A RECEIVER-ORIENTED MESSAGE SCHEDULING ALGORITHM FOR WDM LIGHTWAVE NETWORKS

Ma Maode, Babak Hamidzadeh*, and Mounir Hamdi

Department of Computer Science
Hong Kong University of Science and Technology, Kowloon, HongKong
*Department of Electrical and Computer Engineering
University of British Columbia, Vancouver, Canada

Abstract

In scheduling the transmission of messages in a WDM network, we are facing not only the channel assignment problem but also message sequencing problem. In this paper, we propose and evaluate a new reservation-based message scheduling algorithm for single-hop, WDM passive star lightwave network which addresses both the assignment aspect and the sequencing aspect of the problem. We compare the performance of the algorithm which only addresses the channel assignment issue with the performance of our new algorithm theoretically and experimentally and demonstrate that our new algorithm has significant improvement over the performance of a WDM network with a scheduling algorithm that does not consider the message sequencing problem. As a result, we anticipate that this research can open new directions into the problem of message scheduling on WDM networks.

Keywords: *Optical network, Wavelength-Division-Multiplexing (WDM), Medium access control protocol, Scheduling algorithm.*

1. Introduction

The recent developments in fiber optic technology is making the design of gigabit networks possible. Gigabit networks can provide a very large bandwidth capable of carrying various transmission signals with very different frequencies. The technique of *Wavelength Division Multiplexing* (WDM) is shown to be an effective method to utilize the large bandwidth of an optical fiber. Several system structures of the optical WDM networks have been proposed as in [1, 2]. A typical and simple network is the structure with a single-hop topology which directly connect the network nodes to a passive star coupler [3]. Based on the hardware structure of a WDM optical network, multiple media access control protocols are needed to schedule the messages to be transmitted through the multiple channels of the optical fiber. In particular, the design of an access protocol that make efficient use of the channel resources while satisfying the messages and system constraints is highly expected. The protocols and algorithms proposed on this system structure of WDM optical network can be divided into two different categories: the pre-allocation-based techniques as in [3, 4, 5, 6]; and the reservation based techniques as in [7, 8, 9, 10, 11]. The pre-allocation-based techniques assign transmission rights to different nodes in a static and predetermined manner. The reservation-based algorithms reserve one of the available channels as the control channel to transmit global information about messages to all the nodes in the network. The other channels are used to transmit real messages. Reservation-based techniques are more dynamic in nature and assign the transmission rights to the messages based on run-time availability of the receiving node and the channel in the network. These algorithms are especially beneficial to aperiodical and variable-length messages. The objective of the reservation-based protocol is to efficiently schedule message transmission while effectively avoiding both data channel collisions and message receiver collisions. Our work in this paper focuses on the reservation-based technique.

The channel assignment issue of the reservation-based scheduling has been extensively addressed in the literature. While the message sequencing issue has not received much attention among current available reservation-based scheduling algorithms. Most of the existing reservation-based approaches schedule messages individually and independently of one another. They ignore the fact that the way to choose the order of the message transmission may affect the performance of the protocol of the network. To the best of our knowledge, only our previous paper [12] has addressed the issue of sequencing the messages to be transmitted to improve the performance of a WDM optical network. Using that protocol, the order of message transmission is determined by the message length. This algorithm has been shown to significantly improve the performance of a WDM network. However, if the difference among the lengths of the messages in a transmission frame is small, the effectiveness of the algorithm will be decreased. Also it fails to consider the problem that messages can be blocked when they try to avoid receiver collisions.

In this paper, we propose and evaluate a new reservation-based protocol for scheduling variable-length messages. In our algorithm, we consider to use the information of the receivers' state to impose a priority on the ordering in which messages are transmitted. We also consider to use the *Earliest Available Time Scheduling* (EATS) algorithm [9] to deal with the problem of channel assignment.

