

Minimizing Delay and Packet Loss in Single-Hop Lightwave WDM Networks Using TDM Schedules

George N. Rouskas
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206

Mostafa H. Ammar
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280

Abstract

We consider single-hop WDM networks with stations equipped with tunable transmitters and fixed receivers. Access to the channels is controlled by a weighted TDM scheme. We study the problem of designing TDM frames to minimize the mean packet delay, as well as the mean packet loss probability given a finite buffer capacity. We develop optimization heuristics which, for non-uniform communication patterns common to parallel and distributed computations, represent a significant improvement over I-TDMA*.

1 Introduction

Wave division multiplexing (WDM) is the most promising technology for implementing a new generation of computer communication networks that fully exploit the vast information-carrying capacity of single-mode fiber [4]. By carving the bandwidth of the optical medium into multiple concurrent channels, WDM has the potential of delivering an aggregate throughput that can be in the order of Terabits per second. Single-hop WDM systems [8] provide one hop communication between any source-destination pair, by allowing the various stations to select any of the available channels for packet transmission/reception. Access to the channels can be based on a reservation scheme that requires the use of one [2, 3] or more [6] control channels. Alternatively, a hybrid WDM-TDM approach may be employed. In this case, the optical bandwidth is preallocated by means of a transmission schedule that indicates the slots in which the various transceivers may access the available channels [9, 1].

This work explores the delay and packet loss probability behavior of single-hop networks employing TDM schedules under potentially non-uniform traffic patterns. Because of the complexity of the resulting optimization problems, a large part of our work is devoted to developing heuristics that (a) construct schedules to achieve near-optimal mean queue length, and (b) allocate the buffer capacity so as to attain an acceptable packet loss probability. As a result, substantial gains in performance over I-TDMA* [1] are achieved.

Section 2 presents our system model. Heuristics to minimize the mean packet delay and packet loss probability are developed in Sections 3 and 4, respectively. We present numerical results in Section 5, and a summary in Section 6.

2 System Model

We consider a network of N stations, each equipped with one receiver and one transmitter, interconnected through an optical broadcast medium that can support C wavelengths, $\lambda_1, \lambda_2, \dots, \lambda_C$, $C \leq N$. The stations employ tunable transmitters and/or receivers. For simplicity, we only consider systems with tunable transmitters and fixed receivers; our results can be easily extended to tunable-receiver systems.

The fixed receiver at station i is assigned wavelength $\lambda(i) \in \{\lambda_1, \dots, \lambda_C\}$. The transmitters, on the other hand, are lasers that can be tuned to, and transmit on all wavelengths $\lambda_c, c = 1, \dots, C$. The network operates in a slotted mode, with a slot time equal to the packet transmission time plus the tuning delay (the time it takes a transmitter to tune from one wavelength to another). A *collision* occurs if two or more transmitters access the same channel in a given slot; all packets involved in a collision are lost.

We define σ_i as the probability that a new packet arrives at station i during a slot time, p_{ij} as the probability that a packet arriving at station i has station j as its destination, and $\sum_j p_{ij} = 1$. Whenever $C < N$, a number of receivers have to be assigned the same wavelength $\lambda_c, c = 1, \dots, C$. We let R_c , a subset of $\{1, \dots, N\}$, denote the set of receivers that share channel λ_c , $R_c = \{j \mid \lambda(j) = \lambda_c\}, c = 1, \dots, C$. Then $q_{ic} = \sigma_i \sum_{j \in R_c} p_{ij}$ is the probability that a packet with destination $j \in R_c$ arrives at i within a slot. Each station has C buffers, one for storing packets that need to be transmitted on each channel. The buffer for channel λ_c at station i has a capacity of L_{ic} packets; packets arriving to find a full buffer are lost.

