Performance Evaluation 16 (1992) 223–239
North-Holland

A wavelength division multiple access protocol for high-speed local area networks with a passive star topology *

Jonathan C. Lu and Leonard Kleinrock

Computer Science Department, University of California, Los Angeles, 3732L Boelter Hall, Los Angeles, CA 90024-1596, USA

Abstract

Lu, J.C. and L. Kleinrock, A wavelength division multiple access protocol for high-speed local area networks with a passive star topology, Performance Evaluation 16 (1992) 223–239.

In this paper a multiple access protocol is proposed for a system consisting of many high-speed bursty traffic stations interconnected via an optical passive star coupler. Each station has access over a range of wavelengths. Time is divided into fixed-sized slots. A station must reserve a wavelength first, then transmits the data on that wavelength. New stations can join the system anytime without any system reconfiguration. Broadcast and multicast traffic can also be easily supported. The performance of both the infinite and finite population cases has been modeled and analyzed. 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.

Keywords: wavelength division multiplexing; multiaccess protocol; local area networks; performance analysis; fiber optics.

1. Introduction

The advance of fiber optics technology in the past decade offers a transmission medium of wide bandwidth and low attenuation which is unseen before. In Local Area Networks (LANs), where distance is short and propagation loss is of little concern, it is the high bandwidth that makes fiber so attractive. It is conceivable that we could construct multiple access networks with a total capacity of around 50 terabits per second (Tb/s) by using the low-loss passband of the optical fibers spectrum (1200–1600 nm) [6]. An obstacle to realizing such high-capacity optical networks is the bottleneck at the speed of the electronic interface. High-speed single-channel networks such as Expressnet [17], FDDI [15], and DQDB [12] has a fundamental limitation that in these networks, the maximum throughput of the entire network is limited to the rate that can be supported by the electronics of one of the end user stations. Wavelength Division Multiple Access (WDMA) solves this problem by operating on multiple channels at different wavelengths, with each channel running at full electronic speed. However, a major obstacle lies in the control of the WDMA system to turning the enormous optical bandwidth into user-accessible capacity; it is necessary to develop efficient medium access techniques for packet communications in this environment.

Today's electronically tunable semiconductor lasers and filters can tune from one wavelength to another in a few nanoseconds. However, the tuning range is limited. Therefore, each node can only operate on a small number of wavelengths [1]. One class of WDMA networks can be constructed by the

Correspondence to: Mr. J.C. Lu, Computer Science Department, University of California, Los Angeles, 3732L Boelter Hall, Los Angeles, CA 90024-1596, USA. Email: lu@cs.ucla.edu.

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

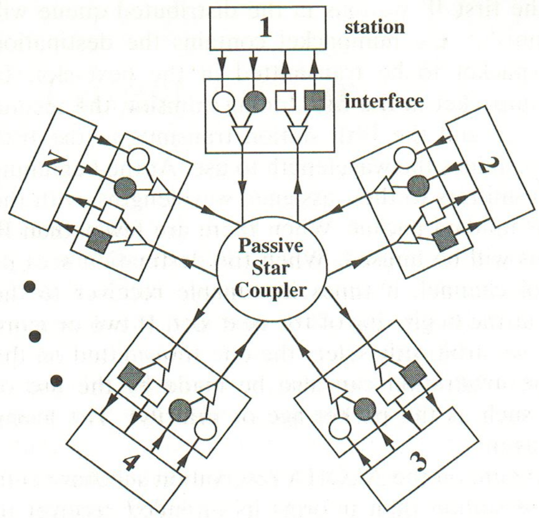

Fig. 1. N stations connected by a passive star coupler.

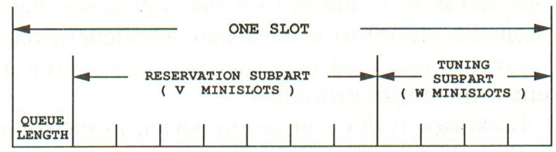

Fig. 2. Structure of a control slot.

of both the infinite and the finite population cases. In Section 4, numerical results from both analysis and simulation are given. Section 5 concludes the paper.

2. Description of protocol

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 is addressed and solved in [13]. Time is divided into fixed-sized slots. Packets are of 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 the 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 Fig. 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 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 is 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 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 transmit the reservation in this next slot with probability p, etc.