The remainder of this paper is organized as follows. Section 2 specifies our system model and the problem to be addressed. Section 3 presents our new algorithm and the techniques involved. Section 4 provides our analytical performance model of the RO-EATS algorithm. Section 5 shows our results from simulation experiments and theoretical analysis to compare the performance of the two algorithms. Finally, section 6 concludes the paper with a summary of the results.

2. WDM Model and Scheduling Problem

In this paper, we consider message transmission in a single-hop WDW optical network whose nodes are connected to a passive star coupler via two ways fibers. Each direction of the fiber supports C+1 WDM channels with the same capacity and there exist N nodes in the network. The C channels, referred to as data channels, are used for message transmission. The remaining channel, referred to as the control channel, is used to exchange global information among nodes about the messages to be sent. The control channel is the basic mechanism for implementing the reservation scheme. Each node in the network has two transmitters and two receivers. One transmitter and one receiver are fixed and are tuned to the control channel. The other transmitter and receiver are tunable and can tune into any data channel to access messages on those channels.

The nodes are assumed to generate messages with variable length which can be divided into several equal-sized packets. The basic time interval on the data channels is the transmission time of one packet. The nodes are divided into two non-disjoint sets of source nodes s_i and destination nodes d_j. A queue of messages to be transmitted is assumed to exist at each source node s_i.

A Time Division Multiple Access (TDMA) protocol is used on the control channel to avoid collision of the control packets belonging to different nodes. According to this protocol, each node can transmit a control packet during a predetermined time slot. The basic time interval on the control channel is the transmission time of a control packet. N control packets make up one control frame on the control channel. Thus, each node has a corresponding control packet in a control frame, during which that node can access the control channel. The length of a control packet is a system design parameter and depends on the number of messages l about which each node is allowed to broadcast, and the amount of control information on each message (e.g., the address of the destination node, message length). Figure 1 illustrates some of the basic concepts in our model.

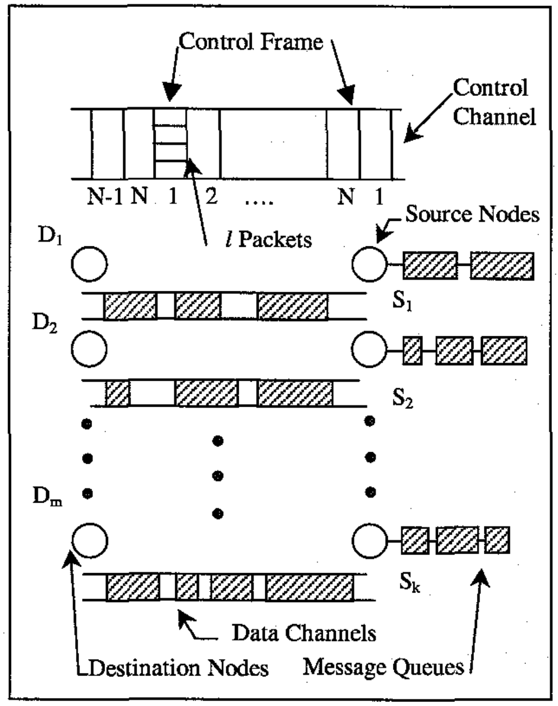

Figure 1. *Data and control channel configuration*

3. Scheduling Algorithm

Like the other reservation-based protocols, the procedure of messages transmission and reception in this system model works as follows: A node has to transmit a control packet on the control channel in its assigned time slot before sending a message to its destination node. After one round-trip propagation delay, the destination node and the other nodes in the network receive the control packet. Then the distributed scheduling algorithm is invoked by each node to determine the data channel and transmission time slots of this message transmission. Once the message is scheduled, the sending node's transmitter will tune to the scheduled data channel and sends this message at the scheduled time. The receiver of this message destination node should tune to that channel to be ready to receive the message. After the propagation delay, the message will arrive at the destination node and it is received.

