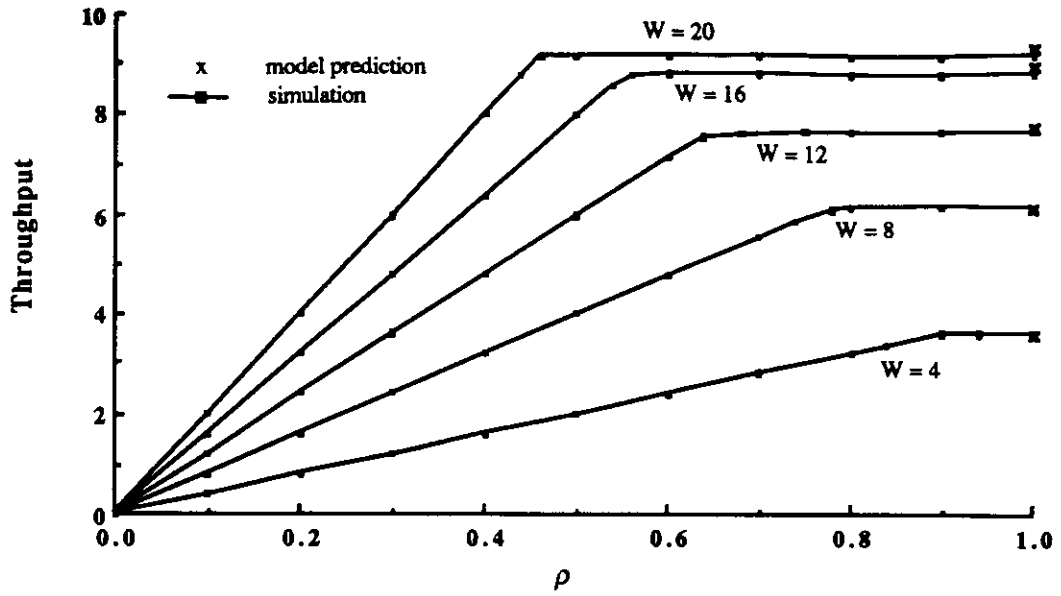


Figure 4.4: Total throughput versus traffic load.

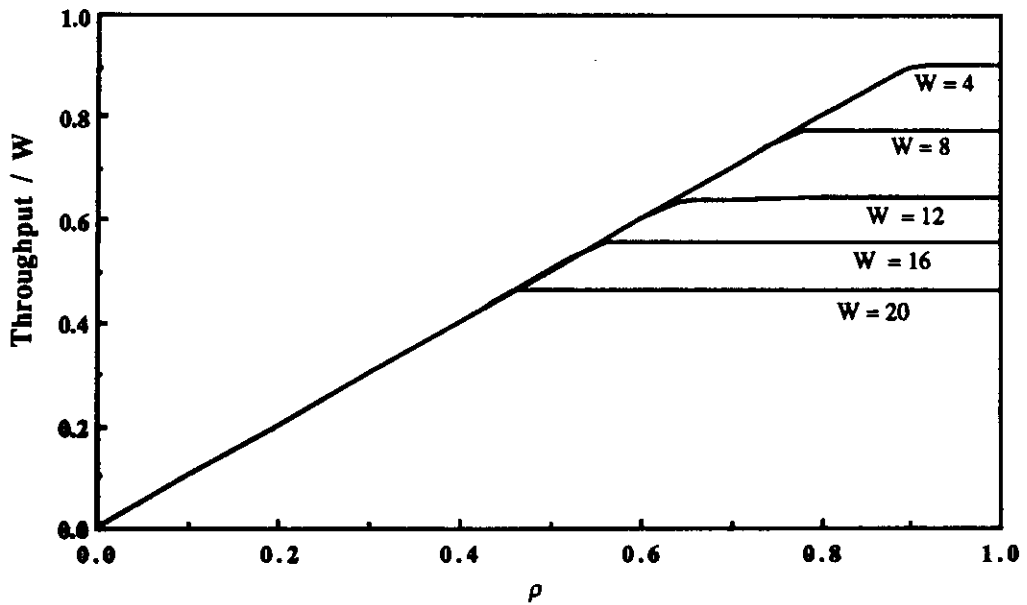


Figure 4.5: Channel efficiency versus traffic load.