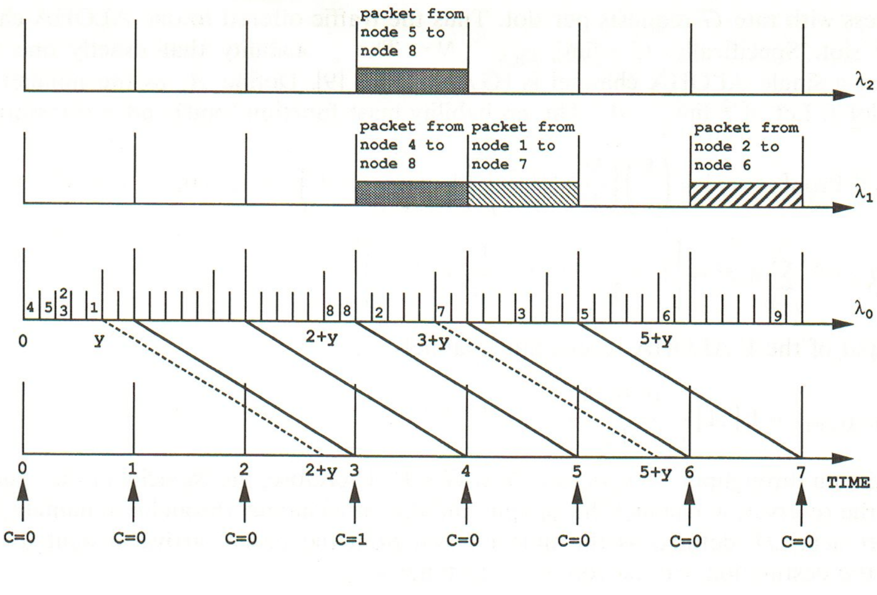

Fig. 3. A scenario of packet transmissions: $N = 10$, $V = 5$, $W = 2$, $R = 2$.

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 node 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 λ_1 and λ_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 λ_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.

3. Performance analysis

3.1. Model assumptions

We assume that there are $(W + 1)$ wavelengths 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 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) will not discard its packet until its successful reception at its destination is recognized.

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

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 [8], we get

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

where $c_j \triangleq \text{Prob}(C = j)$ is the pmf of C. We denote the denominator of $C(z)$ by $D(z)$. Using Rouche's theorem [8], it can be shown that W roots of $D(z)$ are on or inside the unit circle $|z| = 1$. 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 $z_1, z_2, \ldots, z_{(V-W)}$ 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

$$\text{E}[C] = \left.\frac{\mathrm{d}C(z)}{\mathrm{d}z}\right|_{z=1} = -\sum_{i=1}^{V-W} \frac{1}{1 - z_i} \tag{5}$$

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 Chapt. 5 in [8]). Since $\sum_{k=1}^{V} \text{Prob}(K = k)$ must equal to one, we have

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

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

$$\text{Prob}(X = j \mid K = k) = 1/k \quad j = 1, \ldots, k.$$

Unconditioning on k, we get

$$x_j = \sum_{k=j}^{V} \frac{1}{k} \frac{k a_k}{\text{E}[A]} = \frac{1}{\text{E}[A]} \sum_{k=j}^{V} a_k \quad j = 1, \ldots, V$$

and

$$\text{E}[X] = \sum_{j=1}^{V} j x_j = \frac{1}{2} + \frac{\text{E}[A^2]}{2\text{E}[A]} = 1 + \tfrac{1}{2}(V - 1)\frac{G}{V} e^{-G/V}$$