We form our new *receiver-oriented* scheduling algorithm based on the basic channel assignment algorithm, EATS [9]. Our proposed receiver-oriented algorithm (RO-EATS) first considers the earliest available receiver among all the nodes in the network and then selects a message which is destined to this receiver from those which are ready and identified by the control frame. After that, a channel is selected and assigned to the selected message by the EATS algorithm. Our new algorithm makes full use of the global information on each message and the system states of the network to improve the performance of the WDM optical network.

The basic idea of the EATS algorithm is to assign a message to a data channel that has the earliest available time slots among all the channels in the network. In order to keep a record on the channel and receiver usage and their states, two tables are used and reside on each node, which are denoted as *Receiver Available Time* array, RAT, and *Channel Available Time* array, CAT. RAT records the non-available times of the receiver on each node from the current time in the packet slot unit. CAT records the non-available times of each channel from the current time. Both of them are dynamically decreasing with time units. With these global information on each node, the distributed EATS works as follows: Transmit a control packet on the control channel; Choose a channel with the earliest available time; Calculate the transmission time of a message based on two tables; and Update the two tables according to newly scheduled message.

Based on the network structure, the transmission channels and the receivers can be considered as two kinds of resources in series to be exclusively occupied by messages when they are transmitted. This fact may lead messages to be blocked when the scheduler tries to avoid the receivers' collision. To this problem, we have the idea of keeping two or more scheduled consecutive messages away from being destined to the same destination node. As the CAT and RAT provide the information on the states of these two kinds of resources respectively and they are the global information available to all the nodes in the network, our scheme to choose a suitable message to be transmitted is to use the information on the states of the receivers according to RAT to avoid lots of messages being destined to one or a few nodes at the same time. This message's selecting scheme has the ability to sequence the messages presented in the control frame according to the information of the receiver states.

The RO-EATS algorithm can be used to schedule messages belonging to one control frame. The messages' queue at each node can be considered to have only one capacity when this algorithm is in execution.

We assume that there are M nodes and W channels. The messages have variable lengths that follow an Exponential distribution. The messages can be transmitted from source node i to destination node j, where $i \neq j$, and $i, j \in M$. The RAT Table can be expressed as an array of M elements, one for each node. $RAT[j] = n$, where $j = 1, 2, ..., M$, means that node j will be free after n time slots. The CAT Table can be expressed as an array of W elements, one for each channel. $CAT[k] = m$, where $k = 1, 2, ..., W$, means that channel k will be available after m time slots. The RO-EATS algorithm for one message transmission can be expressed in detail as follows:

The RO-EATS algorithm

<u>message choosing scheme:</u>

Sort RAT in non-decreasing order by the value of RAT to form a new $RAT'[j]$;

Check whether the current message is destined to node j, j=0;

If no, jump to **n1**; if yes, jump to channel assignment scheme;

n1:

The current message waits until next time to be checked;

Check other messages in the frame whether one of them destined to node j;

If no, jump to **n2**; if yes, jump to channel assignment scheme;

n2:

$l = l + 1$;

Check other messages in the frame whether one of them destined to node j;

If no, jump to **n2**; If yes, jump to channel assignment scheme;

channel assignment scheme:

Sort CAT in non-decreasing order by the value of CAT to form a new $CAT'[k]$;
Use the channel k to transmit the message;
Calculate $r = RAT[j] + T$, $t_1 = \max(CAT[k], T)$, $t_2 = \max(t_1 + R, r)$;
where T is the Transmitter's tuning time,
R is the propagation delay.
Schedule the message transmission time at $t = t_2 - R$;
Update $RAT[j] = t_2 + m$, $CAT[k] = t_2 - R + m$.
where m is the message length.

4. Performance Analysis

To compare the performance of the scheduling algorithms, we set up an approximate mathematical model for a WDM network using the RO-EATS algorithm. The goal of this model is to study the performance of a WDM network taking our new algorithm under the condition of a limited number of data channels in the system. It is the case that in a real WDM network the number of channels are less than the number of the nodes. The performance metric which we focus on is the average delay time of the messages in the network.