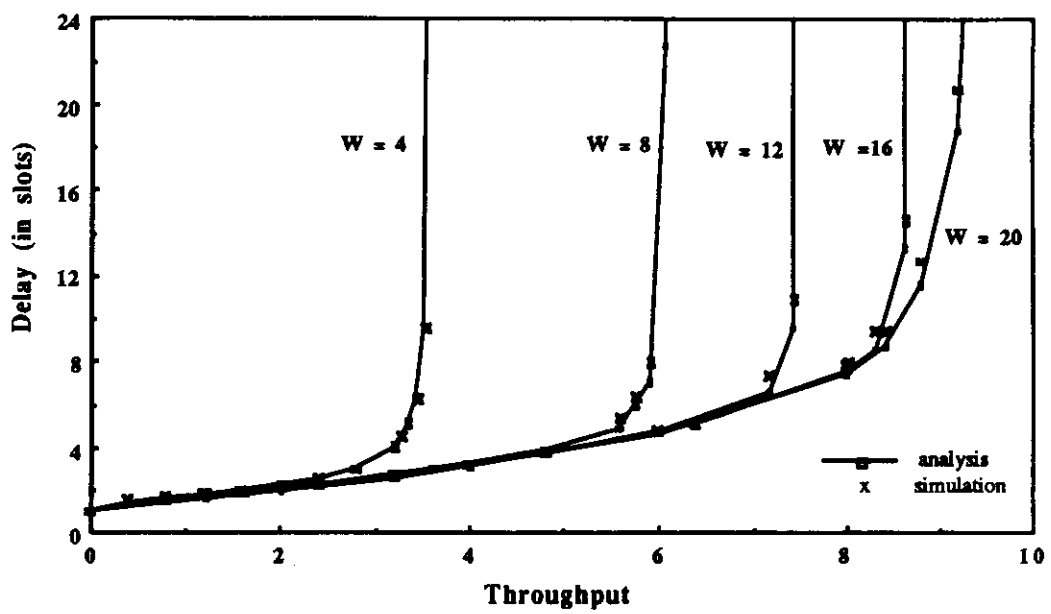


Figure 4.6: Throughput versus delay curves.

upstream nodes by single-channel DQDB, flattens out the delay curve. That is, our protocol achieves better fairness than the DQDB protocol.

4.5 Conclusions

In this chapter we designed a high-speed dual bus network and proposed a WDMA protocol for its medium access control. This is, in effect, a multichannel DQDB system. The numerical results demonstrated that throughput higher than the speed of single electronic interface can be achieved. The capacity of the system increases as more wavelengths are available, but the efficiency of each wavelength drops because, as the ratio of W/N increases, more destination conflicts occur, which require more retransmissions. We also note that the average packet delays for different stations did not differ by much, which is much fairer than the single-channel DQDB network, because of the longer delays experienced by upstream stations to receive the acknowledgments (which offsets the intrinsic unfairness in the single-channel DQDB network).

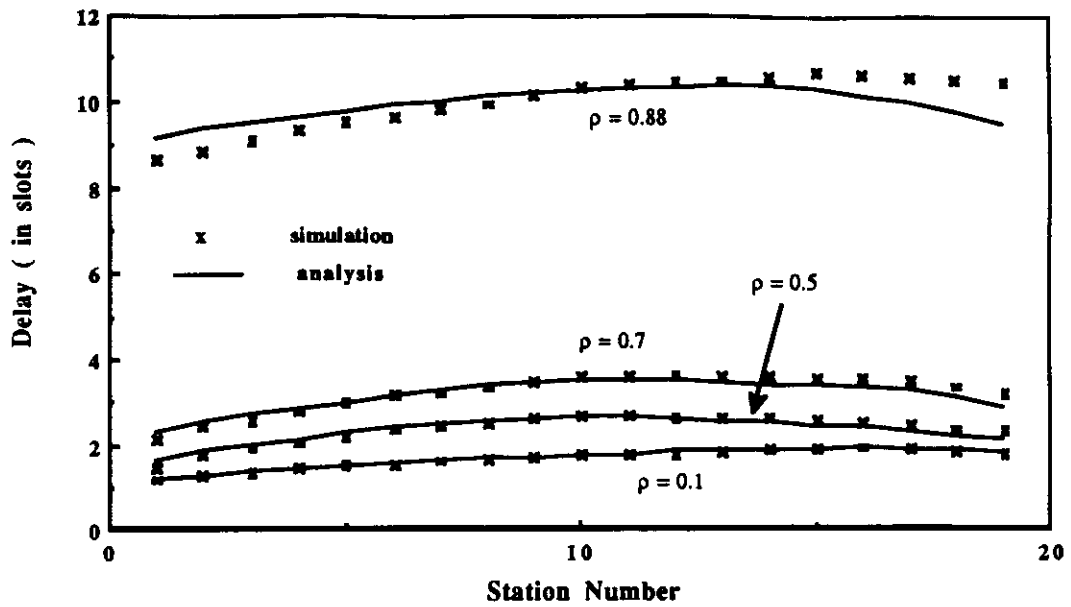


Figure 4.7: Average access delays for individual stations. $W=4$.

