A High Performance Channel Sorting Scheduling Algorithm Based On Largest Packet

P.G.Sarigiannidis, G.I.Papadimitriou, Senior Member, IEEE and A.S.Pomportsis
Department of Informatics, Aristotle University, Box 888, 54124 Thessaloniki, Greece

***Abstract*-We present a new algorithm for wavelength division multiplexing (WDM) star optical networks. The resulting protocol is pre-transmission coordination-based without packet collisions. The proposed algorithm tries to schedule the transmission requests of the network, with the assistant of a prediction mechanism. With the prediction of the packet requests the algorithm manages to decrease the calculation time of building the schedule matrix, which consists of the final schedule of the packets of each node of the network. This reduction is achieved by pipelining the computation process, at the same time with the adoption of a new order in which nodes` requests are serviced. This modification of the service order leads to an increment in terms of network throughput and channel utilization. We compare the performance of two algorithms in terms of channel utilization, network throughput and mean time delay, under different sets of channel values and we present the simulation results.**

I. INTRODUCTION

The today existence of the electronic switching speeds cannot satisfy the constantly increasing demands for high speed within local area networks (LAN), metropolitan area networks (MAN), and wide are networks (WAN). Photonic networks are expanding rapidly, due to the enormous bandwidth of optical fiber technology [1]. Wavelength division multiplexing (WDM) technology transports tens to hundreds of wavelengths per fiber, with each wavelength modulated at 10 Gb/s or more. Such systems may result in gigabit-per-second data rates in independent channels, which transmit simultaneously data flows to a single or multiple users [2]. In this paper, we focus on WDM LANs, based on broadcast and select (BS) architecture. More specifically, in this context the network consists of a number of nodes, a set of channels, and a passive (or physical) star coupler [3]. At a given time every node can select an available channel in order to transmit its data to the appropriate destination. The transmission comes through by the passive star, which broadcasts all input data to all outputs and allows transparent and immediate transfer of data from the transmitters to the receivers.

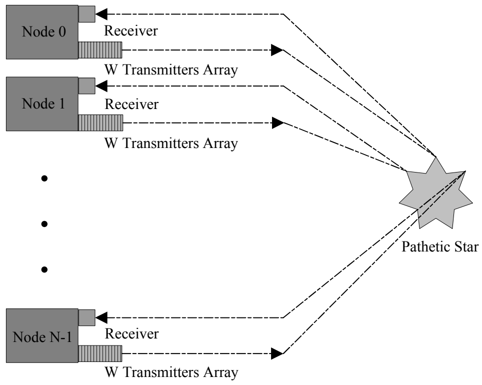

Figure 1. The network model

We consider data transmission in a single-hop WDM optical network whose nodes are connected to a passive star coupler via a two way fiber (Fig. 1). The network comprises N nodes and W channels. Each node disposes an array of tunable transmitters, which provides the transmission of data to the appropriate channels. With this approach the channels are pre-allocated to the nodes. Moreover, the node has a fixed receiver, which allows receiving data in the particular channel, which is dedicated to each node, known also as home channel. A home channel may be shared if the number of nodes exceeds the number of channels within network [2]. Thus, the network is multicasting and unicasting. The connection of the channels is accomplished through the passive star coupler.

A medium access control (MAC) protocol manages to allocate the available channels to the nodes, which are ready to transmit to a specific destination. In order to export conclusions for the functionality of a MAC protocol, we should investigate the two types of collisions that are possible to occur in WDM BS networks [4]. Firstly, a channel collision occurs when two or more nodes try to transmit within the same wavelength simultaneously. Secondly, a receiver collision occurs when two or more nodes try to transmit the data simultaneously to the same node in different wavelengths. Undoubtedly, such a case is possible only in architectures with tunable receivers.

MAC protocols are generally categorized as either pre-transmission co-ordination based or pre-allocation [4]. In the case where in the network there is at least one channel dedicated to the coordination of channels and their transmission time, then the protocol is based on pre-transmission coordination. In a different case, i.e., if the network does not make use of a separate channel for the node transmission control, the protocol is pre-allocation based. Undoubtedly, at many instances it is observed that protocols do not dispose a separate control channel but exert control through control packets. We suppose that the time is divided in time frames. Each frame is composed of a reservation phase and a transmission phase. Also, each frame consists of a number of timeslots, during which the reservation and the packet transmission take place. In pre-transmission co-ordination based protocols the algorithm accepts the time demands of each node of the network and stores them in a transmission frame, called traffic demand matrix, $D=[d_{i,j}]$. The traffic demand matrix show off the load of the network.