In order to make the model show the main characteristics of the system, several assumptions are used and can be summarized as follows:

1) Tuning time is ignored, is set to be 0.
2) The message population is finite, which is the number of the nodes in the system, M, as we consider one message at the head of each queue of each node.
3) For each of the nodes, the arrival messages have their own independent message arrival rate distribution with the same mean value of λ and is a Poisson process.
4) The message transmitted by a node can be destined to every other node with equal probability.
5) The message length of the message stream at each node is a random variable following the Exponential distribution with same mean value of $1/\mu$.

As the RO-EATS algorithm is basically a frame scheduling algorithm mentioned in [12], the number of steady states is finite. A WDM network using the RO-EATS algorithm can be modeled as a k steady states Markov process, which contains k messages in the system. Those messages may be transmitted by the channels which are the earliest available to them or in the queue waiting for the channels. The Markov model can be expressed as follows: The message arrival rate into the system, when the system is in the k state, equals to $(M - k) \times \lambda$, where $k = 0, 1, 2, \ldots, M$. The service rate can be expressed by: when $k \le W$: $\mu_k = \mu \times k$; when $k > W$: $\mu_k = \mu \times W$. Where μ_k is the system service rate when the system is at state k, and μ is the reciprocal of the mean value of message length. We model the system using the RO-EATS algorithm as a load dependent $M/M/m$ server. The service rate depends on the number of messages staying in the system. The steady-state probability can be determined as following formula after the message arrival rate and system service rate have been fixed.

$$P_k = \frac{\prod_{i=1}^{k} \frac{\lambda(M-i+1)}{\mu_i}}{1 + \sum_{k=1}^{M} \prod_{i=1}^{k} \frac{\lambda(M-i+1)}{\mu_i}}$$

Where the M is the number of the nodes in the system; λ is the mean message generation rate at each node; and the P_k is the probability of the system to be in the state k.

As a result, the average delay, D, while the system in the steady state can be given by:

$$D = \frac{\sum_{k=1}^{M} kP_k}{\left(M - \sum_{k=1}^{M} kP_k\right)\lambda}$$

5. Comparison of Results

In this section, we first present the results of the above analytical model of the system using the EATS and the RO-EATS algorithms respectively. We also present results of the simulation experiments of the systems using the different algorithms. We compare these results to verify the accuracy of the analytical model. In both analysis and simulation, we set the parameters of system as following: the number of nodes is fixed as 50, the number of channels is 4, the average messages length is 20 time slots and is following an Exponential distribution. These results are shown in the figures 2 – 4 as following.

In Figure 2, we present the analytical and simulation results. For the analytical results, we can see that when the mean message arrival rate varies from 0.002 to

