A WDMA PROTOCOL FOR MULTICHANNEL DQDB NETWORKS*

Jonathan C. Lu
Bellcore
Piscataway, NJ 08854

Leonard Kleinrock
University of California
Los Angeles, CA 90024

Abstract

Wavelength Division Multiplexing (WDM) provides another dimension of freedom to tap into the vast optical bandwidth. High-speed lightwave networks employing linear fiber bus topology have become attractive with the emergence of erbium-doped optical amplifiers. In this paper we propose a media access protocol for a WDM lightwave LAN/MAN assuming dual bus topology, which is an extension of the Distributed Queue Dual Bus (DQDB) network to the multichannel case. A large number of stations connected to both buses are supported using a small number of wavelengths. An approximate mathematical model is constructed to analyze the performance of the proposed protocol, with its accuracy verified by simulations. Numerical results show that the protocol can achieve very good throughput and delay performance. Moreover, the network demonstrates better fairness in access delays than single-channel DQDB network.

1 Introduction

The advances in lightwave technology over the past two decades have made optical fiber the transmission medium of choice for high-speed communication. The optical fiber provides us with a potential bandwidth of 30 THz in the 1300 nm and 1550 nm low-loss bands. This vast bandwidth can be better utilized by exploiting the emerging dense Wavelength Division Multiplexing (WDM) technology [1], which divides the optical bandwidth into a number of smaller-capacity channels operating at full electronic speed of, say, a few Gb/s. A very high throughput can be achieved by concurrently transmitting traffics belonging to different user pairs on channels at different wavelengths.

In WDM networks, each network interface node has access to multiple wavelengths, either with an array of fixed-tuned optical transmitters (lasers) and receivers (filters), or with a small number of tunable transmitters and receivers. There is a tradeoff between tuning time and tuning range in today's electro-optic technology [2, 3]. Devices tunable over a large portion of the optical spectrum have a slow tuning speed in the order of microseconds, while devices with nanosecond tuning speed can only tune to a limited number of channels.

In general, the physical topology of WDM networks can take the form of a star or a linear bus. There have been many proposed star network designs in recent years [5] – [8]. Although a star topology shows a better power-efficiency [4], it does 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 [1]. 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 Distributed Queue Dual Bus (DQDB) [9] networks in the IEEE 802.6 MAN standard assumes a dual bus topology. Thus, networks based on dual bus topology are expected to be very popular in the future. The design and implementation considerations for photonic dual bus networks have been discussed in [10, 11]. In this paper, we propose a WDMA protocol which is a generalization of the basic single-channel DQDB protocol to the multichannel case. Numerical results show that the protocol achieves a very good throughput/delay performance, and demonstrates better fairness than single-channel DQDB network.

In Section 2 of the paper, we first explain the basic operation of the DQDB protocol, then we extend it to the multichannel case. An approximate queueing model is constructed in Section 3 to solve for the capacity of the system and mean packet delays for each station. In Section 4, numerical results from both simulation and analysis are plotted and compared. Section 5 concludes the paper.

2 System Description

A basic dual bus network is shown in Figure 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 [12] and DQDB [9], have been proposed in the literature. The WDMA protocol proposed here 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.

2.1 DQDB Preliminary

An example DQDB network is drawn in Figure 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 queueing protocol to control the access to the fixed-length slots on the buses. We shall only explain the

*This work was supported in part by the Defense Advanced Research Projects Agency under Contract MDA 903-87-C0663, Parallel Systems Laboratory.

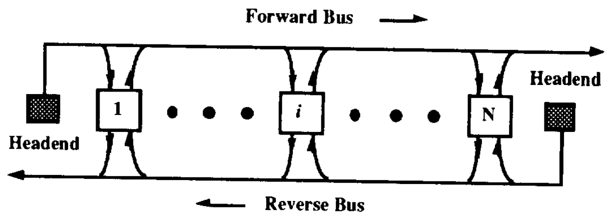

Figure 1: A basic dual bus network.

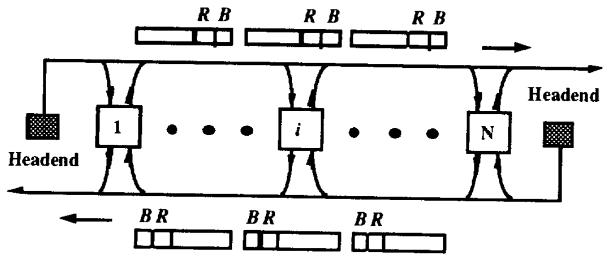

Figure 2: An example DQDB network.

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.