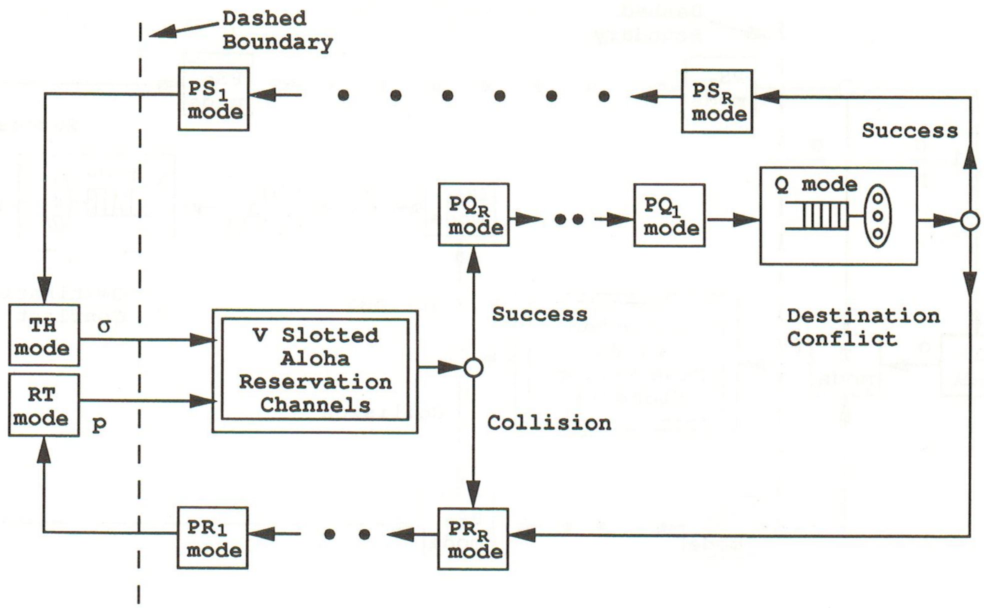

Fig. 6. An approximate model of the system.

following $(3R+3)$ *modes* (here R is assumed to be an interger) at any instant: TH, RT, Q, PQ_m, PR_m, and PS_m $(1 \leqslant m \leqslant R)$. Station 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 in the next slot. The PQ_m, PR_m, and PS_m modes, $m = 1, \ldots, R$, 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 Fig. 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 of 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, $m = 1, \ldots, R$. 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 the model still predicts the performance very well under this assumption.) Define the vector $\boldsymbol{n} \triangleq (n_{RT}, n_Q, i_1, \ldots, i_R, j_1, \ldots, j_R, k_1, \ldots, k_R)$ as the state vector of the system. Then we can see that the vector $\boldsymbol{n}$ 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) [16] to analyze this chain.

3.3.1. The modified model

To simplify the analysis, we first consider a modification of the model in Fig. 6 as suggested in [16], 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 $\sigma \leqslant p$. The modified model is shown in Fig. 7, where the TH mode in Fig. 6 has been decomposed into two modes, I and T, and the RT mode in Fig. 6 has become part of the T mode. A station that has just come

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

$$\delta_Q(\bar{n}) = X(\bar{n}) - \bar{j}_1 = 0 \tag{7}$$

Next we define two other terms. Let $f_T(\bar{n})$ 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 $\bar{n}$. 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 the Appendix) that

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

and

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

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

$$\bar{i}_1 = \bar{i}_2 = \ldots = \bar{i}_R = \bar{n}_T p - f_T(\bar{n}) + X(\bar{n}) - g_Q(X(\bar{n})) \tag{8}$$

$$f_T(\bar{n}) = \bar{j}_R = \ldots = \bar{j}_1 \tag{9}$$

$$\bar{k}_1 = \ldots = \bar{k}_R = g_Q(X(\bar{n})) \tag{10}$$

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

$$f_T(\bar{n}) - X(\bar{n}) = 0 \tag{11}$$

$$g_Q(X(\bar{n}))\left(1 - \frac{\sigma}{p}\right) - \left[N - \bar{n}_T - \bar{n}_Q - R(\bar{n}_T p + f_T(\bar{n}))\right]\sigma = 0 \tag{12}$$

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

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

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. (5)) should be equal to $\bar{n}_Q$, the average number of stations in mode Q. Therefore, we have

$$\bar{n}_Q = -\sum_{i=1}^{V-W} \frac{1}{1 - z_i} \tag{13}$$

where z_i, $i = 1, \ldots, (V - W)$, are the roots of $U(z) - z^W = 0$ outside the unit circle $|z| = 1$. Also, since $X(\bar{n}) = f_T(\bar{n}) = \rho W$, eqs. (11) and (12) become

$$\bar{n}_T p\left(1 - \frac{p}{V}\right)^{\bar{n}_T - 1} - \rho W = 0 \tag{14}$$

$$\frac{g_Q(\rho W)}{\sigma}\left(1 - \frac{\sigma}{p}\right) = N - (1 + pR)\bar{n}_T - \bar{n}_Q - \rho R W \tag{15}$$

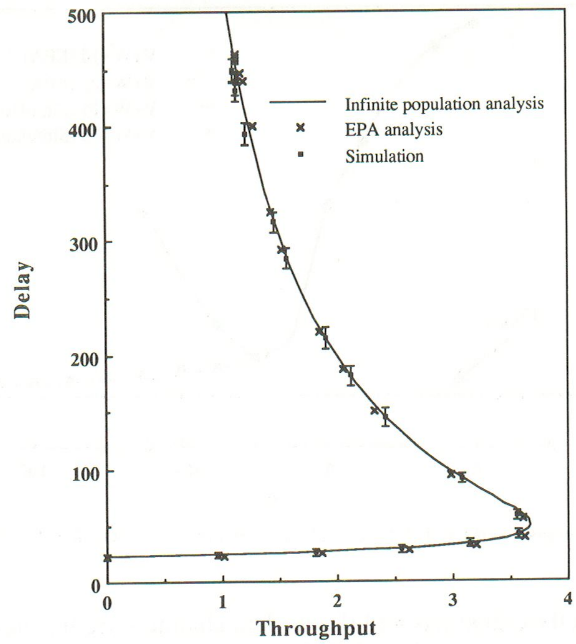

Fig. 10. Throughput versus delay curves: $N = 500$, $W = 4$, $V = 10$, $R = 10$, $p = 0.2$.