2.1 Transmission Schedules

The Interleaved TDMA (I-TDMA*) protocol [1] is an extension of Time Division Multiplexing (TDM) over a multi-channel environment. In I-TDMA*, each station has exactly one chance per frame to transmit on each channel. I-TDMA* exhibits good performance under uniform traffic ($\sigma_i = \sigma_k, p_{ij} = p_{kl} \forall i, j, k, l$), but will be shown to perform poorly under non-uniform loads. Here, we are concerned with *weighted* TDM schemes, a generalization of I-TDMA*, whereby stations do not share the channels equally.

In a weighted TDM scheme, within each frame of length M , source i is allowed to transmit on channel λ_c in exactly

$a_{ic}(M)$ slots. In these slots, i may send a packet to any station $j \in R_c$. A *transmission schedule* indicates, for all i and c , which slots within a frame can be used for transmissions from i on wavelength λ_c , and is described by variables $\delta_{ic}^{(t)}$, $t = 1, 2, \dots, M$, called *permissions*, and defined as

$$\delta_{ic}^{(t)} = \begin{cases} 1, & i \text{ may transmit on } \lambda_c \text{ in slot } t \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Then, $a_{ic}(M) = \sum_{t=1}^M \delta_{ic}^{(t)}$. For $k = 1, \dots, a_{ic}(M)$, we let $d_{ic}^{(k)}$ denote the distance, in slots, between the beginning of the k -th slot that i has permission to transmit on λ_c , and the beginning of the next such slot, in the same or the next frame.

We are interested in schedules in which no collisions are possible. We say that a schedule of frame length M *provides full connectivity in the strong sense* iff it satisfies the following three conditions:

$$q_{ic} > 0 \Rightarrow a_{ic}(M) \geq 1 \quad \forall i, c \quad (2)$$

$$\sum_{c=1}^C \delta_{ic}^{(t)} \leq 1 \quad \forall i, t; \quad \sum_{i=1}^N \delta_{ic}^{(t)} = 1 \quad \forall c, t \quad (3)$$

Condition (2) specifies that, if the traffic originating at station i and terminating at stations listening on wavelength λ_c is nonzero, then there is at least one slot per frame in which i may transmit on wavelength λ_c . This guarantees full connectivity among the network stations. The first of (3) requires that each station be given permission to transmit on at most one channel within a slot t . Finally, the second of (3) implies that exactly one source may transmit on a given channel within a slot t . The last two constraints guarantee a collision-free operation. I-TDMA* is a special case of such a schedule.

By summing over all $t = 1, \dots, M$, (3) implies that:

$$\sum_{c=1}^C a_{ic}(M) \leq M \quad \forall i; \quad \sum_{i=1}^N a_{ic}(M) = M \quad \forall c \quad (4)$$

i.e., that a source may not be given permission to transmit in more than M slots within a frame, and that exactly M slots contain permissions for transmission on each channel. It has been shown [9] that (4) are also sufficient for constructing a schedule that satisfies (3). Thus, a schedule of frame length M providing full connectivity in the strong sense exists iff (2) and (4) are satisfied.

3 Minimization of Average Packet Delay

Let us now suppose that the C buffers at each station have infinite capacity ($L_{ic} = \infty \forall i, c$), and that the sets, R_c , of stations sharing channel λ_c have been decided upon. Observe that the buffers for distinct source-destination pairs do not interact. Consequently, the necessary and sufficient condition for stability is that the number of slots within a frame in which station i may transmit on channel λ_c , be greater

than the number of packets destined to a receiver listening on λ_c that are expected to arrive at station i during a number of slots equal to the frame length [7]:

$$Mq_{ic} < a_{ic}(M) \quad \forall i, c \quad (5)$$

We now define $b_{ic}(M)$ and $f_{ic}(M)$ such that $Mq_{ic} = b_{ic}(M) + f_{ic}(M)$, where $b_{ic}(M)$ is an integer, and $0 \leq f_{ic}(M) < 1$. Therefore, $b_{ic}(M) + 1$ is a lower bound on $a_{ic}(M)$ if the stability condition (5) is to be satisfied:

$$b_{ic}(M) + 1 \leq a_{ic}(M) \quad \forall i, c \quad (6)$$

Because of (4) the following two conditions must hold:

$$\sum_{c=1}^C (b_{ic}(M) + 1) \leq M \quad \forall i; \quad \sum_{i=1}^N (b_{ic}(M) + 1) \leq M \quad \forall c \quad (7)$$

