ELSEVIER

Computer Networks and ISDN Systems 28 (1996) 371-390

COMPUTER NETWORKS and ISDN SYSTEMS

A collision-free MAC protocol for optical star LANs

Ian F. Akyildiz *, David A. Levine

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Accepted 22 November 1994

Abstract

A new multiple access protocol is developed for optical local area networks based on a passive star topology. The protocol uses wavelength division multiplexing (WDM), and combines the advantages of both random-access and scheduled-access sharing schemes, e.g., fixed pre-assigned TDMA and WDM, into a protocol that is highly bandwidth-efficient. No central control is required among users, and the amount of processing required by each station is small. Time is divided in fixed-sized slots. Before transmitting its data, a station must compete with others for the right to use a slot in a pre-assigned wavelength, using a collision-free procedure. This results in a protocol that is suitable for networks where the number of users is larger than the number of available channels. The scheme can operate with at least a single tunable transmitter/receiver pair in each station. The paper includes an analysis of the maximum throughput and delay characteristics of the presented scheme. Several models are also developed and compared to the results obtained from numerous simulations.

Keywords: Optical LANs; Multiaccess protocols; Wavelength division multiplexing

1. Introduction

It is estimated that a single optical fiber has a capacity of at least 30 Terahertz in its low-loss region (1.2–1.6 μm) [2,12]. To put this into perspective, this capacity is close to the equivalent of all the telephone calls in the USA today at the peak busy hour of the year [12]. At the present time, however, only a fraction of this huge capacity can be used in practice. This is usually achieved by dividing the total bandwidth into a number of channels, a technique known as wavelength division multiplexing (WDM). Computer communications have benefited greatly from the introduction of optical fiber as the transmission medium. Very high transmission speeds can now be achieved, limited mainly by bottlenecks in processing and in opto-electronic interfaces. As in any communication system that uses a shared transmission medium, computer networks that use optical fiber require an arbitration protocol for an ordered access to the transmission media. Usually, the main objectives of such communication protocols are to maximize the throughput and to minimize the delay. The characteristics of these protocols are heavily influenced by the span of the network they are designed for. In the past few years, several of these protocols have been proposed for optical Local Area Networks (LANs) [1,3–6,13,15,17,18,24,26]. The designer of this type of protocols has a number of variables to consider. Some of the

* Corresponding author.

0169-7552/96/$15.00 © 1996 Elsevier Science B.V. All rights reserved
SSDI 0169-7552(94)00108-1

most important factors, and their effects on the overall design and performance of the protocol are detailed in what follows:

Medium sharing scheme. Broadly speaking, at one side of the spectrum we find reservation access schemes [13,17], while at the other end we find fixed pre-allocated assigned schemes [3,15]. In a reservation scheme, a node that is preparing a data transfer has to reserve a resource (typically, a slot in a data channel) before the actual packet transmission can take place. The reservations are usually placed in a single special channel (the control channel). Access to the control channel can be in an ordered (pre-allocated or static) or random fashion. Random access schemes [13,18] are generally the simplest to implement, but limit the throughput, delay characteristics, and bandwidth efficiency of the entire network due to collisions in the control channel. Moreover, collisions can also occur in the data channels in those schemes where the data transmission channel is determined at random. Reservation schemes are commonly used in networks where the number of nodes exceeds the number of data channels. In a pre-allocation scheme [7,12], a resource such as a channel, a slot, or a transmission time is held for a user or a group of users. Bandwidth efficiency tends to decrease as the traffic distribution becomes less homogeneous. Pre-allocation protocols are in general the best choice when the number of available channels is equal to or exceeds the number of nodes in the network.

Number and type of transmitters/receivers. At the present time and due to the costs involved, it is in general not feasible for a station in a network to have a number of transmitter/receiver pairs equal to the number of wavelength channels that the station has access to. There are fixed transmitters (FT) and fixed receivers (FR) which can only access a single channel, and tunable transmitters (TT) and tunable receivers (TR) which can access several channels at different times [2]. The number and density of wavelength channels may limit the use of TTs and TRs since these devices have tuning range limits. Also, when designing high speed protocols, it is important to consider that the tuning of a TT or TR to a particular wavelength may involve a non-negligible time. Tuning times depend not only on the distance between accessed channels, but mainly on the channel density. Obviously, all protocols require a minimum of one transmitter/receiver pair, but some require more.

Number of control and data channels. It is suggested in [2] that it may be possible to have a combined total of 1000 high speed data and control channels in a WDM-based network. However, such a large number of channels is not always desirable because of the limitations in the transmitters and receivers explained above and because more channels imply more losses for transmission. While some protocols require one control channel, others require an amount that is equal to the total number of stations in the network. Most protocols can be modified to have a different proportion of control to data channels, trading amount of processing for bandwidth efficiency.

In addition to the above variables, every protocol design should try to achieve fairness, minimize delay and processing, and maximize throughput. Other characteristics, such as scalability, and the ability to easily add and remove stations from the network are highly desirable. Also, central control schemes should be avoided. But obviously, and like in any good engineering problem, it is not possible to have everything in a given design and compromises need to be made. In the proposed protocol, stations reserve bandwidth dynamically. A minimum of one tunable transmitter/receiver pair is required per station. The scheme requires at least one control channel, and depending on the number of available data channels, it can support hundreds and up to a few thousand users. The number of users is generally much larger than the number of available channels.

In contrast to our protocol, in several other schemes the maximum number of users is limited by the number of available channels. In order to support a network of M stations, protocols like [3–6,15] require more than M wavelengths ($2M$ for [15], $M+1$ for [3–6]), imposing a serious restriction on the maximum number of users in the network, typically limiting this number to no more than a hundred stations. Any of these protocols can reduce the number of required wavelengths by using a hybrid TDM scheme, but this adds complexity and augments the amount of processing required in each node. Protocols like [13,17,18,24] present a throughput curve that diminishes after a certain offered load is exceeded, due to collisions in the data channels. Since in our scheme collisions in the data channels do not occur, the throughput is basically a monotonically increasing function of the input load, i.e., the protocol can reach a throughput equal to unity if every station in the network is fit with a number of tunable receivers that equals the number of available data channels.

The hardware architecture of the network assumed for our protocol is based on a central passive star coupler. This is the most frequently used architecture for optical LANs [1,3–6,13,15,17,18,24,26]. This can be attributed to attractive features such as simple passive broadcasting, single-stage control, data rate independence [14,16], flexibility, reliability, and optical power budget advantage over other schemes [10]. Note that our protocol can also be realized on a bus topology.

The remaining of this paper is organized as follows. In Section 2, we present the new protocol. In Section 3, we analyze the maximum achievable throughput of the protocol in terms of the number of stations, number of data channels, and number of receivers per station, assuming a Bernoulli process for the arrival of new packets to every station. Additionally, we provide analytical results for performance measures that are a function of the input and offered loads. In Section 4, we present numerical results of several simulations of the new protocol and compare them with analytical results. A summary and our concluding remarks are given in Section 5.

2. The protocol

Our protocol has its roots in the protocol described in [17]. Contrary to the protocol in [17], in our protocol collisions in the data channels cannot occur, resulting in a protocol with a bandwidth utilization and throughput that basically do not decrease as the number of stations attempting transmission increases (when the load is uniformly distributed). We assume that all stations in the network are synchronized by using a common clock, and that guard times between transmissions from different stations are essentially non-existent. Note that our protocol works but obviously at less than optimal performance without the latter assumption. The synchronization problem has been studied in [20,23]. We do not require a master station in our protocol, i.e., the medium access scheme is distributed.

2.1. Basic protocol

We make the following assumptions:

- There are $N+1$ channels ($\lambda_0, \lambda_1, ..., \lambda_N$) in the fiber (typically, the total number of channels is between 10 and 100).
- There is a single control channel λ_0.
- There are a total of M stations numbered m_1, m_2, ..., m_M in the network.
- The number of stations M is a multiple of the number of data channels N so that $Nq = M$, where q is an integer.
- Each station has at least a tunable transmitter/receiver pair.
- Transmitter and receiver tuning times, propagation delays, and control packet processing times are negligible.

The control channel is divided into equally-sized slots. Each slot consists of two parts, the reservation part and the tuning part, as shown in Fig. 1. The reservation part is divided into N minislots, and each of these minislots is divided into q microslots. This corresponds to dividing all stations into N groups, where each group consists of up to q stations. Therefore, there is a single and unique microslot for each possible station in the network, and each station belongs to a single group.