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, $\lambda_0, \lambda_1, \ldots, \lambda_W$, in the system, where the channel on wave-

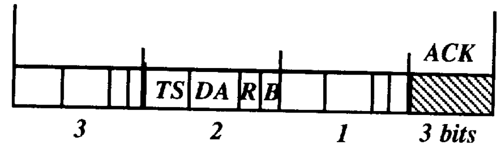

Figure 3: The structure of a slot on the control channel for $W = 3$.

length λ_0, called the *control channel*, is dedicated to the exchange of the control information, and the other W channels, called *data channels*, are used for the transmission of the actual data traffic. For each bus, each station has two lasers: one fixed laser tuned at λ_0 and the other tunable laser which can be tuned to any of the data wavelengths, $\lambda_1, \ldots, \lambda_W$, in a few nanoseconds. The outputs of those two lasers are merged by a coupler before transmitting into the fiber bus. For each bus, each station is also equipped with one filter fixed tuned at λ_0 and one tunable filter tunable to any of the wavelengths, $\lambda_1, \ldots, \lambda_W$, 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 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 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.

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 $1/(N-1)$. 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 $p_{ij} = p/(N-1), 1 \le i,j \le N, i \neq j$. Define p'_{ij} as the total traffic load from i to j including retransmissions. Because 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, for a given j, all the p'_{ij}'s, $i < j$, should be the same. Denote $\gamma_j = p'_{ij}, j = 2, \ldots, N, i < j$. 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 $\gamma_j = p'_{ij} = p_{ij}/q_j = \frac{p}{(N-1)q_j}$. Now consider a *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 one 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:

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

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

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

Solving the above equation, we obtain γ_j,

$$\gamma_j = 1 - \sqrt[j-1]{1-\frac{(j-1)p}{N-1}} \quad (2)$$

3.1 Capacity of the System

Given the γ_j's, the total load (including retransmissions) applied by station $i, 1 \le i \le N-1$, 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.

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., $1-(\sum_{j=1}^{j=i-1}\Gamma_j/W)$. 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 [13], the average number of packets (waiting to access minislots) in node i's buffer can be found equal to

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

$$= \frac{\Gamma_i}{W} \cdot \frac{1 - \sum_{j=i}^{N-1} \frac{\Gamma_j}{W}}{1 - \sum_{j=1}^{N-1} \frac{\Gamma_j}{W}}$$

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 in 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 [13], we have that the average number of packets at node i waiting for acknowledgments is equal to

$$N_{WA} = d_i(\Gamma_i - \frac{p(N-i)}{N-1})/W \qquad (3)$$

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

$$T_i = \frac{N_{WM} + N_{WA}}{\frac{p(N-i)}{N-1}} \quad \text{minislots}$$

$$= \frac{1}{W} \cdot \frac{N_{WM} + N_{WA}}{\frac{p(N-i)}{N-1}} \quad \text{slots}$$

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 $\rho \triangleq Np/2W$ as the ratio of the total new packet generation rate over the maximum capacity of the system.

In Figure 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 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 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 7 and 8 for W=4 and W=16, respectively. It is well-known [14] that single-channel DQDB without any bandwidth balancing scheme

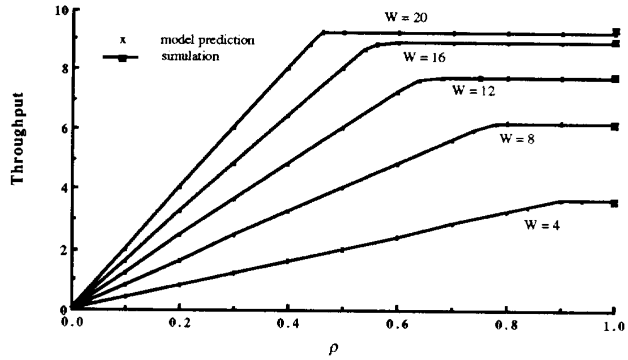

Figure 4: Total throughput versus traffic load.

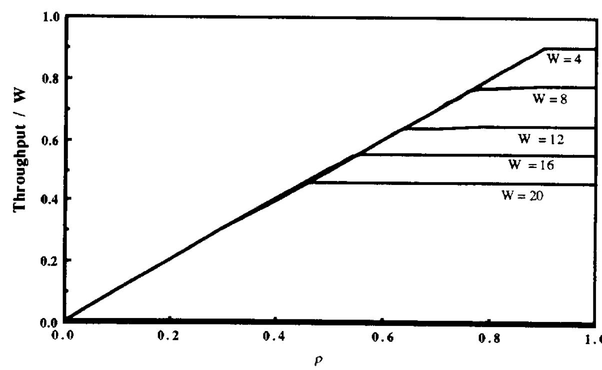

Figure 5: Channel efficiency versus traffic load.