II. BACKGROUND

A very important scheduling algorithm of pre-transmission coordination based networks is online interval-based scheduling (OIS) [5]. OIS is an online algorithm and that means that begins schedule computation just after reading the requests of the first node. So the algorithm needs only a part of the demand matrix to function. That has the advantage to save schedule computation time, because starts to build the schedule matrix, before the total acceptance of the nodes' demands.

It is important to refer that OIS starts to operate once a set of requests by node n is known. Let us suppose that the node n demands t_1 timeslots to accomplish the transmission, via channel w_1. OIS searches for availability between timeslot t and timeslot $t + (t_1\text{-}1)$. If during this time gap node n is not engaged to any other channel w_1 ($w \neq w_1$) then the coordination is accomplished and the time gap t to $t + (t_1\text{-}1)$ is registered to node n, with channel w. In the next step OIS refresh the lists and examines the remaining demands of the rest N-1 nodes of the network. In this way, every timeslots within the final schedule matrix of OIS contains a registration of transmission by a specific node via a specific channel. Of course if there are no registered nodes for some specifics spaces in the schedule matrix the equivalent channel stays idle and the time gap unused.

A notable continuation of OIS is predictive online scheduling algorithm (POSA) [6]. Its aim is to extend and improve OIS, by decreasing the computation time of the scheduling matrix. This is accomplished with the aim of hidden Markov chains, which allows to algorithm to predict the demands of the nodes for the next frame. In the same time POSA transmits the data of the current time. This parallel action leads to a significant time saving, since the algorithm does not lose time, by waiting the delivery of the node's demands. In addition, constructs the scheduling matrix immediately, based on the prediction, made during the previous frame. It is clear that such a prediction is based on the real requests of the nodes. POSA collects these requests from the passed frames and stores them into history queues. POSA uses two different algorithms. In the beginning of each frame POSA runs the learning algorithm and collects and informs its history queues about the new demands. After the learning algorithm POSA apply the second algorithm, known as prediction algorithm. This part of POSA is responsible for the prediction of the following frame. Thus during the prediction algorithm POSA tries to predict the demands of the frame which follows.

III. FM-POSA

The proposed new protocol is called first max-predictive online scheduling algorithm (FM-POSA). The first part of the name (FM) is an acronym related to the operation of the scheduling algorithm while the second part (POSA) states that the proposed algorithm is based on POSA which it evolves and improves [6]. The algorithm operates in three independent phases, the learning phase, the shifting phase and the prediction phase. In the first phase, the algorithm monitors the network traffic and tries to learn about the variations in traffic. The algorithm also updates the history queues about the changes in traffic so that it can make accurate predictions later. In the second phase the algorithm stops monitoring and updating and enters the prediction phase. Finally, in the last and most important phase the algorithm predicts the nodes' requests for the next frame based on its learning phase. In addition the algorithm performs the forwarding of packets to their destinations.

The new element that is introduced by FM-POSA is the order in which the nodes' requests are processed. It is easy to see that POSA process the nodes' requests one after the after (serially) starting from the first node and finishing with the last one. This means that POSA does not examine the size or the behavior of the nodes in detail but always processes the requests in the same order. FM-POSA on the other hand records and compares the nodes requests based on a key point i.e. the largest packet size. Thus, FM-POSA searches for the largest packet and sorts nodes according to this. The first node that is served is the one with the largest packet following by the one with the second largest packet and ending with the one with the smallest packet. It is worth to mention that if two nodes have packets of equal size the node that will be served first is randomly selected.

It would be useful to see an example of the processing of a demand matrix in order to understand the operation of both algorithms. Consider the following demand matrix: $D = \begin{bmatrix} 2..2..1 \\ 2..3..3 \\ 6..4..1 \end{bmatrix}$. We must remind you that the rows in this table are the network nodes and the columns are the network channels. It is obvious that in this table that we study we have nine independent predictors where every one of them predicts for a node-channel pair. With the above data, predictor $p_{0,0}$ has predicted that node N_0 will require two timeslots for the scheduling of packets to their destinations using channel W_0. The same logic stands for all other predictors until the last one $p_{2,2}$ which predicts one timeslot for channel W_2 and N_2.