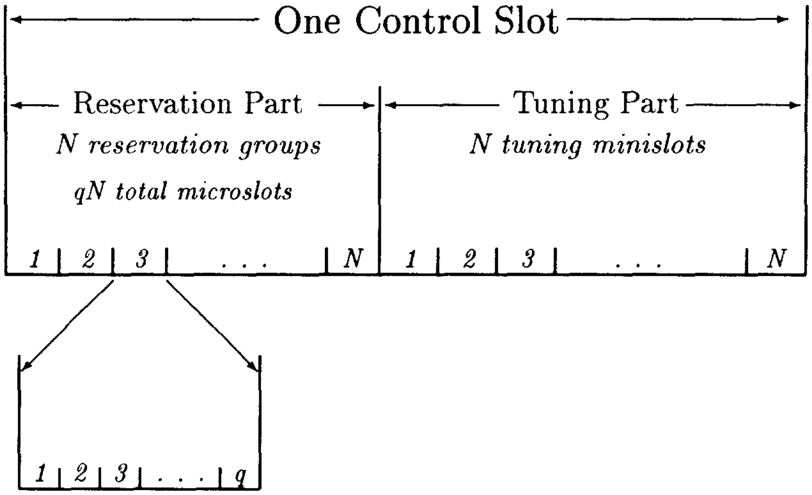

Fig. 1. Structure of a control slot.

The tuning part of the control channel is also divided into N minislots. For each of the N groups of stations there is a corresponding minislot in the tuning part, e.g., group i uses tuning minislot i, for $i = 1, 2, 3, \dots, N$. In each of the N groups, and in every control slot, up to q stations compete for the opportunity to use the tuning minislot assigned to their group. The tuning minislot is required because it is in this minislot where a station specifies its intended receiver. Apart from the intended receiver's address, the tuning minislot could also contain other information, such as the sender's address, or type of traffic. However, if the amount of information in the control slot is to be minimized, the sender's address can be included in the corresponding data packet. The sender's address can also be determined implicitly. At the expense of extra processing, each station can track the winner of each group and then wait for a winner to write that receiver's address in the tuning minislot.

Each microslot in the reservation part is one bit wide. Whenever a station needs to send data to another station, it first raises a flag in its assigned microslot, as shown in Fig. 2. The purpose of the flag is to tell the other stations in the same group that a given station needs to use the minislot in the tuning part assigned to the group. Since up to $(q-1)$ other stations may also be trying to use the tuning minislot at the same time, i.e., in the same control slot, we use the following contention scheme.

Associated with each of the N groups, there is a pointer called the "group pointer" that uniquely determines the station that can use the tuning minislot in the present control slot. At any given time, the group pointer indicates the number of the station that has the right to use the tuning minislot, provided that the station raised the flag in this control slot. If the station in question did not raise the flag, then another candidate is sought by "rotating" a copy of the group pointer until a raised flag is found. Accordingly, in Fig. 2, stations m_3, m_5, and m_{10} in groups 1, 2, and 4 respectively, will use their tuning microslot in this control slot. Note that since no station raised its flag in group 3, tuning minislot 3 will be unused during this control slot.

To prevent a condition where only the first of two stations with contiguous microslots gets a better opportunity to send a large number of data packets just because it is closer to the group pointer, each station that has been selected to use the tuning minislot has to count the number of stations C_{fr} in its group that raised their flag during the current control slot. Then, the station using the tuning minislot can raise its flag again only after at least C_{fr} control slots have passed. This mechanism improves the fairness in the selection of the next user of a tuning minislot within a given group.

Each of the stations in a network can implement the group pointer as an "ON" bit in a "pointer" register that has a size equal to q. The bit in the pointer register is logically right-shifted with each control slot, so that when the "ON" bit reaches the rightmost position in the register, the next control slot results in the leftmost bit of

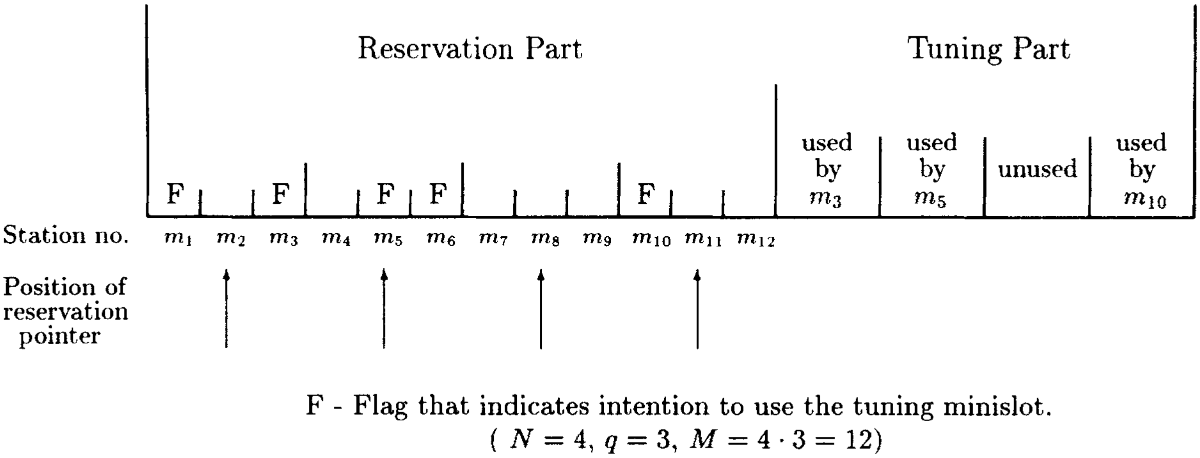

Fig. 2. An example of stations trying to seize a tuning minislot.

the register being turned "ON", while all others are "OFF". Each station is continually monitoring the control channel. If it has raised its reservation flag, it has to save in a "flag status" register (also of width q) the flags that other stations may have raised. Each station then ANDs the contents of the pointer and flag status registers. If a non-zero result is obtained, then the station that has a number equal to the position of the pointer will use the tuning minislot of the present control slot. Otherwise, a copy of the bit in the pointer register is shifted and the AND operation is performed in search of another station with the right to use the tuning minislot. This process may be repeated for up to $(q-1)$ times. Obviously, this scheme only works if all stations in a given group have a consistent value of the pointer at all times. Also, and if not equal, the offsets of the pointers of other groups, in relation to the local pointer, must be known by all stations if dynamic addition and deletion of stations is to be available (explained below).

For the basic protocol described here, the winning station of each group needs only to fill its tuning minislot with the address of the intended receiver. The wavelength to use for the transmission is implied by the tuning minislot number. For example, stations using tuning minislot 1 would send their data on wavelength λ_1. The length of the data slots is equal to the length of the control slots. The transmission of the data slots needs to occur at the end of the control slot. Since two or more stations may be sending data to a given station during the same slot, a scheme is required to resolve conflicts that could arise if the designated station has a number of receivers that is smaller than the number of expected data packets. In the simplest approach, the addressed station picks one of the sending stations at random; conversely, and following the ideas described above for the group pointers, all stations could implement a second pointer (that would be synchronized among all stations) for this purpose. Regardless of the approach, the checking of correct delivery of a data packet should be implemented in each station by higher layer functions. In other words, it should not be assumed that a successful reception by the intended receiver occurred just because access of a tuning slot was achieved followed by the transmission of a corresponding data packet.

As in [17], broadcast and multicast can easily be supported in this protocol. For either of these services, once a station has gained access to a tuning minislot, it sends a special symbol consisting of a pre-defined bit-pattern that other stations interpret as either a broadcast or multicast code. This code substitutes the usual destination address of a receiver and has higher priority than regular destination addresses. Obviously, in a broadcast situation, all data packets sent by other stations may be lost (this depends on the number of receivers available in each of the addressed stations). In addition to a broadcast situation, and discarding hardware failure and noise, there are other conditions where a data packet could be lost. For example, a station with only a single tunable receiver will miss a packet if its address was transmitted in a tuning minislot in the control channel while it was receiving data in another channel.

For a total of M possible stations, N data channels, and Y special symbols, the minimum size of the control slot (and therefore the minimum size of a data slot) is given by $M + N\lceil \log_2(M+Y) \rceil$. For example, a network with 20 data channels, a maximum of 240 stations, broadcast capabilities, and up to 15 other special symbols, would require a control slot of only 400 bits wide. Such configuration would be able to support traffic composed of unsegmented ATM cells.

3. Performance evaluation

In this section, we study several aspects that describe the behavior of the network under a variety of conditions. We begin by investigating the effects that the number of channels, total number of stations, and receivers per station have on the maximum achievable network throughput. We also study the behavior of the throughput of the network as a function of the rate of arrival of new packets and the total number of packets presented to the network. Finally, we propose a model for the delay characteristics of a network under different loading conditions.

Modeling Assumptions:

- Packets are generated in each station as independent Bernoulli processes, and are uniformly destined to any station. The probability of a new packet arriving to any station at the end of a slot is equal to σ, the ratio of the input load I to the number of stations M.
- Each station has an FT/FR pair for the control channel. For the data channels, each station has a single FT for its assigned data channel, and one or more TRs. Moreover, the number of tunable receivers per station is the same and is denoted by R.
- Each station has a large buffer size for data packets, i.e, packets cannot be lost due to buffer overflow.