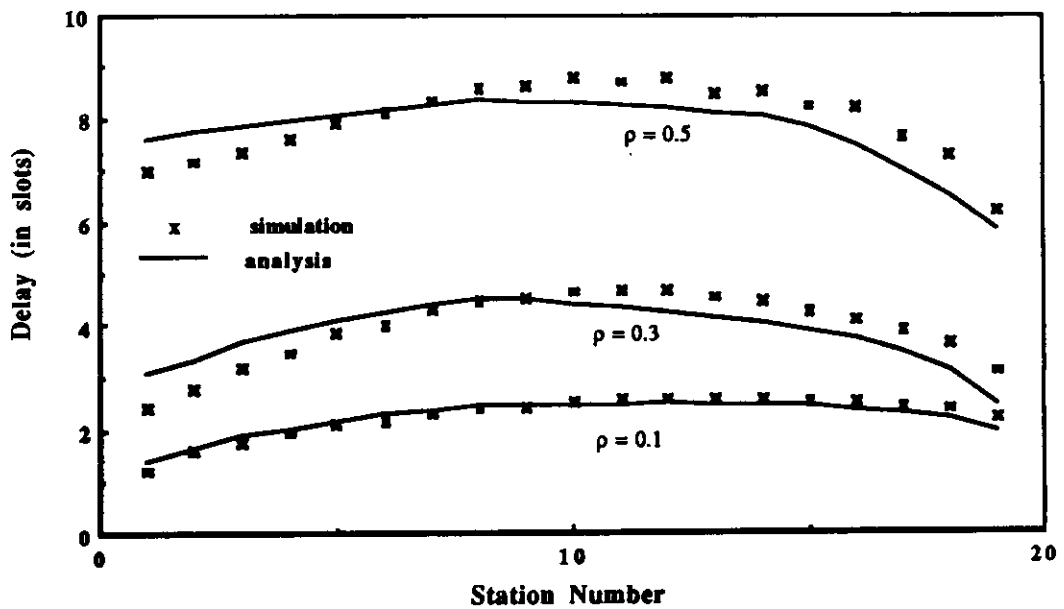


Figure 4.8: Average access delays for individual stations. $W=16$.

CHAPTER 5

Conclusions and Future Research

In this dissertation we have addressed the fundamentals, advantages, and limitations of WDMA systems. In Chapter 1 we presented a number of future applications which will most likely require the support of a high-capacity communication network, and pointed out the bottleneck which exists at the speed of the electronic interface. WDMA is a very promising concurrency technique because it avoids this bottleneck by not requiring faster electronics. Also the feasibility of key technology components (e.g., tunable lasers, tunable filters, couplers, optical amplifiers, etc.) has been demonstrated in the laboratories. The deployment of this technology will depend upon cost considerations. There are many WDMA network designs and experiments, some of which are described in this dissertation, that have been conducted in various universities and research organizations. We expect to see more in the future.

In Chapter 2, we constructed an approximate mathematical model to study the effects of resource contention (of receivers, transmitters, and wavelengths) in WDMA networks with a general receiver/transmitter configuration, an arbitrary traffic patterns, and an ideal access mechanism. An important result we observed is that, given uniform traffic and a number of wavelengths much fewer than the number of stations (which is most likely the near-term case), it is better to have both tunable transmitters and tunable receivers at each node

instead of having only one of them tunable; moreover, only a small number of tunable transmitters and tunable receivers at each node is enough to produce a performance close to the upper bound.

Based on this observation, in the next two chapters we proposed two WDMA protocols in both of which we assumed that the number of available wavelengths is fewer than the number of stations, and that each station is equipped with a small number of tunable transmitters and tunable receivers. In Chapter 3, we adopted the star topology for its power-efficiency, and showed that high throughput and low delay could be achieved. The protocol, however, is intrinsically unstable because of the slotted ALOHA employed, and some kind of dynamic control procedure is required to stabilize it. Considering the disadvantages of extending the star network to serve a large number of stations in a large area, and the recent impressive progress in broadband optical amplifiers, in Chapter 4 we presented a WDMA protocol for networks with a dual bus topology suitable for a metropolitan area. The proposed protocol is a multichannel generalization of the DQDB protocol, and numerical results showed that the protocol not only could support a throughput larger than the full electronic speed, but also achieved better fairness than the single-channel DQDB network. The performance of the protocols in those two chapters were mathematically analyzed and the correctness of these models was verified by comparing numerical results to simulation.

In summary, WDMA has the potential to provide us with the concurrency required to tap into the enormous optical bandwidth. We next point out some directions that require future research work (we exclude from these directions those pure technology breakthroughs that will allow us to access more wave-

lengths at higher speeds and better power budgets — these are beyond the scope of this dissertation).