However, since $f_{ic}(M) < 1$, it is easy to see that, unless M is sufficiently large, (7) may be violated, making it impossible to have $a_{ic}(M) \geq b_{ic}(M) + 1$. We now show how to select M so that (7) are satisfied. Consider channel λ_c , and select M'_c such that: $M'_c \sum_{i=1}^N q_{ic} \leq M'_c - N$, or $M'_c \geq \frac{N}{1 - \sum_{i=1}^N q_{ic}}$. For

this frame length, M'_c , we can easily get: $\sum_{i=1}^N (b_{ic}(M'_c) + 1) \leq M'_c$. Thus, by selecting $M' \geq \max_c \{M'_c\}$ we ensure that the second of (7) is satisfied.

By proceeding as above, we can see that frame length $M'' \geq \max_i \{M''_i\}$, where $M''_i \geq (C/(1 - \sigma_i))$, ensures that the first of (7) is satisfied. In order to satisfy both of (7), M has to be such that

$$M \geq \max\{M', M''\} \quad (8)$$

We now turn our attention to the issue of constructing schedules such that the average packet delay is minimized. Thus, we are seeking a solution to the following problem.

Problem 1 Given the number of stations, N , the number of available wavelengths, C , and the traffic parameters, $\sigma_i p_{ij}$, $i, j = 1, \dots, N$, find a schedule such that the network-wide average packet delay is minimized, assuming that buffers of infinite capacity are available at each station.

There are three dimensions to this problem: (a) the sets of receivers, R_c , sharing wavelength λ_c , $c = 1, \dots, C$, must be constructed, (b) the number of slots per frame, $a_{ic}(M)$, allocated to each source-channel pair (i, λ_c) must be obtained, and (c) a way of placing the $a_{ic}(M)$ slots within the frame, for all i, λ_c , must be determined. A similar study of a single-channel network [5] has shown that the optimization yields a hard allocation problem. The corresponding multi-channel optimization problem is even harder as the minimization is over all possible partitions of the set of receivers, $\{1, 2, \dots, N\}$, into sets R_c , $c = 1, \dots, C$. Our approach, then, is to first present a heuristic to obtain near-optimal schedules assuming that sets R_c are known. The issue of constructing these sets to achieve load balancing is has been addressed in [11].

Recall that the buffers for each source-channel pair are independent; therefore, if we consider each channel in isolation, all the results obtained in [5] are applicable, and provide a lower bound on the multi-channel problem, given a partition of $\{1, \dots, N\}$ into sets R_c . Consider channel λ_c ; the average packet delay is minimized [5] when (a) the percentage of time station i may transmit on channel λ_c is

$$x_{ic} = q_{ic} + (1 - \sum_{k=1}^N q_{kc}) \frac{\sqrt{1 - q_{ic}}}{\sum_{k=1}^N \sqrt{1 - q_{kc}}} \quad \forall i \quad (9)$$

and (b) for each source, i , the $a_{ic}(M)$ permissions for i to transmit on channel λ_c are equally spaced within the frame:

$$\forall i : d_{ic}^{(k)} = d_{ic} = \frac{1}{x_{ic}}, \quad k = 1, \dots, a_{ic}(M) \quad (10)$$

Note that x_{ic} and d_{ic} are independent of M . Given a frame length, M , satisfying (8), we assign a number of slots to the source-channel pair (i, λ_c) so that constraints (4) hold and

$$b_{ic}(M) + 1 \leq a_{ic} \leq \lceil Mx_{ic} \rceil \quad (11)$$

Once $a_{ic}(M)$ have been determined for all i and λ_c , we need to construct the schedule so that the permissions assigned to each source-channel pair are placed within the frame according to (10). This is not feasible in general, as d_{ic} may not be integers. Even if they are, scheduling the transmissions between all sources and channels in equally spaced slots may violate constraints (3). To overcome this problem, a golden-ratio policy was developed in [5], which requires that the frame length be a Fibonacci number. This policy places the permissions within the frame in intervals close to the ones dictated by (10), and achieves a delay very close to the lower bound.