Here, the input load of the network I is considered to be the average rate of data packets generated by all stations in the network. The offered load to the network G is equal to the average rate of data packets being carried by the network.

The fact that each station has independent receivers for the control and data channels allows a station to receive a data packet and to watch the control channel at the same time, reducing the possibility of a lost packet. Also, a station transmitting a data packet can simultaneously place a request for a tuning minislot in the control channel.

3.1. Maximum achievable throughput of the network

As explained in Section 2.1, if two or more stations send a packet during the same data slot to a station that has a single tunable receiver, all but one of the data packets sent to that station will be lost. Obviously, increasing the number of tunable receivers reduces the number of packets that are lost (assuming that the station has the hardware to process messages that arrive at the same time), resulting also in a higher network throughput. Here, the network throughput is considered to be the total rate of data being transmitted between all stations and normalized by the total network capacity. Note that the network throughput does not involve the bandwidth that is waisted in using a channel for control purposes.

We want to determine the value of the maximum achievable throughput S_{MAT} of any network configuration, as a function of the total number of stations M, N data channels, and R receivers per station. We assume that the network operates under heavy load, i.e., in every slot, all N data channels are being used.

Let U be a set that contains all permutations that can be formed with the destination addresses of the N data packets being sent during any slot. We can generalize and consider that U is composed of disjoint subsets of the form $E_{k_1=a,k_2=b,k_3=c,\ldots}$. The elements of these subsets are permutations that have certain characteristics. Now, consider that every possible permutation of the stations' addresses appearing in the N data channels can be divided into disjoint groups, where the contents of each group is a repeated address. Obviously, there can be at least one group with N elements (if there is only a single address repeated N times), and up to N groups with 1 element each (if all of the N addresses are different). For a given permutation, the number of disjoint groups that can be formed with the N addresses determine the subset that the permutation belongs to. The expression k_i for $i = 1, 2, \ldots$ denotes the number of groups in a permutation that have i and only i repeated addresses. Obviously, the sum of the k_i values for $i = 1, 2, \ldots$ must always satisfy $\sum_{i=1}^{\infty} ik_i = N$. As an example, if in a network with $N = 11$ data channels and during a particular slot, packets are sent to addresses $3, 2, 20, 4, 4, 2, 4, 1, 19, 3, 12$ using channels $1, 2, 3, \ldots, 11$, respectively, then the permutation formed with these addresses has four groups with one element (addresses 1, 12, 19, and 20 are not repeated), two groups with two elements (addresses 2 and 3 are repeated twice), and one group with three elements (address 4 is repeated three times). Therefore, this permutation is an element of subset $E_{k_1=4,k_2=2,k_3=1}$.

Clearly, if during a slot a particular destination is targeted by i data packets, then the number of packets that can be received by this destination is a function of the number of tunable receivers R in the station. In fact, the number of data packets with a common address that can be received is either i or R, whichever is smaller. We can apply this partial result to the analysis of any permutation of N addresses to determine the S_{MAT} of the network. To do this, we divide a given permutation of N addresses into disjoint groups that contain identical addresses, and determine the number of data packets that can be received for each group, which is the number of elements in the group or the number of tunable receivers per station, whichever is smaller. Then, we add the number of receivable data packets for every group, and denote this as the number of *maximum receptions* for all these groups. Clearly, this permutation of N addresses is only an element of a subset of the form $E_{k_1=a,k_2=b,k_3=c,\ldots}$, and every other element in this subset must have the same value of maximum receptions.

Moreover, for a subset $E_{k_1=a,k_2=b,k_3=c,\ldots}$, we can find the total number of maximum receptions of all the elements in this subset simply by finding the number of maximum receptions of any element in this group and multiplying this quantity by the number of elements in the subset.

We determine the maximum achievable throughput of the network $S_{MAT}(N, M, R)$ by finding the sum of all maximum receptions of all disjoint subsets of U of the form $E_{k_1=a,k_2=b,k_3=c,\ldots}$, and dividing this number by $|U| \cdot N$, where $|U|$ is the cardinality of U (total number of elements), and $|U| \cdot N$ is the total number of maximum receptions of all possible permutations of N addresses assuming that every station has N tunable receivers.

Therefore, in order to compute the maximum achievable throughput of the network, we need to find the number of disjoint subsets of U of the form $E_{k_1=a,k_2=b,k_3=c,\ldots}$ that can be formed, as well as the number of elements in each of these subsets. Clearly, a station that has gained access rights to a data packet cannot send a packet to itself. However, this consideration does not yield simple analytical expressions for the determination of the number of elements in each of the subsets that compose U. Fortunately, simple and accurate expressions can be obtained if we remove the above restriction and consider that any station of the network can send a packet to any other station, including itself. Obviously, with this consideration, the total number of elements in U and in every of its subsets increases, but the end result, the maximum throughput of the network, is very close to the value obtained if we consider that no station can transmit data packets to itself. No formal proof is provided here, but the maximum throughput obtained in this way appears to be a lower bound that approaches the true value as the number of stations in the network tends to infinity.

With this consideration in mind, we proceed to find the number of elements in any subset $E_{k_1=a,k_2=b,k_3=c,\ldots}$ as follows. First, we determine the total number of disjoint groups N_{dg} that can be formed within any element that belongs to a subset $E_{k_1=a,k_2=b,k_3=c,\ldots}$. This is simply

$$N_{dg}(E_{k_1=a,k_2=b,k_3=c,\ldots}) = \sum^{n} k_i \qquad (1)$$

where n is chosen so that $\sum_{i=1}^{n} ik_i = N$. We are interested in finding the total number of permutations of N addresses (where an address can be any number between 1 and M) so that Eq. (1) is satisfied. The number of possible ordered arrangements (permutations) of $N_{dg}(E_{k_1=a,k_2=b,k_3=c,...})$ different elements given that each element in the permutation may take a value between 1 and M is simply (i.e., sampling without replacement):

$$P(M, N_{dg}(E_{k_1=a,k_2=b,k_3=c,...})) = M!/(M - N_{dg}(E_{k_1=a,k_2=b,k_3=c,...}))! \tag{2}$$

Since we can also interleave addresses from different groups, we need to multiply this number by other factors. The first factor is found by taking any disjoint group that has two or more elements (groups that have a single element are excluded here) and by finding the number of distinguishable permutations of the elements of the chosen group among the total of N addresses. That is, if this group contains r_1 identical addresses, we find

$$C(N, r_1) = \binom{N}{r_1}. \tag{3}$$

The next factor is found in a similar way, but now from a total of $(N - r_1)$ addresses. As expected, and if the second group had r_2 identical addresses, the third factor is the number of distinguishable permutations of the elements of the third group among $(N - r_1 - r_2)$ addresses. The process is repeated for all remaining groups that have more than one element. Finally, to complete the expression we are seeking, we need the following consideration. If there are $k_2 = b, k_3 = c, k_4 = d, \ldots$ groups that contain $2, 3, 4, \ldots$ elements respectively, any distinguishable permutation of $E_{k_1=a,k_2=b,k_3=c,...}$ would be repeated $k_2!k_3!k_4!\ldots$ times if we just used the factors developed so far above. Therefore, we need to divide the partial result by $k_2!k_3!k_4!\ldots$ to get the correct answer. Some manipulation of the expressions defined above yields a simplified equation for the number of elements in any subset $E_{k_1=a,k_2=b,k_3=c,...}$:

$$|E_{k_1=a,k_2=b,k_3=c,...}| = \frac{N!\ M!}{\left(\prod_{i=2}^{n}(i!)^{k_i}\right)\left(N - \sum_{i=1}^{n} ik_i\right)!\left(M - N + \sum_{i=1}^{n}(i-1)k_i\right)!\left(\prod_{i=1}^{n} k_i!\right)} \tag{4}$$

where

$$\sum_{i=1}^{n} ik_i = N$$

If there is a single k_i group with i identical addresses, and all other groups have only a single address, Eq. (4) reduces to

$$|E_{k_1=N-i,k_i=1}| = \binom{N}{i}\frac{M!}{(M - N + i - 1)!}, \quad 2 \le i \le N. \tag{5}$$