5.1 WDMA Wide Area Networks

All the work covered in this dissertation focuses on optical LANs and MANs. An obvious next step is to apply the Wavelength Division Multiplexing (WDM) technology to build high-speed Wide Area Network (WAN) backbones to interconnect LANs and MANs. There have been a number of papers describing WAN designs employing WDM in the literature [Bra90][Ste90][CGK90][Che91]. The two general architectures for WDM WANs are *wavelength routing* and *wavelength switching*. Wavelength routing networks are composed of one or more passive wavelength-selective elements (e.g., wavelength division multiplexer/demultiplexer) plus some optical amplifiers. Signals transmitted on a wavelength stay on that wavelength while being routed through possibly several fiber trunks to the destination. There is no electro/opto conversion except at the source and the destination. The proposals in [Ste90][Che91] belong to this category.

In wavelength switching networks, signals transmitted on one wavelength may be effectively transferred to another wavelength in a switching office, which can be done by detecting the signals on the first wavelength and using the detected signals to modulate a laser for the second wavelength. One example is the *Lightnet* proposed in [CGK90] where a virtual link, called a *lightpath*, is created by allocating a single wavelength on all the physical links composing the virtual link. Messages may traverse several lightpaths before getting to their destinations.

Wavelength routing network enjoys the advantage of no electro/opto conversion inside the network. However, it does suffer from the noise accumulation in a cascade of optical amplifiers which is necessary to compensate for the power loss in transmitting in a wide area. On the contrary, a wavelength switching network does not have this noise accumulation problem since the signal coming in on an optical fiber can be detected and regenerated at the switching office before being reinjected onto another optical fiber. The price paid in this case is the time spent in opto/electro/opto conversions. More research on the design of high-speed WANs is necessary.

5.2 The Case When $W = N$

Another topic worth more investigation is the case where the number of wavelengths is as large as the number of stations in the system (i.e., $W = N$). In this case, wavelength is no longer the scarce source since each station can be assigned a unique wavelength to transmit on. Destination conflict is now the major factor that limits the system's throughput. Given uniform traffic and one tunable receiver at each station, it has been shown [CDR90] that the peak throughput of each station is $(1 - 1/(N - 1))^{N-1}$, which approaches $1 - 1/e \approx 0.632$.

Two schemes have been proposed [CY91][CZA92] which try to improve the above throughput limit without increasing the number of receivers at each node. Both require that each node maintain an $N \times N$ backlog matrix B , the element b_{ij} of which indicates the number of packets at station i destined for station j . All packet arrivals must first be announced on a TDM-based control channel to all stations in the network. At the beginning of each slot an identical algorithm is executed by all stations to determine which packets are allowed to be trans-

mitted in this slot. Several scheduling algorithms were proposed in these two papers. Numerical results in both indicate that the maximal throughput of each data channel can approach unity. There exist a number of disadvantages associated with these types of schemes: First, there is a minimum of one end-to-end propagation delay because of the need to announce new packet arrivals on the control channel. Second, new stations cannot easily join the system since they have no idea of the current value of the backlog matrix. Third is the scheduling overhead. The algorithm in [CY91] has a computational complexity of $O(N^2)$, while the one in [CZA92] requires the execution of an identical random number generator with the same seed by all stations for each data channel in each slot. Clearly new mechanisms are required to achieve better performance.

5.3 Protocols to Support Integrated Traffic

In all the studies of this dissertation, only bursty data traffic is considered. However, with the advent of BISDN it is anticipated that future high-speed networks will carry bursty data traffic as well as digitized voice, video, facsimile, and images. These source types possess different characteristics (such as stream traffic), and different requirements (such as real-time delivery and/or higher tolerance of errors), which make the integration of mixed-media traffic in a single network challenging. There has been much work [Muk91][Jsa83][Jsa87] proposed on the design of high-speed single-channel optical networks for integrated communications. However, very little work on multichannel integrated networks has been reported in the literature. Efficient multichannel access protocols are indispensable to support the requirements of high-speed applications in the future multimedia communications environment.

5.4 High-Speed Circuit/Packet Switch

WDM technology can also be applied to the construction of high-speed circuit/packet switches. As the data transmission rate in the proposed standards (e.g., Synchronous Optical NETWORK (SONET), DQDB, BISDN) increases, it is foreseeable that the central office of future telecommunication networks must possess a switching capacity of tens, or even hundreds, of Gb/s. Wavelength is another domain that can be exploited in order to alleviate the burden of pushing the speed of the electronic processing faster and faster. Several designs of high-speed packet switches employing WDM techniques have been proposed [AGKV88][Eng88][CB91]. More work is still necessary to build better high-speed switches.

APPENDIX A

A.1 Proof That the Denominator $D(z)$ Has W Roots on or Inside the Unit Circle $|z| = 1$

In this section we prove that, given $V/e < W$, the denominator, $D(z)$, of Equation (3.5) has W roots on or inside the unit circle $|z| = 1$. We first repeat the Rouché's theorem [Kle75] here:

Rouché's Theorem : If $f(z)$ and $g(z)$ are analytic functions of z inside and on a closed contour C , and also if $|g(z)| < |f(z)|$ on C , then $f(z)$ and $f(z) + g(z)$ have the same number of zeroes inside C .