Given this demand matrix the OIS and POSA algorithms would construct a scheduling matrix, which the service of requests starts from the first node N_0 and continues with node N_1 and finishes with node N_2. On the other hand, FM-POSA operates gradually in two actions:

Action 1: Search and locate the largest packet from the requests of the nodes and save it in a vector called MAX. By executing this action FM-POSA forms the following vector MAX. MAX = $\begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$. By observing the vector MAX, it is easy to see that the longest request for node N_0 is equal to two timeslots, for node N_1 three timeslots and for node N_3 six timeslots.

Action 2: Reorder the elements in MAX in descending order.

By executing this action FM-POSA forms the final vector S_MAX: S_MAX = $\begin{bmatrix} 6 \\ 3 \\ 2 \end{bmatrix}$

The reordering of vector MAX means that the processing of requests in order to produce the scheduling matrix is dictated by the final vector S_MAX. In other words the first node that is processed is the one with the longest request in timeslots, i.e. node N_2 with 6 timeslots, the second node that is processed is node N_1 and the last one is node N_0.

In this context let us examine the final form of the schedule matrix, constructed by POSA-OIS and FM-POSA. Given the demand matrix D of the above example the final schedule matrix, constructed by POSA will be as this in Figure 1.

timeslots

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
W_0	N_0	N_0	N_1	N_1	N_2	N_2	N_2	N_2	N_2	N_2	idle	idle	idle	idle
W_1	idle	idle	N_0	N_0	N_1	N_1	N_1	idle	idle	idle	N_2	N_2	N_2	N_2
W_2	N_2	idle	idle	idle	N_0	idle	idle	N_1	N_1	N_1	idle	idle	idle	idle

Figure 2. The final schedule matrix, constructed by POSA

It is easy to observe that POSA spends a total 14 timeslots to finalize the schedule matrix. Also, POSA wastes a total 18 idle sub-timeslots. So POSA losses a 43% percent of matrix's cells:

$$\frac{18...idle...subtimeslots}{14...overall...timeslots * 3...channels} = 0.43$$

In contrary FM-POSA will construct the following schedule matrix (Fig. 2):

timeslots

	0	1	2	3	4	5	6	7	8	9	10
W_0	N_2	N_2	N_2	N_2	N_2	N_2	N_1	N_1	N_0	N_0	idle
W_1	N_1	N_1	N_1	N_0	N_0	idle	N_2	N_2	N_2	N_2	idle
W_2	N_0	idle	idle	N_1	N_1	N_1	idle	idle	idle	idle	N_2

Figure 3. The final schedule matrix, constructed by FM-POSA

If we examine the schedule matrix of Figure 3, we will notice that FM-POSA starts to construct the matrix, by processing the requests of N_2. Then the algorithm continues with N_1's requests and completed the construction with N_0's requests. This service order shifting leads to an evident gain, in terms of timeslots, which can be translated in real time. FM-POSA spends a total of 11 timeslots, and concurrently wastes only 9 idle subtimeslots. Hence, FM-POSA losses only 27% percent of matrix's cells:

$$\frac{9...idle...subtimeslots}{11...overall...timeslots * 3...channels} = 0.27$$

IV. Simulation Results

This section presents the simulation results of the two algorithms POSA and FM-POSA. The two algorithms have been studied in the terms of utilization, network throughput and mean time delay, under uniform traffic. We consider two network models. The first consists of 4 channels and a row of nodes (10, 20, 30, 40, and 50). The second consists of 8 channels and a row of nodes (10, 20, 30, 40, and 50). It is important to refer that N symbols the number of nodes, W symbols the number of channels, K is the maximum value for incoming packets. Also, it is important to pinpoint that the speed of the line has been defined at 2.4 Gbps. The tuning latency is considered to be equal to zero for simplicity reasons. The duration of the simulation is 10000 frames, from which 1000 belongs to learning phase and the rest to the prediction phase. Finally, K is not constant but equal to FLOOR(NW/5).