Our approach is to use the golden ratio policy to place the permissions within each channel independently of the others. Considering channels in isolation, however, may cause a source to be assigned to transmit on two or more channels in the same slot, violating the first of (3). If this occurs, we must rearrange the schedule to remove these violations (this is always possible, since $a_{ic}(M)$ satisfy (4), and thus, a schedule providing full connectivity in the strong sense always exists). To this end, we use algorithm REARRANGE, described in [11], with a worst case complexity of $O(N^2M^2)$. We now propose the following Slot Allocation Heuristic.

Slot Allocation Heuristic (SAH)

1. If $C < N$, use the heuristic in [11] to determine the set of receivers, R_c , $c = 1, \dots, C$, that share each channel.
2. Select the smallest Fibonacci number, M , that satisfies (8), and obtain $a_{ic}(M)$ from (11) so that (4) hold.
3. For $c = 1$, use the golden ratio policy [5] to place the $a_{ic}(M)$, $i = 1, \dots, N$, slots for transmissions on λ_c . Repeat for $c = 2, \dots, C$, to obtain an initial schedule, $S_0(M)$.

4. Run algorithm REARRANGE [11], to perturb S_0 , producing a schedule, $S(M)$, satisfying constraints (3).
5. Repeat Steps 2 through 4 for the next Fibonacci number, M , up to M_{max} . Select the frame length, M , and schedule, $S(M)$, that yields the lowest average delay.

4 Minimization of Packet Loss Probability

Typically, the total number of buffers available at each station, which we will denote by $L_{i,max}$, is finite. The problem that arises then can be stated concisely as:

Problem 2 Given the number of stations, N , the number of available wavelengths, C , the traffic parameters, $\sigma_i p_{ij}$, $i, j = 1, \dots, N$, and the maximum number of buffers at each station, $L_{i,max}$, $i = 1, \dots, N$, (a) find a schedule, and (b) for all stations i , determine the buffer size, L_{ic} , for packets waiting for transmission on channel λ_c , so that $\sum_{c=1}^C L_{ic} = L_{i,max}$, and the network-wide probability of packet loss is minimized.

Thus, in addition to the three subproblems that comprise Problem 1, the $L_{i,max}$ buffers at each station have to be optimally partitioned into C queues. Recall, however, that a solution to Problem 1 minimizes the average packet delay, or, equivalently, the expected queue size of the CN buffers. We conjecture that, regardless of the buffer sizes, $L_{ic} \forall i, c$, the packet loss probability is minimized when the conditions specified by (9) and (10) are satisfied. Our approach, then, will be to construct the schedule for Problem 2 using SAH; we now consider how to partition the buffers at each station so as to minimize the packet loss probability.

4.1 Analysis

We first derive a lower bound for the packet loss probability, based on the observation that the mean queue length for source-channel pair (i, λ_c) is minimized when i is assigned to transmit on λ_c in slots which are exactly d_{ic} slots apart (see (10)). Since the buffers for each source-channel pair are independent, we may consider pair (i, λ_c) in isolation.

We observe the system at the instants just before the beginning of slots in which i may transmit on λ_c . Consider the l -th such slot. We define $r_{ic}^{(l)}(n, L_{ic})$ as the probability that i has n packets in its buffer (of size L_{ic}) for λ_c at the beginning of the l -th slot, $0 \leq n \leq L_{ic}$. We also define $P_{ic}(v)$, $0 \leq v \leq d_{ic}$, as the probability that v packets for λ_c arrive at i in the d_{ic} slots between the beginning of the l -th slot and the beginning of the $(l+1)$ -th slot, $P_{ic}(v) = \binom{d_{ic}}{v} q_{ic}^v (1 - q_{ic})^{d_{ic}-v}$.