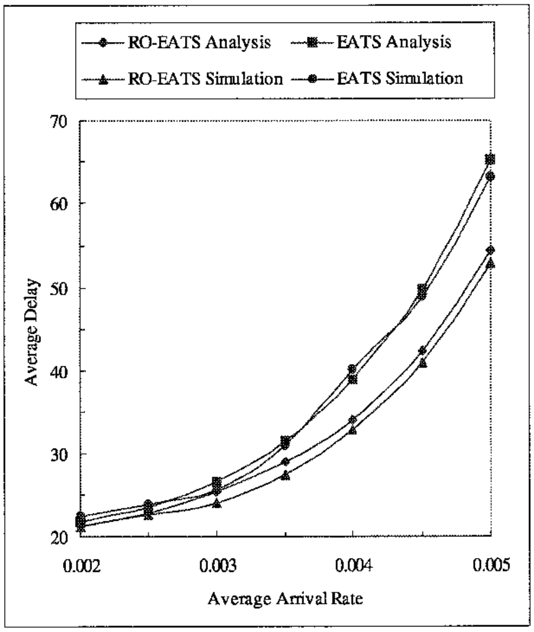

***Figure 2.** Average Delay vs. Average Arrival Rate*

0.005 for each node, the average delay of the messages using the EATS algorithm ranges from 21 to 65 time slots; while that of the RO-EATS algorithm ranges from 22 to 55 time slots. It shows that the RO-EATS algorithm has improved the system performance as the average delay of the messages in the system using the RO-EATS algorithm is quite less than that in the system using the EATS algorithm. For the results of experiments, it is shown that when the mean message arrival rate changes from 0.002 to 0.005 for each node, the average delay of the messages using the EATS algorithm ranges from 23 to more than 60 time slots; while that of the RO-EATS algorithm ranges from 22 to slightly more than 50 time slots. The difference on the average delay between two algorithms reaches up to nearly 20% of the whole average delay value at the critical point. The simulation results confirm the fact that the RO-EATS algorithm can improve the system performance quite a lot as we did by using our simple mathematical model. This fact has validated our approximated analytical model and also verified our simulation results.

We have also conducted some other simulation experiments to show the improvements on the other system performance metrics by the RO-EATS algorithm. One of the results is to show how the system throughput is affected by the average message arrival rate in Figure 3.

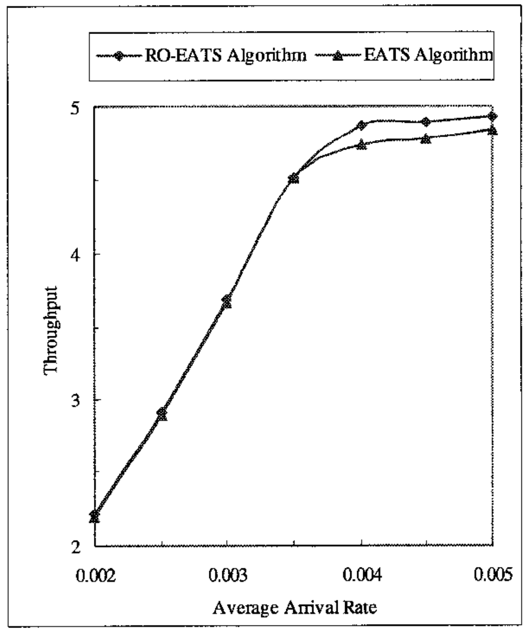

***Figure 3.** Throughput vs. Average Arrival Rate*

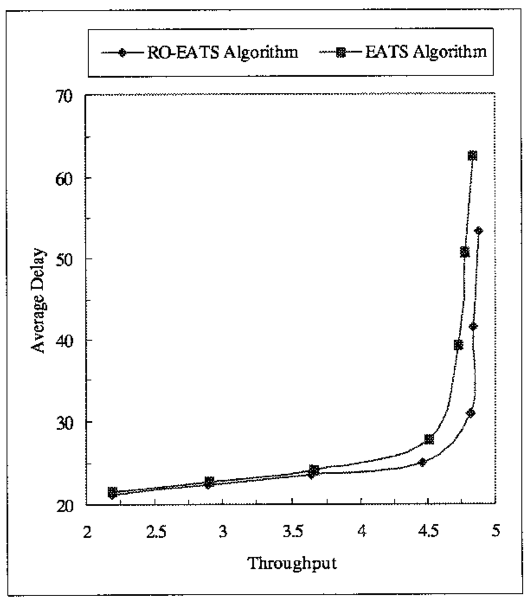

***Figure 4.** Average Delay vs. Throughput*

In Figure 3, the system throughput is measured by number of packets being transmitted per unit time slot. We can see that the throughput of the WDM network using RO-EATS algorithm is larger than the throughput of the WDM network using the EATS algorithm, especially when the average message arrival rate reaches the point of 0.004. This fact implies that the system using the RO-EATS algorithm has better capacity than the system using the EATS algorithm to balance the work load. The RO-EATS algorithm can not only improve the average message delay but also improve the system throughput.