The number of permutations that have N different addresses is equal to the permutation of M addresses within N positions (where $M \ge N$). That is,

$$P(M, N) = \frac{M!}{(M - N)!}. \tag{6}$$

To determine the expressions for the total number of disjoint subsets of U (subsets of the form $E_{k_1=a,k_2=b,k_3=c,...}$) as a function of N, we construct a table with the number of channels N and the corresponding number of subsets of U. We divide the table in two sections, one with even values of N, the other with the odd values. We find the first and second differences of the number of subsets in each row of N, and note that the second differences yield a constant value equal to 2. Denoting the number of disjoint subsets of U by $Y_U(N)$, we simply need to solve the difference equation $\Delta^2 Y_U(N) = 2$. Then

$$Y_U(N) = \begin{cases} 13/6(N/2) - 1/2(N/2)^2 + 1/3(N/2)^3, & N = 2,4,6,\ldots \\ -1 + 8/3(2N-1) - (2N-1)^2 + 1/3(2N-1)^3, & N = 1,3,5,\ldots \end{cases} \quad (7)$$

With these expressions, we are now capable of determining the total number of maximum receptions of all elements in U. For an element of a subset $E_{k_1=a,k_2=b,k_3=c,\ldots}$, the number of maximum receptions MR is given by

$$MR = \sum_{i=1}^{n} k_i \cdot \rho_i$$

where

$$\rho_i = \begin{cases} i, & \text{if } R \geq i, \\ R, & \text{otherwise} \end{cases}.$$

Finally, the total number of maximum receptions TMR for all elements in U is given by

$$TMR = \sum_{\text{disjoint subsets of } U} \left(|E_{k_1=a,k_2=b,k_3=c,\ldots}| \sum_{i=1,2,3,\ldots}^{n} \rho_i \cdot k_i \right). \quad (8)$$

Figure 3 shows the maximum achievable throughput for a network with $N = 30$ channels computed for different numbers of stations, and for a number of tunable receivers per station between one and four. From Fig. 3, it is clear that increasing the total number of stations has a favorable effect on the achievable maximum throughput of the network, since the overall probability of directing multiple data packets to a single station during the same control slot diminishes. Also in Fig. 3, increasing the number of tunable receivers to 2 guarantees a maximum achievable throughput of at least 0.9, even for only a total of thirty stations in the network. Similar conclusions can be drawn for networks with a different number of channels. Two tunable receivers per station result in maximum throughputs that are near or above 0.9. Also, for a constant overall input load, increasing the number of stations improves the maximum achievable throughput.

3.2. Throughput and load characteristics

Input load versus throughput. Since no collisions can occur in the data channels, the network throughput usually does not decrease as the input or offered loads increase. In fact, as long as the normalized input load I/N does not exceed the maximum achievable throughput of the network S_{MAT}, the throughput is almost directly proportional to the normalized input load. When the normalized input load is smaller than S_{MAT}, all packets being sent are eventually successful in being received by their destinations. When the normalized input load reaches the value of S_{MAT}, the network becomes saturated, and the throughput curve reaches a maximum. For subsequent increases of the input load, the throughput characteristic becomes constant at a value equal to S_{MAT}. Thus, the network throughput can be given by

$$S = \begin{cases} I/N, & \text{if } I \leq N \cdot S_{MAT}, \\ S_{MAT}, & \text{otherwise.} \end{cases} \quad (9)$$

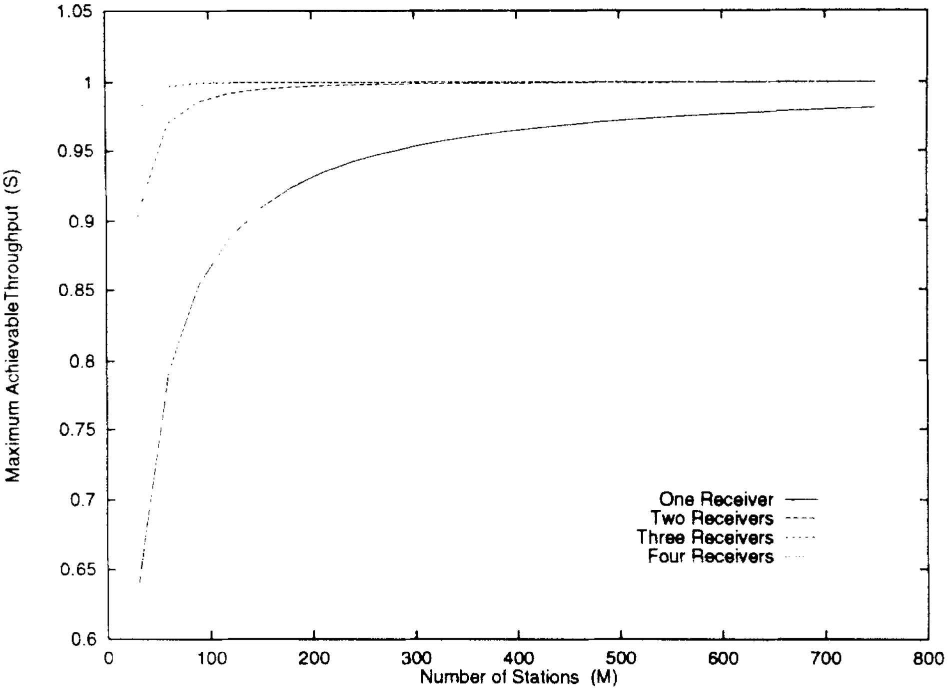

Fig. 3. Maximum achievable throughput of the network with $N = 30$.

Input loads that result in throughputs near or equal to S_{MAT} can only be sustained for a short time without packet losses. This is because at these values the rate of data packet generation is close to or exceeds the maximum rate of data packet reception. This, in turn, results in the accumulation of incoming packets at the stations' buffers, resulting also in the dropping of packets when the buffers become full. Needless to say, the normalized input load should never exceed the maximum achievable throughput if a stable operation of the network is to be achieved.

Offered load versus channel efficiency. The channel efficiency of the network $\eta_{network}$ is the ratio of the number of data channels that carry packets that are received successfully to the total number of data channels that are carrying data. If during extremely low load conditions, only one of the N channels in the network is carrying a data packet, then the network is assured to be operating with a channel efficiency equal to unity since every transmitted packet will be received by the targeted station. As the offered load increases, two or more channels are likely to be carrying data packets during any slot. If the stations have at most one tunable receiver, the channel utilization of the network will become less than unity since now it is possible for a station to be targeted by two or more data packets during a given slot. The worst-case scenario for the channel utilization occurs when the offered load is maximum, that is, $|G| = |N|$. At this point, the channel efficiency is equal to the maximum achievable throughput of the network. From these observations, we can conclude that the channel efficiency of the network: (1) is a function of the offered load, (2) it starts with a value of unity for light loads, and (3) it decreases until it reaches a value equal to S_{MAT} for the peak offered load. We approximate this relation with a straight line, yielding the following expression that relates η_{NW} and S_{MAT}:

$$\eta_{NW}(G) = \frac{S_{MAT} - 1}{N} G + 1. \tag{10}$$

Input load versus offered load. Since in most networks not all stations have a number of receivers equal to the number of available data channels, eventually some data packets will not be delivered successfully during a given slot, i.e., those that could not be received by a station that is receiving data packets in other channels, resulting in different values for the input and offered loads. As a consequence of this, for values of the input load smaller than the total capacity of the network, the offered load is always greater than or equal to the input load, with equality occurring when every station has a receiver for every channel in the network, assuming that no data packet can be lost due to hardware failure or noise. During every slot, the number of packets that are sent successfully is proportional to the channel efficiency value for that offered load. The packets that were not delivered successfully will be retransmitted at a future time, following the arbitration scheme described earlier. During any slot, additional packets will be generated at the stations (with constant rate I). With these observations in mind, we proceed to approximate this situation with the following difference equation:

$$G_{k+j} = I + G_k(1 - \eta_{NW}(G_k)) \tag{11}$$

where k denotes the difference interval, G_k denotes the value of the offered load on the k slot, and j is the average number of slots required for a retransmission ($1 < j \le q$). Substituting Eq. (10) in Eq. (11), and after some simplifications, we obtain

$$G_{k+j} = I + G_k\left(\frac{1 - S_{MAT}}{N}G_k\right). \tag{12}$$

We are only interested in finding the steady state solution of the equation, that is, the value of $G_{k+j\to\infty}$. Since only G depends on the value of the step, we can take the limit of Eq. (12) as $k \approx k + j \to \infty$, and solve the resulting equation as a simple second-order equation. Note that the value of G_0 is irrelevant, and Eq. (12) always converges to the same value. Note also that the solution of Eq. (12) can yield values of $G_{k+j\to\infty}$ greater than the maximum capacity of the network N for sufficiently large values of I. Therefore, we must restrict the value of $G_{k+j\to\infty}$ to the maximum network capacity. Excluding the root that yields values greater than the channel capacity, the resulting final expression is

$$G = \lim_{k+j\to\infty} G_k = \begin{cases} \dfrac{1 - \sqrt{1 - 4I(\frac{1-\eta_{NW}}{N})}}{2(\frac{1-\eta_{NW}}{N})}, & \text{if } \dfrac{1 - \sqrt{1 - 4I(\frac{1-\eta_{NW}}{N})}}{2(\frac{1-\eta_{NW}}{N})} \le N, \\ N, & \text{otherwise.} \end{cases} \tag{13}$$

Offered load versus throughput. Since for a range of values of the input load I network throughput and normalized input load are in general the same, and since offered and input loads are also relatively similar, it is not surprising to find that the normalized offered load and the network throughput are alike for the values that G can take. Using Eqs. (9), (12) and (13), we get

$$S(G) = \frac{G\left(1 - G\left(\frac{1-S_{MAT}}{N}\right)\right)}{N}. \tag{14}$$

Equation (14) is composed of a linear and a quadratic term. These terms reflect the fact that for small values of G, normalized offered load and throughput are almost equal, but then, as the offered load approaches the network limit, $G/N \to 1$ and $S \to S_{MAT}$.