$P_{ic}(> v)$, the probability that more than v packets arrive at i in the d_{ic} slots can be similarly defined. Now, i will have n , $n = 1, \dots, L_{ic} - 1$, packets for λ_c at the beginning of the l -th slot if (a) i had $n + 1$ packets at the beginning of slot $l-1$, transmitted one on λ_c in the $(l-1)$ -th slot, and no packets arrived since, and (b) i had $n - v$ packets, transmitted one, and $v + 1$ packets arrived¹. Similar observations can be

¹Except when $v = n$, in which case we require that n packets arrive. In (12) this case is covered by using $\min\{n, v + 1\}$.

made for $r_{ic}^{(l)}(L_{ic}, L_{ic})$. We can then write the following set of recursive equations for $l = 2, 3, \dots$. The initial conditions (14) are obtained by assuming that the frame starts at a slot in which i may transmit on λ_c .

$$r_{ic}^{(l)}(n, L_{ic}) = r_{ic}^{(l-1)}(n+1, L_{ic})P_{ic}(0) + \sum_{v=0}^n r_{ic}^{(l-1)}(n-v, L_{ic})P_{ic}(\min(n, v+1)) \quad n < L_{ic} \quad (12)$$

$$r_{ic}^{(l)}(L_{ic}, L_{ic}) = \sum_{v=0}^{L_{ic}} r_{ic}^{(l-1)}(L_{ic}-v, L_{ic})P_{ic}(\geq \min(L_{ic}, v+1))$$

$$r_{ic}^{(l)}(0, L_{ic}) = 1 - \sum_{n=1}^{L_{ic}} r_{ic}^{(l)}(n, L_{ic}) \quad (13)$$

$$r_{ic}^{(l)}(n, L_{ic}) = 0 \quad 1 \leq n \leq L_{ic}, \quad r_{ic}^{(1)}(0, L_{ic}) = 1 \quad (14)$$

If the stability condition (5) is satisfied, the system will eventually reach a steady state such that:

$$r_{ic}(n, L_{ic}) = \lim_{l \rightarrow \infty} r_{ic}^{(l)}(n, L_{ic}), \quad n = 1, \dots, L_{ic} \quad (15)$$

Then the probability of a packet arriving at station i been lost, given that the packet is destined to a receiver listening on λ_c and the buffer for that channel has a capacity of L_{ic} packets is

$$Q_{ic}(L_{ic}) = \sum_{n=0}^{L_{ic}} r_{ic}(n, L_{ic})P_{ic}(> L_{ic} - n) \quad (16)$$

The probability that a packet arriving at station i has to be transmitted on channel λ_c is just q_{ic}/σ_i . Therefore, the probability of packet loss at station i given a partition of the $L_{i,max}$ buffers into C queues of sizes L_{i1}, \dots, L_{iC} is

$$Q_i(L_{i1}, \dots, L_{iC}) = \frac{1}{\sigma_i} \sum_{c=1}^C q_{ic}Q_{ic}(L_{ic}) \quad (17)$$

4.2 Buffer Partitioning

Our problem now reduces to obtaining, for all i , queue sizes L_{i1}, \dots, L_{iC} , such that Q_i as given in (17) is minimized. Since enumerating all potential partitions, L_{i1}, \dots, L_{iC} , of the $L_{i,max}$ buffers into C queues and evaluating the packet loss probability of each using (17) is not feasible in general, we adopt a heuristic approach aimed at obtaining a near-optimal partition. Observe that $Q_{ic}(L') \leq Q_{ic}(L)$ for $L' > L$, and that $L_{ic} \geq 1 \forall c$. The following heuristic starts with queue sizes $L_{ic} = 1 \forall c$ and iteratively increments the queue size for the channel λ_c that results in the highest decrease for the packet loss probability as given in (17).

Buffer Partitioning Heuristic (BPH)

1. Initialize $L_{i1} \leftarrow \dots \leftarrow L_{iC} \leftarrow 1$ and $L_{i,max} \leftarrow L_{i,max} - C$. Repeat Step 2 while $L_{i,max} \geq 1$.