From Equation (3.5), we have $D(z) = (1 - \frac{G}{V}e^{-\frac{G}{V}} + \frac{Gz}{V}e^{-\frac{G}{V}})^V - z^W$. Let $f(z) = -z^W$ and $g(z) = (1 - \frac{G}{V}e^{-\frac{G}{V}} + \frac{Gz}{V}e^{-\frac{G}{V}})^V$. For simplicity, define $x = \frac{G}{V}e^{-G/V}$. We have $0 \leq x \leq e^{-1}$ for $G \in [0, \infty)$. Now $g(z)$ becomes $(1 - x + xz)^V$. Define a closed contour C as the circle $|z| = 1 + \delta$, where δ is a very small positive number. On the contour C ,

$$|f(z)| = |z|^W = (1 + \delta)^W = 1 + W\delta + o(\delta^2)$$

and

$$\begin{aligned}
|g(z)| &= |1 - x + xz|^V \\
&\leq (|1 - x| + |xz|)^V = (1 - x + x(1 + \delta))^V \\
&= (1 + x\delta)^V = 1 + Vx\delta + o(\delta^2) \\
&< 1 + \frac{V\delta}{e} + o(\delta^2)
\end{aligned}$$

Given that $V/e < W$, there always exists a value of δ , say $\delta^* > 0$, such that $\forall \delta \in (0, \delta^*), 1 + \frac{V\delta}{e} + o(\delta^2) < 1 + W\delta + o(\delta^2)$. If we choose C as $|z| = 1 + \delta, \delta \in (0, \delta^*)$, then we have $|g(z)| < |f(z)|$ on C . Because $f(z) = -z^W$ has W zeroes inside C , from Rouché's theorem we know that $D(z) = f(z) + g(z)$ also has W zeroes inside C . Since δ can take on any value between 0 and δ^* for the contour C , this is equivalent to say that those W roots of $D(z)$ must be on or inside the unit circle $|z| = 1$.

Q.E.D.

A.2 Derivations of $f_T(\bar{n})$ and $g_Q(i)$

In this section we derive the functions $f_T(\bar{n})$ and $g_Q(i)$ in Equations (3.9) and (3.10), respectively. We first derive $f_T(\bar{n})$, the average number of stations that successfully transmit a reservation minipacket in a slot, given that the system is in state \bar{n} . Suppose that i reservation minipackets are transmitted in a slot. The probability that a particular one of them is successful is equal to $(1 - 1/V)^{i-1}$ and the average number of successful reservations is $i(1 - 1/V)^{i-1}$. Also, the conditional probability that i reservation minipackets are transmitted in a slot

given that the system is in state \bar{n} is

$$\binom{\bar{n}_r}{i} p^i (1-p)^{\bar{n}_r-i}$$

Therefore, we have

$$\begin{aligned} f_r(\bar{n}) &= \sum_{i=1}^{\bar{n}_r} i \left(1 - \frac{1}{V}\right)^{i-1} \binom{\bar{n}_r}{i} p^i (1-p)^{\bar{n}_r-i} \\ &= \bar{n}_r p \left(1 - \frac{p}{V}\right)^{\bar{n}_r-1} \end{aligned}$$

Next we derive $g_q(i)$, the average number of successfully transmitted data packets given that i data packets are transmitted, under the assumption that a station does not transmit to itself. Consider a destination station, say station k , for example. The probability that station k is also the source of one of the i transmitted packets is i/N , and given this, the probability that none of the other $(i-1)$ packets are destined to station k is $(1 - \frac{1}{N-1})^{i-1}$. The probability that none of the i packets is transmitted by station k is $(1 - i/N)$, and given this the probability that none of the i packets are destined to station k is $(1 - \frac{1}{N-1})^i$. Therefore, the probability that at least one among those i packets is going to station k (which is also equal to the average number of packets successfully received by station k) is equal to

$$1 - \left[\frac{i}{N} \left(1 - \frac{1}{N-1}\right)^{i-1} + \left(1 - \frac{i}{N}\right) \left(1 - \frac{1}{N-1}\right)^i \right]$$

Therefore, we have

$$\begin{aligned} g_q(i) &= N \left[1 - \left[\frac{i}{N} \left(1 - \frac{1}{N-1}\right)^{i-1} + \left(1 - \frac{i}{N}\right) \left(1 - \frac{1}{N-1}\right)^i \right] \right] \\ &= N \left[1 - \left(1 - \frac{1}{N-1}\right)^{i-1} \left(\frac{N^2 - 2N + i}{N(N-1)}\right) \right] \end{aligned}$$