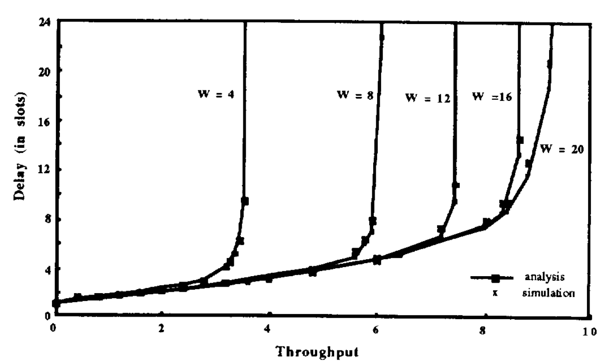

Figure 6: Throughput versus delay curves.

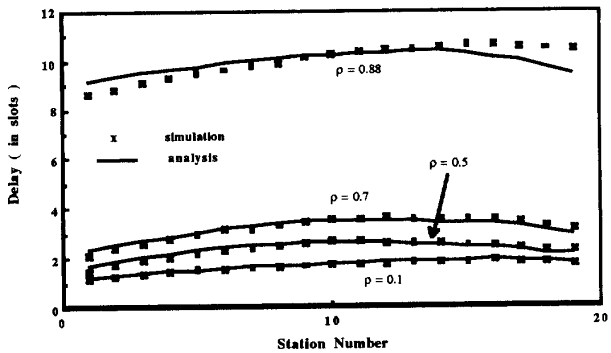

Figure 7: Average access delays for individual stations. W=4.

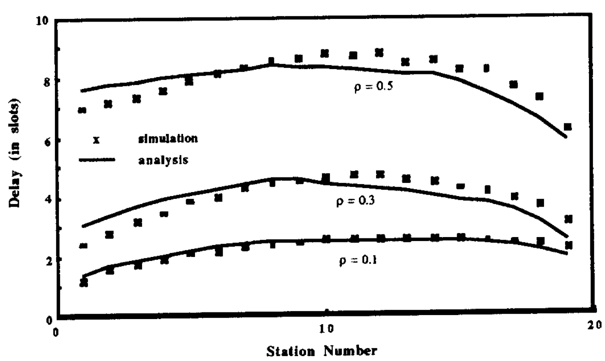

Figure 8: Average access delays for individual stations. W=16.

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 upstream nodes by single-channel DQDB, flattens out the delay curve. That is, our protocol achieves better fairness than the DQDB protocol with no bandwidth balancing.

5 Conclusions

In this paper 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 without bandwidth balancing, because of the longer delays experienced by upstream stations to receive the acknowledgments (which offsets the intrinsic unfairness in the single-channel DQDB network).

References

[1] C. A. Brackett, "Dense wavelength division multiplexing networks: Principles and applications," *IEEE J. Select. Areas Commun.*, vol. 8, no. 6, pp. 948–964, Aug. 1990.

[2] T. P. Lee and C. E. Zah, "Wavelength-tunable and single-frequency semiconductor lasers for photonic communications networks," *IEEE Commun. Mag.*, pp. 42–52, Oct.1989.

[3] H. Kobrinski and K. W. Cheung, "Wavelength-tunable optical filters: Applications and technologies," *IEEE Commun. Mag.*, pp. 53–63, Oct. 1989.

[4] P. S. Henry, "High-capacity lightwave local area networks," *IEEE Commun. Mag.*, pp. 20–26, Oct. 1989.

[5] I. M. I. Habbab, M. Kavehrad, and C. W. Sundberg, "Protocols for very high-speed optical fiber local area networks using a passive star topology," *J. Lightwave Technol.*, vol. LT-5, no. 12, pp.1782–1794, Dec. 1987.

[6] M. S. Chen, N. R. Dono, and R. Ramaswami, "A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks," *IEEE J. Select. Areas Commun.*, vol. 8, no. 6, pp.1048–1057, Aug. 1990.

[7] N. R. Dono, P. E. Green, Jr., *et al*, "A wavelength division multiple access network for computer communication," *IEEE J. Select. Areas Commun.*, vol. 8, no. 6, pp.983–994, Aug. 1990.

[8] J. C. Lu and L. Kleinrock, "A wavelength division multiple access protocol for high-speed local area networks with a passive star topology," *Performance Evaluation*, vol. 16, pp. 223–239, Nov. 1992.

[9] R. M. Newman, Z. L. Budrikis, and J. L. Hullett, "The QPSX MAN," *IEEE Commun. Mag.*, pp. 20–28, Apr. 1988.

[10] K. W. Cheung, "Design and implementation considerations for wavelength-division multiplexed (WDM) photonic dual bus," *ICC '92*, pp. 848–854, 1992.

[11] K. W. Cheung and V. M. Mak, "EQEB – a multi-channel extension of the DQDB protocol with tunable channel access," *GLOBECOM '92*, pp. 1610–1617, 1992.

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

[13] L. Kleinrock, *Queueing Systems, Vol. I: Theory*, John Wiley and Sons, New York, 1975.

[14] H. Kaur and G. Campbell, "DQDB - an access delay analysis," *IEEE INFOCOM '90*, pp. 630–635, 1990.