2. Find the channel λ_c for which $|q_{ic}(Q_{ic}(L_{ic} + 1) - Q_{ic}(L_{ic}))|$ has the greatest value. Set $L_{ic} \leftarrow L_{ic} + 1$ and $L_{i,max} \leftarrow L_{i,max} - 1$.

Assuming that $L_{i,max} > C$, the C values $Q_{ic}(1)$ and $Q_{ic}(2)$, have to be computed initially, for all c ; from then on, Step 2 will be repeated $L_{i,max} - C - 1$ times, which is the number of times (16) will be used with $L_{ic} > 2$. It is possible to trade the quality of the buffer partition for speed by allocating buffers to the various queues in chunks of size greater than one in Step 2.

5 Numerical Results

We consider the 8-station disconnected type traffic matrix with probabilities p_{ij} as shown in Figure 1. We let $\sigma_i = \sigma \forall i$. Figure 1 also shows the weighted TDM schedule constructed by SAH for this matrix. We used SAH ($M_{max} = 2,584$) to construct optimized schedules for $C = 2, 4$, and 8 channels, and values of σ from 0.01 to 0.99. We then obtained through simulation the delay and throughput curves of these schedules shown in Figure 2; the delay is given in slots, and the throughput in packets received per slot. A value of 100 in the delay plots denotes an infinite value for the average packet delay. In Figure 2 we also plot the delay and throughput curves of the I-TDMA* schedule. It is immediately evident that the schedules constructed by SAH outperform I-TDMA* by a wide margin, in terms of both delay and throughput; very similar behavior has been obtained for a variety of other traffic matrices [10].

We now assume that each station employs a finite number of buffers, $L_{i,max}$, and let $L_{i,max} = L_{max} \forall i$. Our objective is to compare the packet loss probability under two scenarios: (a) when the $L_{i,max}$ buffers available at each station are allocated according to the Buffer Partitioning Heuristic (BPH), and (b) when the $L_{i,max}$ buffers are equally partitioned among the various channels. Only optimized schedules are considered here. Figure 3 plots the packet loss probability curves, obtained through simulation, against the total number of buffers at each station, L_{max} , for $C = 10, 20, \sigma = 0.3$, and a 20-station ring-type matrix matrix (see [10]). Label "EP" denotes the equal sharing of buffers among the channels. The plots indicate that, the buffer allocation determined by BPH results in a performance improvement between one and two orders of magnitude over an equal partitioning scheme; similar results were obtain for many other traffic matrices.

6 Concluding Remarks

We presented heuristics to minimize the delay and packet loss probability for single-hop networks using TDM schedules. Techniques such as these are the first step towards WDM networks that dynamically adapt to the traffic pattern.

References

- [1] K. Bogineni, *et al.* Low complexity multiple access protocols for wavelength-division multiplexed photonic networks. *IEEE JSAC*, 11(4):590-604, 1993.

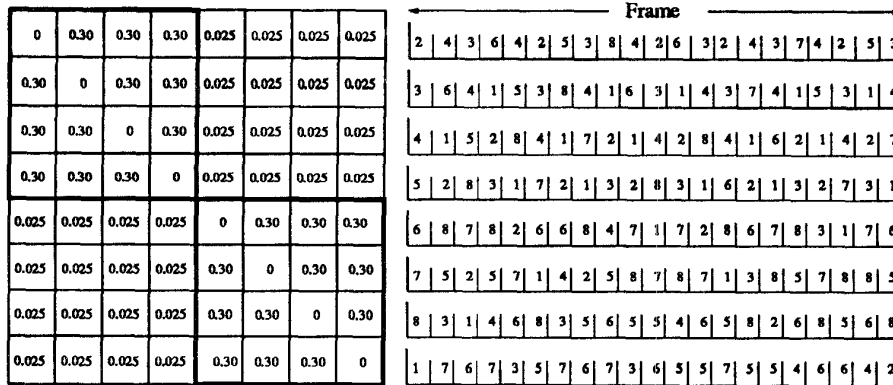


Figure 1: 8-station disconnected-type traffic matrix, and weighted TDM schedule for $M = 21, C = 8, \sigma = 0.7$