REFERENCES

- [Aca87] Anthony S. Acampora. "A Multichannel Multihop Local Lightwave Network." In *GLOBECOM '87*, pp. 37.5.1–37.5.9, 1987.
- [ACG⁺88] Edward Arthurs, J. M. Cooper, Matthew S. Goodman, Haim Kobrinski, M Tur, and Mario P. Vecchi. "Multiwavelength Optical Crossconnect for Parallel-Processing Computers." *IEE Electron. Lett.*, **24**:119–120, 1988.
- [AGKV88] Edward Arthurs, Matthew S. Goodman, Haim Kobrinski, and Mario P. Vecchi. "HYPASS: An Optoelectronic Hybrid Packet Switching System." *IEEE J. Select. Areas Commun.*, **6**(9):1500–1510, December 1988.
- [AK89] Anthony S. Acampora and Mark J. Karol. "An Overview of Lightwave Packet Networks." *IEEE Network*, **3**(1):29–41, January 1989.
- [Bel91] T. E. Bell. "Technology 1991: Telecommunications." *IEEE Spectrum*, **28**(1):44–47, January 1991.
- [BFG90] Joseph A. Bannister, Luigi Fratta, and Mario Gerla. "Topological Design of the Wavelength-Division Optical Network." In *IEEE INFOCOM '90*, pp. 1005–1013, 1990.
- [BKNS89] William R. Byrne, Toomas A. Kilm, Bruce L. Nelson, and Marius D. Soneru. "Broadband ISDN Technology and Architecture." *IEEE Network*, pp. 23–28, January 1989.
- [BM92] Subrata Banerjee and Biswanath Mukherjee. "An Efficient and Fair Protocol for a Multichannel Lightwave Network." In *Tech. Digest, OFC '92*, p. 183, 1992.
- [Bra90] Charles A. Brackett. "Dense Wavelength Division Multiplexing Networks: Principles and Applications." *IEEE J. Select. Areas Commun.*, **8**(6):948–964, August 1990.
- [Cat92] Charles E. Catlett. "In Search of Gigabit Applications." *IEEE Commun. Mag.*, pp. 42–51, April 1992.
- [CB91] Arturo Cisneros and Charles A. Brackett. "A Large ATM Switch Based on Memory Switches and Optical Star Couplers." *IEEE J. Select. Areas Commun.*, **9**(8):1348–1360, October 1991.

- [CDR90] Mon-Song Chen, Nicholas R. Dono, and Rajiv Ramaswami. "A Media-Access Protocol for Packet-Switched Wavelength Division Multiaccess Metropolitan Area Networks." *IEEE J. Select. Areas Commun.*, 8(6):1048–1057, August 1990.
- [CF91] Imrich Chlamtac and Andrea Fumagalli. "QUADRO-Stars: High Performance Optical WDM Star Networks." In *GLOBECOM '91*, pp. 34.4.1–34.4.6, 1991.
- [CG88] Imrich Chlamtac and Aura Ganz. "A Multibus Train Communication (AMTRAC) Architecture for High-Speed Fiber Optic Networks." *IEEE J. Select. Areas Commun.*, 6(6):903–912, July 1988.
- [CG89] Imrich Chlamtac and Arua Ganz. "Design Alternatives of Asynchronous WDM Star Networks." In *ICC '89*, pp. 739–744, 1989.
- [CG90] Imrich Chlamtac and Aura Ganz. "Toward Alternative High-Speed Network Concepts: The SWIFT Architecture." *IEEE Trans. Commun.*, 38(4):431–439, April 1990.
- [CGK90] Imrich Chlamtac, Aura Ganz, and G. Karmi. "Lightnet: Lightpath Based Solutions for Wide Bandwidth WANs." In *IEEE INFOCOM '90*, pp. 1014–1021, 1990.
- [Che90] Kwok-Wai Cheung. "Acoustoopic Tunable Filters in Narrowband WDM Networks: System Issues and Network Applications." *IEEE J. Select. Areas Commun.*, 8(6):1015–1025, August 1990.
- [Che91] Kwok-Wai Cheung. "Scalable, Fault-Tolerant 1-Hop Wavelength Routing Networks." In *GLOBECOM '91*, pp. 1240–1245, 1991.
- [CY91] Ming Chen and Tak-Shing Yum. "A Conflict-Free Protocol for Optical WDMA Networks." In *GLOBECOM '91*, pp. 35.6.1–35.6.6, 1991.
- [CZA92] R. Chipalkatti, Z. Zhang, and A. S. Acampora. "High-Speed Communication Protocols for Optical Star Networks Using WDM." In *IEEE INFOCOM '92*, pp. 2124–2133, 1992.
- [DGL+90] Nicholas R. Dono, Paul E. Green, Jr., Karen Liu, Rajiv Ramaswami, and Franklin Fuk-Kay Tong. "A Wavelength Division Multiple Access Network for Computer Communication." *IEEE J. Select. Areas Commun.*, 8(6):983–994, August 1990.