3.3. Delay characteristics

As explained in Section 2.1, a station must follow the proposed protocol before it can send a data packet to any other station. A station can generate a data packet during any slot with probability $\sigma = I/M$. For every

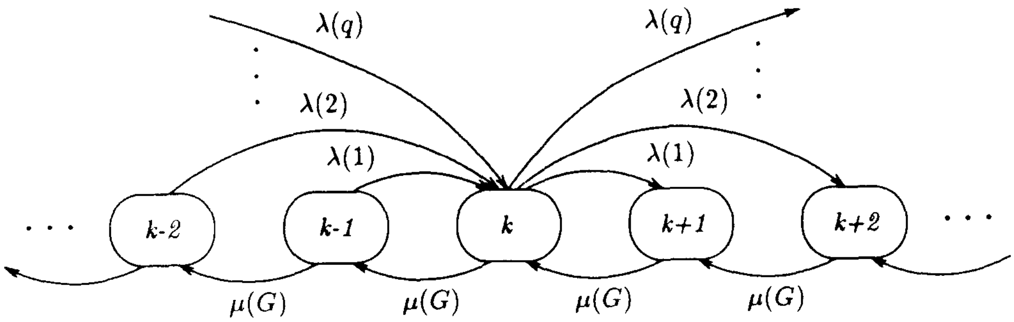

Fig. 4. The bulk arrival state-transition-rate diagram.[1]

data channel in the network, there are up to q stations that may compete during any slot for the right to use the data channel assigned to the group. Obviously, any station that was not granted transmission rights during a given slot must wait for one or more additional slots for the opportunity to use the data channel assigned to its group. Moreover, additional packets may arrive to a station that is awaiting transmission. Clearly, this situation constitutes a queueing system with a round-robin queueing discipline. Since it is assumed that all groups of stations have the same number of stations q, with the same input load I/N per group, all groups in the network behave identically.

For modeling purposes, we approximate the queueing system that is formed in every group of stations with a Markovian system with bulk arrivals. Figure 4 presents the bulk arrival state-transition-rate diagram for this model, showing the transitions for a state E_k. In this model, each state represents the total number of packets (in all stations within the group) waiting to be transmitted. Since each station within a group can generate up to 1 data packet per slot, we set the maximum number of bulk arrivals equal to q. Clearly, the number of new arrivals to the group in each slot follows a binomial distribution, so we calculate the transition probability from state E_k to state E_{k+i} $(i \geq 0)$ as

$$\lambda(\sigma, q, i) = \binom{q}{i} \sigma^i (1 - \sigma)^{q-i} \tag{15}$$

For a given network configuration and a given input load, the parameters σ and q are fixed, so we express $\lambda(\sigma, q, i)$ simply as $\lambda(i)$. The transition from a given state to a lower state can only be to the nearest neighbor; this is because at most 1 packet from a station in the group can successfully be transmitted during any slot. Note that not all attempts for transmission become successful, so clearly the departure rate μ is not equal to 1. It is clear that the departure rate is equal to the channel efficiency of the network, and thus depends on the value of the offered load. That is

$$\mu(G) = \eta_{NW}(G) = \frac{S_{MAT} - 1}{N} G + 1. \tag{16}$$

If we define p_k to be the equilibrium probability for having a total of k data packets within a given group, the equilibrium equations that describe the bulk arrival model are

$$(1 - \lambda(0)) p_0 = \mu(G) p_1, \tag{17}$$

[1] For clarity, only details for state E_k are shown.

$$\begin{aligned} (1-\lambda(0)+\mu(G))p_k &= \mu(G)p_{k+1}+\sum_{i=0}^{k-1}p_i\lambda(I,k-i), \quad \text{if } k\le q,\\ (1-\lambda(0)+\mu(G))p_k &= \mu(G)p_{k+1}+\sum_{i=k-q}^{k-1}p_i\lambda(I,k-i), \quad \text{if } k\ge q, \end{aligned} \tag{18}$$

$$p_0=\frac{1}{1+\sum_{k=1}^{\infty}p_k} \tag{19}$$

where Eq. (17) is the single boundary equation for the state E_0.

Conventional methods such as the z-transform do not give a close form solution for the above equations. Nevertheless, the values of $p_1, p_2, p_3, \ldots$ can be approximated by solving the equations numerically, considering a large number of possible states. For the solution to converge, the input load has to be kept below the maximum load that the network can handle, i.e., $I < M \cdot S_{MAT}$. Obviously, when the solution converges, as $k \to \infty$, $p_k \to 0$. Once the values for $p_1, p_2, p_3, \ldots$ are found, it is simple to find the average number of customers in the system:

$$\overline{N}=\sum_{k=1}^{\infty}kp_k. \tag{20}$$

The average time spent in the system, i.e., the average delay incurred by each packet before transmission, can be calculated using the well-known Little's Law:

$$D=\frac{\overline{N}}{\overline{\lambda}}=\frac{\sum_{k=1}^{\infty}kp_k}{q\sigma} \tag{21}$$

where $\overline{\lambda}=q\sigma$ is the average arrival rate of customers to the queue.

Using the expressions obtained above, it will be shown in the next section that values for the normalized input load less than two thirds of S_{MAT} typically result in packet delay values of 2 slots or less.

4. Numerical results

Here, we present numerous results obtained from simulations and from the application of the models developed in the previous section. The results are presented in a sequence of plots that show existing relationships between the input load, offered load, network throughput, channel efficiency, and packet delay, and compare the curves obtained through simulation and through the applications of the developed models. We simulated a network with 30 data channels ($N = 30$), a varying number of stations per group q, and different conditions for the input load I. In addition, each station had an FT/FR pair for the control channel, an FR for its assigned data channel, and a single TT for transmission to any data channel. Most plots show the results for networks with $q = 2$, 5, and 20 stations per group. While results obtained with a fixed number of channels in the network may appear to be too specific, this is not necessarily the case. We ran simulations with a different number of channels and verified that the results are proportionally similar to each other.

For the simulation program, we considered that any transmitting station could determine by the end of a control slot if the data packet about to be sent could be received by the intended destination. Also, if the intended receiver was going to be busy listening to another channel, the transmitting station could contend for its corresponding tuning minislot during the next control slot.

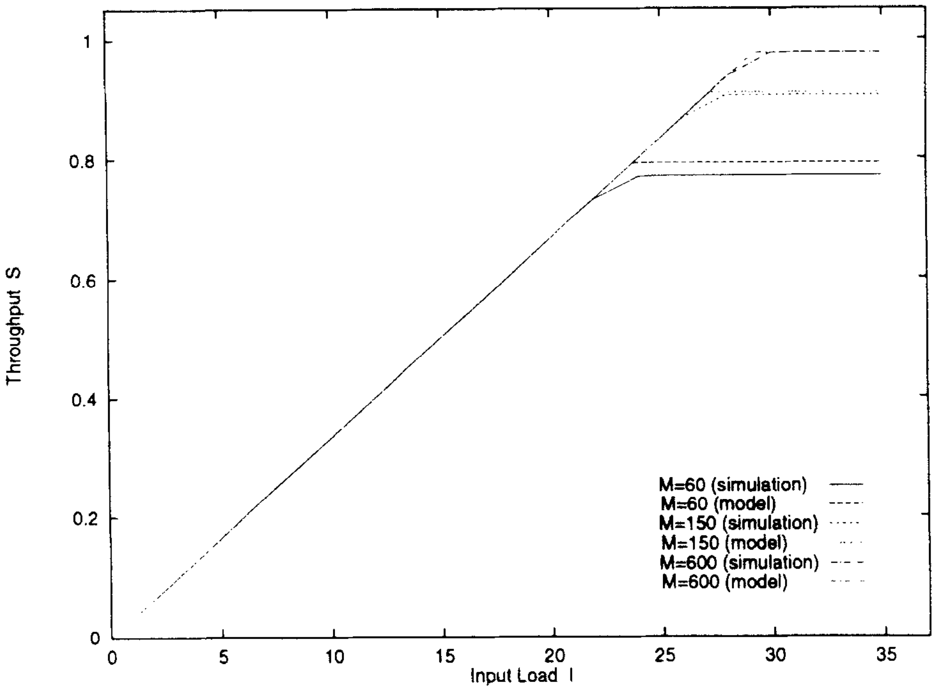

Fig. 5. Throughput S versus input load I, with $N = 30$.