- [2] M-S Chen, *et al.* A media access protocol for packet switched wavelength division multiaccess metropolitan area networks. *IEEE JSAC*, 8(6):1048-1057, 1990.
- [3] R. Chipalkatti, *et al.* Protocols for optical star coupler network using WDM: Performance and complexity study. *IEEE JSAC*, 11(4):579-589, 1993.
- [4] P. E. Green. The future of fiber optic computer networks. *IEEE Computer*, 24(9):78-87, 1991.
- [5] M. Hofri and Z. Rosberg. Packet delay under the golden ratio weighted TDM policy in a multiple-access channel. *IEEE Trans. on Inform. Theory.*, 33(3):341-349, 1987.
- [6] P. A. Humblet, *et al.* An efficient communication protocol for high-speed packet-switched multichannel networks. *IEEE JSAC*, 11(4):568-578, 1993.
- [7] L. Kleinrock. *Queueing Systems, Volume 1: Theory.* John Wiley & Sons, New York, 1975.
- [8] B. Mukherjee. WDM-Based local lightwave networks Part I: Single-hop systems. *IEEE Network Magazine*, pages 12-27, May 1992.
- [9] G. N. Rouskas. *Single-Hop Lightwave WDM Networks and Applications to Distributed Computing.* PhD thesis, Georgia Inst. Technology, Atlanta, GA, May 1994.
- [10] G. N. Rouskas and M. H. Ammar. Minimizing Delay and Packet Loss in Single-Hop Lightwave WDM Networks Using TDM Schedules. Tech. Report TR-94-12, North Carolina State University, Raleigh, NC, 1994.
- [11] G. N. Rouskas and M. H. Ammar. Analysis and Optimization of Transmission Schedules for Single-Hop WDM Networks. *IEEE/ACM Transactions on Networking*, April 1995.

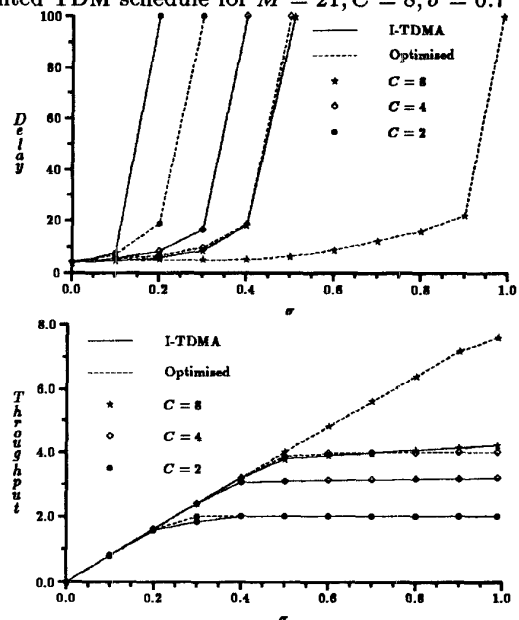


Figure 2: Delay/throughput curves for matrix of Figure 1

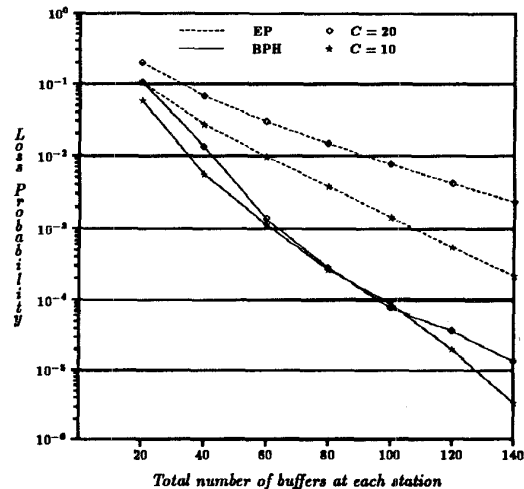


Figure 3: 20-station loss prob. curves ($C = 10, 20, \sigma = 0.3$)