We combine the simulation results shown in Figure 2 and Figure 3 into Figure 4 to show the relationship between the average message delay and the system throughput. From this figure, we can see that the performance of RO-EATS algorithm is better than that of the EATS algorithm in the sense that in a range of certain values of system throughput, the average message delay of the system using the RO-EATS algorithm is always less than that of the system taking the EATS algorithm.

6. Conclusions

In this paper, we proposed a new reservation-based algorithm for scheduling variable-length messages in a single-hop WDM passive star network, denoted RO-EATS. Unlike many existing reservation-based techniques, the proposed algorithm addresses both message sequencing and channel assignment aspects of the scheduling problem simultaneously. RO-EATS is shown to reduce the average messages delay and increase the system throughput of a WDM network. The results of our analysis and simulation show that the more global information on the system states and the messages to be transmitted are used, the more improvements to the system performance can be expected. The main contribution in this paper is that the global information on the states of message destination nodes has been used to analyze and solve the messages blocking problem caused by avoiding receiver collision, which is ignored by previous scheduling algorithms.

References

1. B. Mukherjee, "WDM-based Local Lightwave Networks- Part I: Single-Hop Systems," *IEEE Network*, pp. 12-27, May 1992.
2. B. Mukherjee, "WDM-based Local Lightwave Networks- Part II: Multi-Hop Systems," *IEEE Network*, pp. 20-32, July 1992.
3. K. Bogineni, K. M. Sivalingam, and P. W. Dowd, "Low-Complexity Multiple Access Protocols for Wavelength - Division Multiplexed Photonic Networks," *IEEE Journal on Selected Areas of Communications*, 11 (4), pp. 590-603, May 1993.
4. A. Gantz and Y. Gao, "Time-Wavelength Assignment Algorithms for High Performance WDM Star Based Systems," *IEEE Transactions on Communications*, 42(2/3/4), pp. 1827-1836, Feb/Mar/April 1994.
5. G. N. Rouskas and M. H. Ammar, "Analysis and Optimization of Transmission Schedules for Single-Hop WDM Networks," *Proceedings of IEEE INFOCOM*, pp. 1342-1349, March 1993.
6. M.S. Borella, and B. Mukherjee, "Efficient Scheduling of Nonuniform Packet Traffic in a WDM/TDM Local Lightwave Network with Arbitrary Transceiver Tuning Latencies," *Proceedings of IEEE INFOCOM*, pp. 129-137, 1995.
7. K. Bogineni and P. W. Dowd, " A Collisionless Multiple Access Protocol for a Wavelength Division Multiplexed Star-Coupled Configuration: Architecture and Performance Analysis, " *Journal of Lightwave Technology*, 10(11), pp.1688-1699, November 1992.
8. R. Chipalkatti, Z. Zhang, and A. S. Acampora, " Protocols for Optical Star-Coupler Network Using WDM: Performance and Complexity Study," *IEEE Journal on Selected Areas of Communications*, 11(4), pp.579-589, May 1993.
9. F. Jia, B. Mukherjee, and J. Iness, "Scheduling Variable-Length Messages in a Single-Hop Multichannel Local Lightwave Network," *IEEE/ACM Transactions on Networking*, Vol. 3, No. 4, pp. 477-487, August 1995.
10. C.S. Li, M.S. Chen, and F.F.K. Tong, "POSMAC: A Medium Access Protocol for Packet-Switched Passive Optical Networks using WDMA," *Journal of Lightwave Technology*, 11(5/6), pp. 1066-1077, May/June 1993.
11. N. Mehravari, "Performance and Protocol improvements for Very High-Speed Optical Fiber Local Area Networks using a Passive Star Topology," *Journal of Lightwave Technology*, 8(4), pp. 520-530, April 1990.
12. B. Hamidzadeh, M. Maode and M. Hamdi, "Message Sequencing Techniques for On-Line Scheduling in WDM Networks."*Proceedings of GLOBECOM' 1997.*