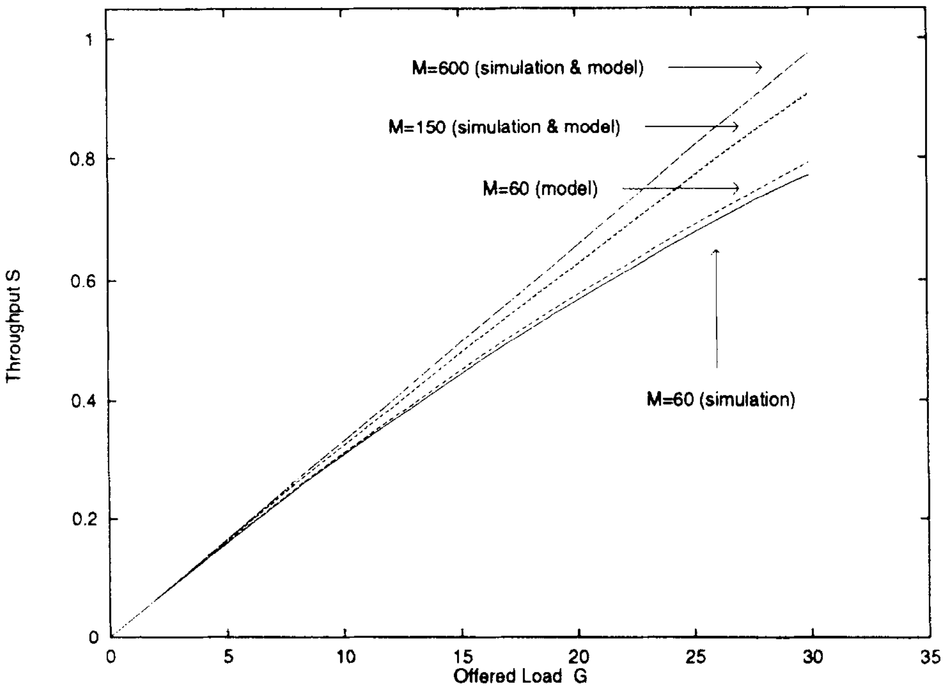

Fig. 6. Throughput S versus offered load G, with $N = 30$.

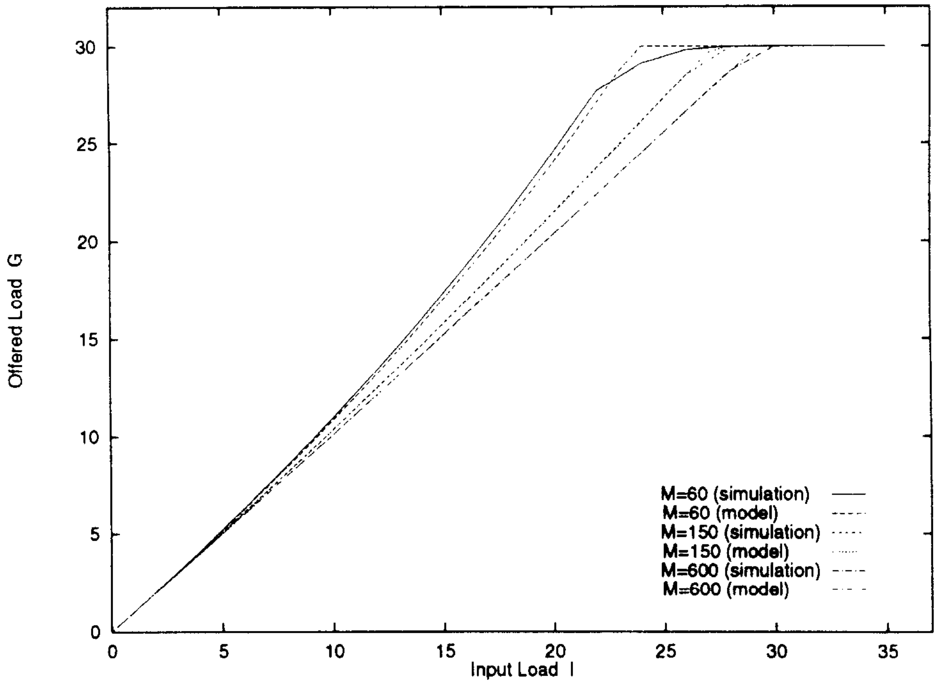

Fig. 7. Offered load G versus input load I, with $N = 30$.

Figure 5 presents the throughput versus input load characteristics for different network configurations with a total of $M = 60, 150, 600$ stations. From Fig. 5, it is evident that for small input loads the throughput is equal to the normalized input load, i.e., the packet generation rate is equal to the rate that packets are delivered successfully. When the normalized input load is greater than the S_{MAT} for the network configuration, the throughput remains at this value, the maximum network capacity is reached, and packets start to accumulate in the stations at a higher rate than the rate at which they can be delivered successfully. The analytical model that describes this relation yields a very accurate approximation, especially as the number of stations in the network increases. This is because the analytical model is based on the S_{MAT} for the network configuration, and as explained in Section 3.1, the calculated S_{MAT} is an upper bound of the real maximum achievable throughput that tends to the real value as the number of stations tends to infinity. From Fig. 5, it can be observed that a network with $M = 600$ stations can reach a throughput that is very close to unity.

The throughput versus offered load curves are presented in Fig. 6. It is obvious that, in contrast to ALOHA and CSMA-based protocols [13,17,18], the throughput of our protocol does not decrease as the offered load increases. Similar to the throughput vs. input load characteristic, increasing the number of stations in the network while maintaining the offered load has a positive effect on network throughput. It is worth to point out that in a network with N TRs per station, and regardless of the total number of stations M in the network, the values for throughput, normalized input, and normalized offered loads are identical for $0 \leq I \leq N$, and the network is able to achieve a throughput equal to unity.

For a given input load, a network with a large number of stations has more successful transmissions than a network with a smaller number of stations, requiring on the average less retransmissions per packet and thus resulting in smaller values for the offered load as shown in Fig. 7. For a network with $M = 600$ stations, the number of packets that need to be retransmitted are minimum, and the offered load versus input load curve is practically a straight line with slope equal to unity.

Figure 8 shows the channel efficiency versus offered load characteristic for the network. As explained in Section 3.2, for small values of offered load, η_{NW} is equal or close to unity, and decreases gradually until it

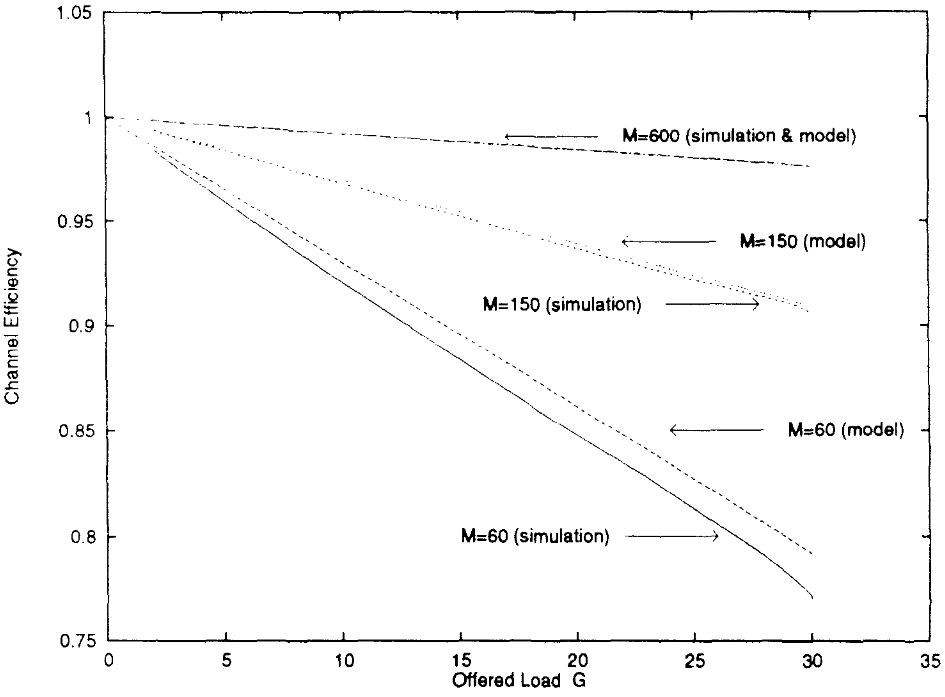

Fig. 8. Channel efficiency η_{NW} versus offered load, with $N = 30$.