- [EM88] Martin Eisenberg and Nader Mehravari. "Performance of the Multichannel Multihop Lightwave Network Under Nonuniform Traffic." *IEEE J. Select. Areas Commun.*, 6(7):1063–1078, August 1988.
- [Eng88] Kai Y. Eng. "A Photonic Knockout Switch for High-Speed Packet Networks." *IEEE J. Select. Areas Commun.*, 6(7):1107–1116, August 1988.
- [GC89] Aura Ganz and Imrich Chlamtac. "Path Allocation Access Control in Fiber Optic Communication Systems." *IEEE Trans. Comput.*, 38(10):1372–1382, October 1989.
- [GG92] Aura Ganz and Yao Gao. "A Time-Wavelength Assignment Algorithm for a WDM Star Network." In *IEEE INFOCOM '92*, pp. 2144–2150, 1992.
- [GK91] Aura Ganz and Zahava Koren. "WDM Passive Star - Protocols and Performance Analysis." In *IEEE INFOCOM '91*, pp. 991–1000, 1991.
- [GKV⁺90] Matthew S. Goodman, Haim Kobrinski, Mario P. Vecchi, Ray M. Bulley, and James L. Gimlett. "The LAMBDANET Multiwavelength Network: Architecture, Applications, and Demonstrations." *IEEE J. Select. Areas Commun.*, 8(6):995–1004, August 1990.
- [Goo89] Matthew S. Goodman. "Multiwavelength Networks and New Approaches to Packet Switching." *IEEE Commun. Mag.*, pp. 27–35, October 1989.
- [Gre91] Paul E. Green. "The Future of Fiber-Optic Computer Networks." *IEEE Comput. Mag.*, pp. 78–87, September 1991.
- [Hen89] Paul S. Henry. "High-Capacity Lightwave Local Area Networks." *IEEE Commun. Mag.*, pp. 20–26, October 1989.
- [Hil89] Godfrey R. Hill. "Wavelength Domain Optical Network Techniques." *Proceedings of the IEEE*, 77(1):121–132, January 1989.
- [HK88] Michael G. Hluchyj and Mark J. Karol. "ShuffleNet : An Application of generalized Perfect Shuffles to Multihop Lightwave Networks." In *IEEE INFOCOM '88*, pp. 379–390, 1988.

- [HKS87] Isam M. I. Habbab, Mohsen Kavehrad, and Carl-Erik W. Sundberg. "Protocols for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology." *J. Lightwave Technol.*, LT-5(12):1782-1794, December 1987.
- [JRS92] Frank J. Janniello, Rajiv Ramaswami, and David G. Steinberg. "A Prototype Circuit-Switched Multi-Wavelength Optical Metropolitan-Area Network." In *ICC '92*, pp. 818-823, 1992.
- [Jsa83] "Special Issue on Packet Switching Voice and Data Communication." *IEEE J. Select. Areas Commun.*, SAC-1, December 1983.
- [Jsa87] "Special Issue on Switching Systems for Broadband Networks." *IEEE J. Select. Areas Commun.*, SAC-5, October 1987.
- [KC89] Haim Kobrinski and Kwok-Wai Cheung. "Wavelength-Tunable Optical Filters: Applications and Technologies." *IEEE Commun. Mag.*, pp. 53-63, October 1989.
- [KC90] Harbans Kaur and Graham Campbell. "DQDB - An Access Delay Analysis." In *IEEE INFOCOM '90*, pp. 630-635, 1990.
- [KL75] Leonard Kleinrock and Simon S. Lam. "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation." *IEEE Trans. Commun.*, COM-23(4):410-423, April 1975.
- [Kle75] Leonard Kleinrock. *Queueing Systems, Vol. I: Theory*. John Wiley and Sons, New York, 1975.
- [Kle76] Leonard Kleinrock. *Queueing Systems, Vol. II: Computer Applications*. John Wiley and Sons, New York, 1976.
- [Kob89] Haim Kobrinski. "Simultaneous Fast Wavelength Switching and Direct Data Modulation Using a 3-Sectioned DBR Laser with 2.2 nm Continuous Tuning Range." In *Tech. Dig. OFC '89*, 1989. Paper PD3.
- [KS88] Mark J. Karol and Salman Shaikh. "A Simple Adaptive Routing Scheme for ShuffleNet Multihop Lightwave Networks." In *GLOBECOM '88*, pp. 1640-1647, 1988.
- [LA90] Jean-Francois P. Labourdette and Anthony S. Acampora. "Wavelength Agility in Multihop Lightwave Networks." In *IEEE INFOCOM '90*, pp. 1022-1029, 1990.