Next we compare both the infinite and finite population analyses for a system consisting of a large number of stations. Figure 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 simulation results.

Figures 11 and 12 show the effect of varying V and W while keeping their sum constant by fixing the slot size. We can 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 $V \approx eW$

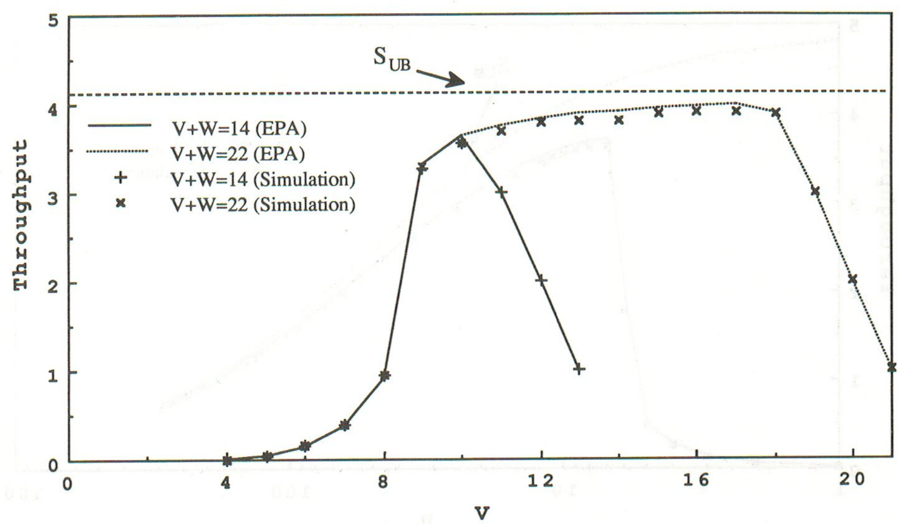

Fig. 11. Throughput versus V for fixed slot sizes ($V + W$ constant): $N = 500$, $R = 10$, $\sigma = 0.01$, $p = 0.2$.

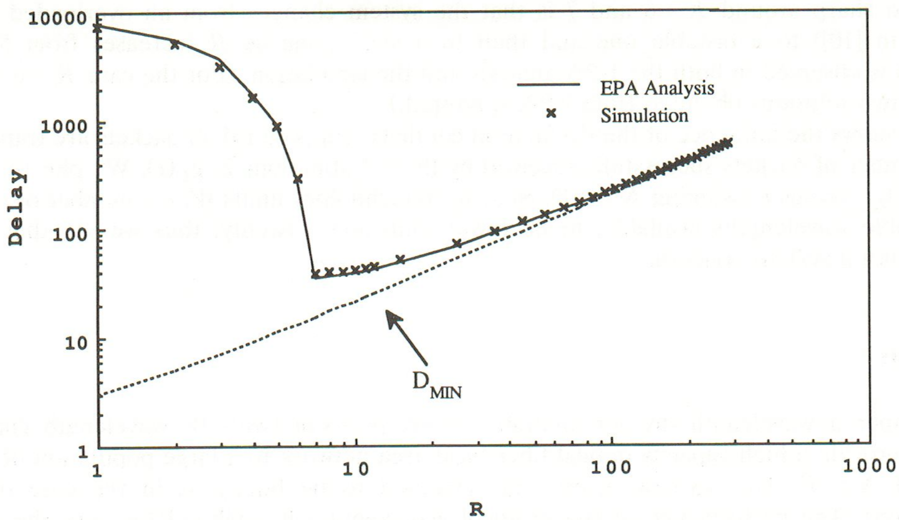

Fig. 14. Delay versus propagation delay: $N=500$, $W=4$, $V=10$, $\sigma=0.01$, $p=0.2$.

and a lower bound for the delay,

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

which are the "flat" regions in Figs. 11 and 12, respectively. Note that when $S_{\mathrm{UB}} > \min(V/e, W)$ (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 Figs. 13 and 14. Although the physical distance of the network is usually fixed, R 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