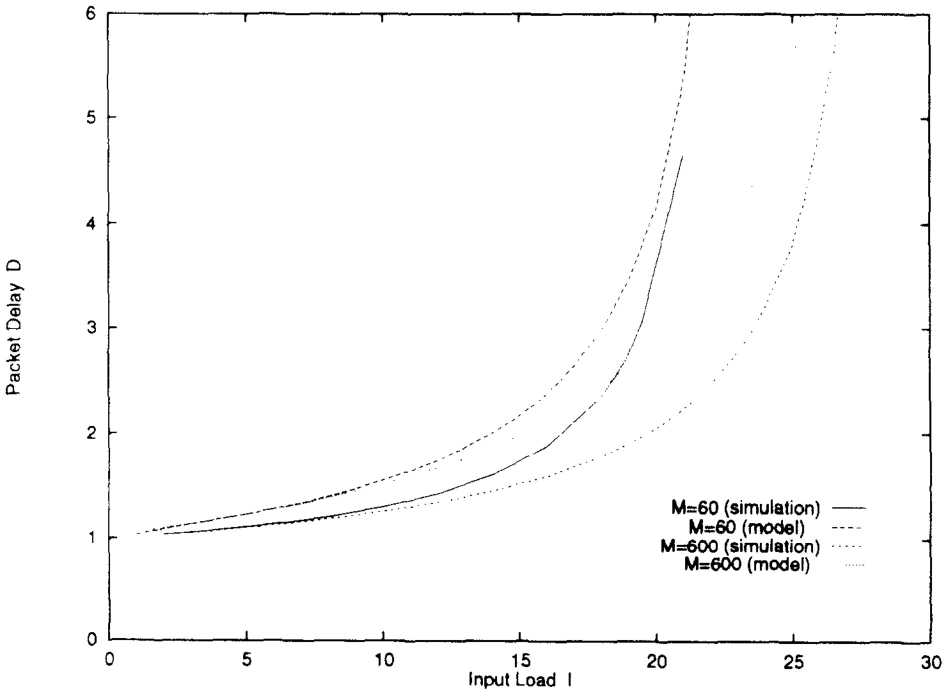

Fig. 9. Packet delay D versus input load I, with $N = 30$.

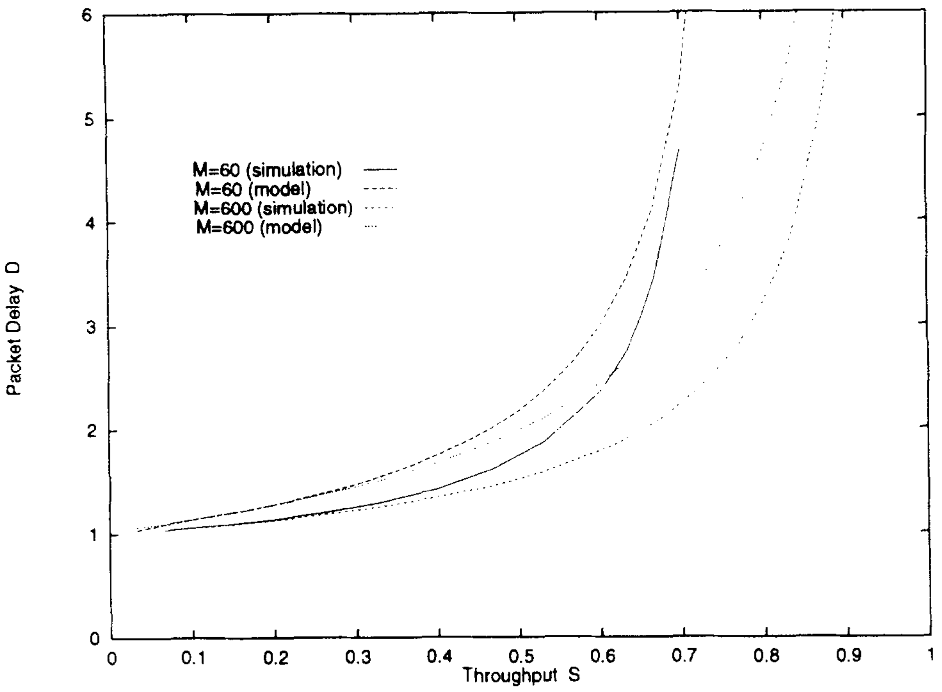

Fig. 10. Packet delay D versus throughput S, with $N = 30$.

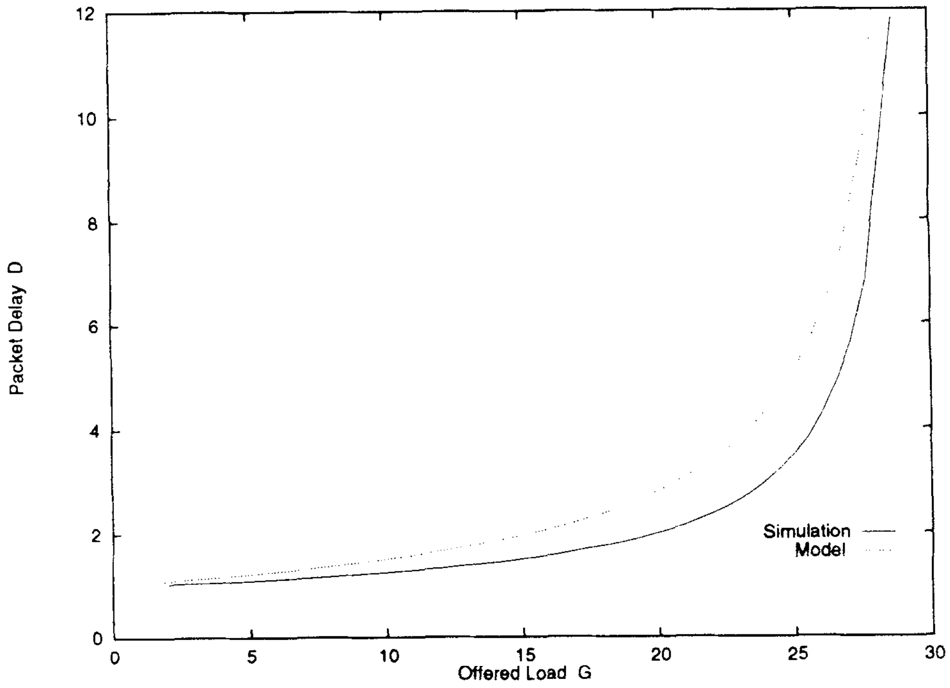

Fig. 11. Packet delay D versus offered load G, with $N = 30$.

reaches the S_{MAT} value as the offered load reaches its maximum value N packets/slot. Recall that the model developed for this characteristic assumes that channel efficiency and offered load are related according to a straight line equation. From Fig. 7, it is obvious that this is a valid assumption.

Delay versus input load is plotted in Fig. 9. It is interesting to observe that values of the normalized input load less than two thirds of the maximum achievable throughput result in packet delays of two slots or less! Obviously, higher values of the input load place the corresponding delay values in the steep portions of the curve. All delay versus throughput curves show an asymptotic behavior, with the asymptote located at $I = N \cdot S_{MAT}$. The asymptote and the curves shift to the right as more receivers per station are added. When every station in the network has N tunable receivers, the asymptote is located at $I = N$. In this case, every station is guaranteed a minimum bandwidth of $1/q$ of a channel, and a maximum delay of q slots before it can use a data channel.

Since throughput and normalized input load are equal for a range of values of the latter parameter, it is not surprising to find that the curves displayed in Fig. 10 for throughput versus delay are proportionally identical to those displayed in Fig. 9. In this case, the location of the asymptotes is simply $S = S_{MAT}$.

The number of stations in the network appears not to have an effect on the packet delay versus offered load characteristic. This was concluded after plotting this characteristic using several values for the number of stations in a group q, and by observing that the obtained curves were practically identical. Accordingly, Fig. 11 presents the general packet delay versus offered load characteristic for a network with $N = 30$.

5. Summary and conclusions

In this paper we developed a new collision-free media access protocol for optical networks where the number of users can surpass the number of available data channels. The proposed protocol is totally distributed and allows small data slots. The usage of network bandwidth is distributed among stations within a group in a round-robin-like fashion. Every station in the network is guaranteed a transmission opportunity in at most every q slots. The proposed protocol requires moderate processing, and stations only need to monitor certain sections of a control slot. We showed that the common assumption that "the number of packets lost due to simultaneous transmission to a single station is negligible" is not always valid, and can have a significant effect on the throughput of those networks with a small number of stations. In a real network with specialized servers, it is evident that to achieve high throughput, the servers will be required to have multiple receivers. Moreover, with enough receivers, a network using the proposed protocol can achieve a throughput equal to unity (in the data channels). This is in sharp contrast to other protocols that actually have their throughput reduced as the offered load is increased. The protocol presented here is fair to every station within a group in that on the average, every station among competing nodes gets the same opportunity for access to the network.