- [Lam74] Simon S. Lam. "Packet Switching in a Multi-Access Broadcast Channel with Applications to Satellite Communication in a Computer Network." Technical Report UCLA-ENG-7429, Computer Science Department, School of Engineering and Applied Science, March 1974.
- [LF82] John O. Limb and C. Flores. "Description of Fasnet – A Unidirectional Local Area Communications Network." *Bell Syst. Tech. J.*, 61:1413–1440, 1982.
- [LZ89] Tien Pei Lee and Chung-En Zah. "Wavelength-Tunable and Single-Frequency Semiconductor Lasers for Photonic Communications Networks." *IEEE Commun. Mag.*, pp. 42–52, October 1989.
- [Meh90] Nader Mehravari. "Performance and Protocol Improvements for Very High Speed Optical Fiber Local Area Networks Using a Passive Star Topology." *J. Lightwave Technol.*, 8(4):520–530, April 1990.
- [MMK87] S. Murata, I. Mito, and K. Kobayashi. "Over 720 GHz (5.8nm) Frequency Tuning by a 1500 nm DBR Laser with Phase and Bragg Wavelength Control Regions." *IEE Electron. Lett.*, 23:403–405, April 1987.
- [Muk91] Biswanath Mukherjee. "Integrated Voice-Data Communication Over High-Speed Fiber Optic Networks." *IEEE Comput. Mag.*, pp. 49–58, February 1991.
- [Muk92] Biswanath Mukherjee. "WDM-Based Local Lightwave Networks Part I: Single-Hop Systems." *IEEE Network*, 6(3):12–27, May 1992.
- [NBH88] R. M. Newman, Z. L. Budrikis, and J. L. Hullett. "The QPSX Man." *IEEE Commun. Mag.*, pp. 20–28, April 1988.
- [OHLJ85] N. A. Olsson, J. Hegarty, R. A. Logan, and L. F. Johnson. "68.3 km Transmission with 1.37Tbit km/s Capacity Using Wavelength Division Multiplexing of Ten Single-Frequency Lasers at 1.5 μm ." *IEE Electron. Lett.*, 21:105–106, 1985.
- [OS91] Yoram Ofek and Moshe Sidi. "Design and Analysis of a Hybrid Access Control to an Optical Star Using WDM." In *IEEE INFOCOM '91*, pp. 20–31, 1991.

- [Pat81] Janak H. Patel. "Performance of Processor-Memory Interconnections for Multiprocessors." *IEEE Trans. Comput.*, **30**(10):771-780, October 1981.
- [Per85] Stewart D. Personick. *Fiber Optics*. Plenum Press, New York, 1985.
- [PS85] D. B. Payne and J. R. Stern. "Single Mode Optical Local Networks." In *GLOBECOM '85*, pp. 39.5.1-39.5.6, 1985.
- [Ros86] F. E. Ross. "FDDI - A Tutorial." *IEEE Commun. Mag.*, pp. 10-17, May 1986.
- [RP90] Rajiv Ramaswami and Rajesh Pankaj. "Tunability Needed in Multi-Channel Networks: Transmitters, Receivers, or Both?" Technical Report 16237, IBM, October 1990.
- [Shu89] P. W. Shumate. "Optical Fibers Reach Into Homes." *IEEE Spectrum*, **26**(2):43-47, February 1989.
- [Sta88] William Stallings. *Data and Computer Communications*. Macmillan Publishing Company, New York, second edition, 1988.
- [Ste90] Thomas E. Stern. "Linear Lightwave Networks: How Far Can They Go?" In *IEEE GLOBECOM 90*, pp. 1866-1872, 1990.
- [Sto71] Harold S. Stone. "Parallel Processing with the Perfect Shuffle." *IEEE Trans. Comput.*, **20**(2):153-161, February 1971.
- [TBF83] Fouad A. Tobagi, F. Borgonovo, and Luigi Fratta. "Expressnet: A High Performance Integrated-Services Local Area Network." *IEEE J. Select. Areas Commun.*, **SAC-1**(5):898-913, November 1983.
- [TI84] Shuji Tasaka and Yutaka Ishibashi. "A Reservation Protocol for Satellite Packet Communication - A Performance Analysis and Stability Considerations." *IEEE Trans. Commun.*, **COM-32**(8):920-927, August 1984.
- [WD83] R. Wyatt and W. J. Delvin. "10kHz Linewidth 1.5 um InGaAsP External Cavity Laser with 55 nm Tuning Range." *IEE Electron. Lett.*, **19**:110-112, 1983.
- [WND83] L. D. Westbrook, A. W. Nelson, and C. Dix. "New Approach to the Manufacture of Low Threshold 1500 nm Distributed Feedback Lasers." *IEE Electron. Lett.*, **19**:423-424, May 1983.