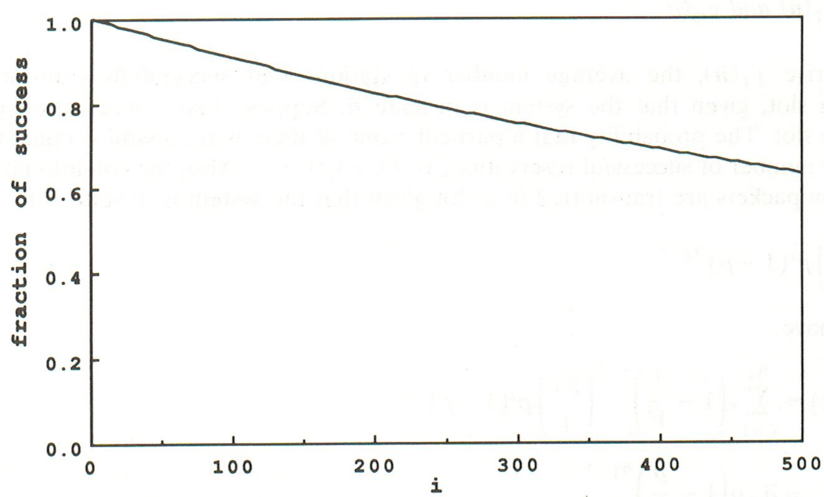

Fig. 15. The fraction of success versus the number of data packets transmitted: $N=500$.

Now 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 $(1-1/(N-1))^{i-1}$. 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 $(1-1/(N-1))^{i}$. 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

[1] C.A. Brackett, Dense wavelength division multiplexing networks: principles and applications, *IEEE J. Selected Areas Comm.* **8**(6) (1990) 948–964.

[2] M.-S. Chen, N.R. Dono and R. Ramaswami, A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks, *IEEE J. Selected Areas Comm.* **8**(6) (1990) 1048–1057.

[3] I. Chlamtac and A. Ganz, Design alternatives of asynchronous WDM star networks, *Proc. ICC'89,* pp. 23.4.1–23.4.6.

[4] I. Chlamtac and A. Ganz, Toward alternative high-speed network concepts: the SWIFT architecture, *IEEE Trans. Comm.* **38** (1990) 431–439.

[5] A. Ganz and Z. Koren, WDM passive star—protocols and performance analysis, *Proc. IEEE INFOCOM'91*, Bal Harbour, FL, April 1991, pp. 9A.2.1–9A.2.10.

[6] P.S. Henry, High-capacity lightwave local area networks, *IEEE Comm. Mag.* **27** (1989) 20–26.

[7] M.G. Hluchyj and M.J. Karol, ShuffleNet: an application of generalized perfect shuffles to multihop lightwave networks, *Proc. IEEE INFOCOM'88*, pp. 4B.4.1–4B.4.12.

[8] L. Kleinrock, *Queueing Systems, Vol. 1: Theory* (Wiley, New York, NY, 1975).

[9] L. Kleinrock, *Queueing Systems, Vol. 2: Computer Applications* (Wiley, New York, NY, 1976).

[10] L. Kleinrock and S.S. Lam, Packet switching in a multiaccess broadcast chanel: performance evaluation, *IEEE Trans. Comm.* **23** (1975) 410–423.

[11] S.S. Lam, Packet switching in a multi-access broadcast channel with applications to satellite communication in a computer network, Computer Science Department, School of Engineering and Applied Science, Engineering Report UCLA-ENG-7429, March 1974.

[12] R.M. Newman, Z.L. Budrikis and J.L. Hullet, The QPSX MAN, *IEEE Comm. Mag.* **26** (1988) 20–28.

[13] Y. Ofek and M. Sidi, Design and analysis of a hybrid access control to an optical star using WDM, *Proc. IEEE INFOCOM'91*, Bal Harbour, FL, April 1991, pp. 2A.3.1–2A.3.12.

[14] L.G. Roberts, Dynamic allocation of satellite capacity through packet reservation, *Proc. AFIPS Conference* **42** (1973) 711–716.

[15] F.E. Ross, FDDI—a tutorial, *IEEE Comm. Mag.* **24** (1986) 10–17.

[16] S. Tasaka and Y. Ishibashi, A reservation protocol for satellite packet communication—a performance analysis and stability considerations, *IEEE Trans. Comm.* **32** (1984) 920–927.

[17] F.A. Tobagi, F. Borgonovo and L. Fratta, Expressnet: a high performance integrated-services local area network, *IEEE J. Selected Areas Comm.* **1**(5) (1983) 898–913.