In this work, we considered that propagation delays were negligible. Fortunately, the protocol can easily be extended to consider non-negligible propagation delays. In this case, stations may send numerous requests using different control slots before they can verify if their first request was successful. This occurs when the corresponding control packet returns after a round-trip propagation delay through the network. Preliminary work shows that most performance parameters (with the exception of packet delay times) are basically unaffected by propagation delays. The average packet delay is increased by at least two times the round-trip propagation delay time (one for the reservation, the other for the data transmission). We pursue this in detail in a future paper.

References

[1] K. Bogineni, K.M. Sivalingam and P.W. Dowd, Low-Complexity Multiple Access Protocols for Wavelength-Division Multiplexed Photonic Networks, *IEEE Journal of Selected Areas in Comm.*, Vol. 11, No. 4, pp. 590–604, May 1993.

[2] C.A. Brackett, Dense Wavelength Division Multiplexing Networks: Principles and Applications, *IEEE Journal of Selected Areas in Comm.*, Vol. 8, No. 6, pp. 948-964, August 1990.
[3] M-S Chen, N.R. Dono and R. Ramaswami, A Media-Access Protocol for Packet-Switched Wavelength Division Multiaccess Metropolitan Area Networks, *IEEE Journal of Selected Areas in Comm.*, Vol. 8, No. 6, pp. 1048-1057, August 1990.
[4] M. Chen and T-S. Yum, A Conflict-Free Protocol for Optical WDMA Networks, *IEEE GLOBECOM'91*, 1991, pp. 1276-1281.
[5] R. Chipalkatti, Z. Zhang and A.S. Acampora, High Speed Communication Protocols for Optical Star Coupler Using WDM, *IEEE INFOCOM'92*, 1992, pp. 2124-2133.
[6] R. Chipalkatti, Z. Zhang and A.S. Acampora, Protocols for Optical Star-Coupler Network Using WDM: Performance and Complexity Study, *IEEE Journal of Selected Areas in Comm.*, Vol. 11, No. 4, pp. 579-589, May 1993.
[7] I. Chlamtac and A. Ganz, Design Alternatives of Asynchronous WDM Star Networks, *IEEE ICC'89*, June 1989, Vol. 2, pp. 739-744.
[8] I. Chlamtac and A. Fumagalli, A Solution to Packet Switching in Optical Transmission Networks, *Computer Networks and ISDN Systems*, Vol. 26, Nos. 6-8, March 1994.
[9] I. Chlamtac and A. Fumagalli, QUADRO-Star: A High Performance Optical WDM Star Network, *IEEE Transactions on Communications*, Vol. 42, No. 8, August 1994.
[10] P.W. Dowd and K.K. Bogineni, Switching Latency Overlap Techniques for WDM Star-Coupled Media Access Protocols, *IEEE INFOCOM'93*, April 1993, Vol. 1, pp. 65-74.
[11] S. Goldberg, *Introduction to Difference Equations*, Wiley, New York (1958).
[12] P.E. Green, *Fiber Optic Networks*, Prentice-Hall, Englewood Cliffs, N.J. (1992).
[13] I.M.I. Habbab, M. Kavehrad and C-E.W. Sundberg, Protocols for Very High Speed Optical Fiber Local Area Networks using Passive Star Topology, *IEEE Journal of Lightwave Technology*, Vol. LT-5, No. 12, pp. 1782-1794, 1987.
[14] P.S. Henry, High-Capacity Lightwave Local Area Networks, *IEEE Comm. Mag.*, Vol. 27, No. 10, pp. 20-26, 1989.
[15] P.A. Humblet, R. Ramaswami and K.N. Sivarajan, An Efficient Communication Protocol for High-Speed Packet-Switched Multichannel Networks, *IEEE Journal of Selected Areas in Comm.*, Vol. 11, No. 4, pp. 568-578, May 1993.
[16] H. Kobrinski and K-W. Cheung, Wavelength-Tunable Optical Filters: Applications and Technologies, *IEEE Comm. Mag.*, Vol. 27, No. 10, pp. 53-63, 1989.
[17] 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, Nos. 1-3, pp. 223-239, 1992.
[18] N. Mehravari, Performance and Protocol Improvements for a Very High Speed Optical Fiber Local Area Networks Using a Passive Star Topology, *IEEE Journal of Lightwave Technology*, Vol. 8, No. 4, pp. 520-530, April 1990.
[19] B. Mukherjee, WDM-Based Local Lightwave Networks Part I: Single-Hop Systems, *IEEE Network*, pp. 12-27, May 1992.
[20] Y. Ofek and M. Sidi, Design and Analysis of a Hybrid Access Control to an Optical Star Using WDM, *IEEE INFOCOM '91*, Vol. 1, pp. 20-31, April 1991.
[21] P.R. Prucnal, M.A. Santoro and S.K. Sehgal, Ultrafast All-Optical Synchronous Multiple Access Fiber Networks, *IEEE Lightwave Telecomm. Systems*, Vol. 3, No. 2, pp. 16-26, May 1992. *IEEE Journal of Selected Areas in Comm.*, Vol. SAC-4, No. 9, pp. 1484-1492, December 1986.
[22] C. Rolland, L.E. Tarof and A. Somani, Multigigabit Networks: The Challenge, *IEEE Lightwave Telecomm. Systems*, Vol. 3, No. 2, pp. 16-26, May 1992.
[23] G. Semaan and P. Humblet, Timing and Dispersion in WDM Optical Star Networks, *IEEE INFOCOM '93*, April 1993, Vol. 2, pp. 573-577.
[24] G.N.M. Sudhakar, N.D. Georganas and M. Kavehrad, Slotted Aloha and Reservation Aloha Protocols for Very High-Speed Optical Local Area Networks Using Passive Star Topology, *Journal of Lightwave Technology*, Vol. 9, No. 10, pp. 1411-1422, 1991.
[25] H. Toba, K. Oda, K. Nose and N. Takato, Factors Affecting the Design of Optical FDM Information Distribution Systems, *IEEE Journal of Selected Areas in Comm.*, Vol. 8, No. 6, pp. 965-971, August 1990.
[26] K.A. Williams, T.Q. Dam and D.H.-C. Du A Media-Access Protocols for Time- and Wavelength-Division Multiplexed Passive Star Networks, *IEEE Journal of Selected Areas in Comm.*, Vol. 11, No. 4, pp. 560-567, May 1993.

Ian F. Akyildiz received his B.S., M.S., and Ph.D. degrees in Computer Engineering from the University of Erlangen-Nuernberg, Germany, in 1978, 1981 and 1984, respectively. Currently, he is a Full Professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology. He has held visiting professorships at the Universidad Tecnica Federico Santa Maria, Chile, Université Pierre et Marie Curie (Paris VI) and Ecole Nationale Supérieure Télécommunications in Paris, France. He has published over hundredfifty technical papers in journals and conference proceedings. He is a co-author of a textbook entitled "Analysis of Computer Systems" published by Teubner Verlag in Germany in 1982. He is an editor for "Computer Networks and ISDN Systems", for "IEEE Transactions on Computers", and for "ACM-Baltzer Journal of Wireless Networks". He guest-edited several special issues, such as on "Parallel and Distributed Simulation Performance" for "ACM Transactions on Modeling and Simulation"; on "Teletraffic Issues in ATM Networks" for "Computer Networks and ISDN Systems", and on "Networks in the Metropolitan Area" for "IEEE Journal of Selected Areas in Communications".
Dr. Akyildiz is a senior member of IEEE since 1989, member of ACM (SIGCOMM) and is a National Lecturer for ACM since 1989. He received the "Don Federico Santa Maria Medal" for his services to the Universidad of Federico Santa Maria in Chile. Dr. Akyildiz is listed on "Who's Who in the World (Platinum Edition)". He received the ACM Best Distinguished Lecturer Award for 1994.
His current research interests are in B-ISDN (ATM) networks, wireless networks, optical local area networks, performance evaluation, parallel simulation, high performance computing and network security. His E-mail address is: ian@armani.gatech.edu.

David A. Levine received the B.S. in Electronics and Communications in 1988 from the Universidad La Salle, Mexico City, and the M.S.E.E. degree in 1991 from the University of Virginia. He is currently a Ph.D. student in Computer Engineering at the Georgia Institute of Technology, where he is the recipient of a scholarship from the National Science Foundation.
His current research interests are in wireless B-ISDN (ATM) networks, Optical LANs, and performance evaluation. His email address is: dlevine@nemesis.mirc.gatech.edu.