The results from the comparison between the two algorithms in terms of channel utilization proves that FM-POSA remains constantly better that POSA, either for 4 or for 8 channels (Figures 4, 5). The results from the comparison between the two algorithms in terms of throughput proves that FM-POSA remains again constantly better that POSA, either for 4 or for 8 channels (Figures 6, 7). Lastly, the results from the comparison between the two algorithms in terms of throughput vs. delay shows that as the time delay is increased FM-POSA precedes POSA. Concurrently, FM-POSA presents a lower mean time delay than POSA, for each value of the workload (Figures 8, 9).

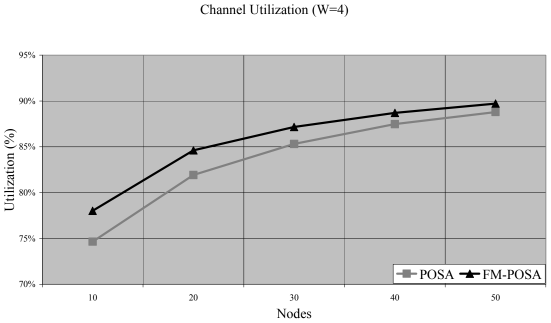

Figure 4. Channel Utilization with 4 channels

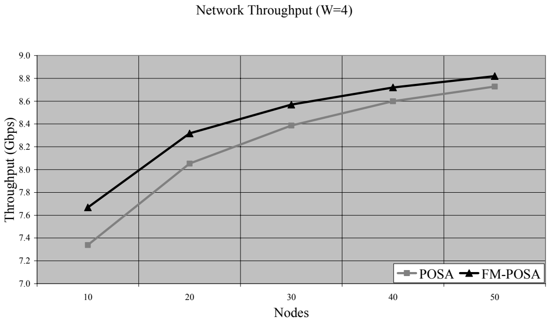

Figure 5. Network Throughput with 4 channels

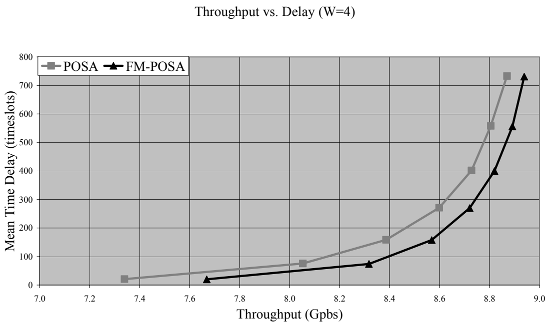

Figure 6. Throughout vs. Delay with 4 channels

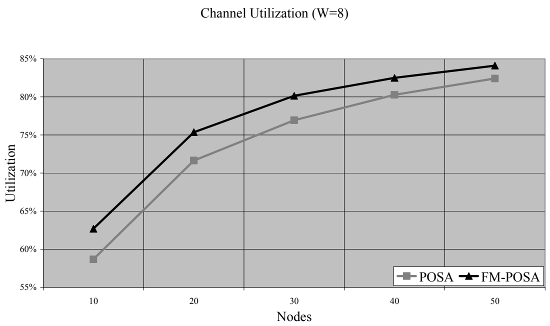

Figure 7. Channel Utilization with 8 channels

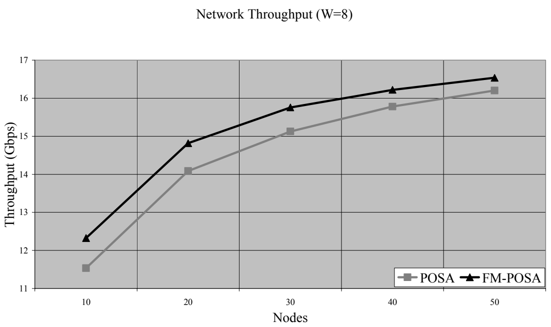

Figure 8. Network Throughput with 8 channels

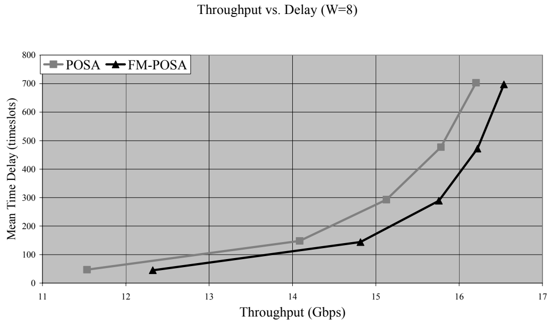

Figure 9. Throughput vs. Delay with 8 channels

V. CONCLUSIONS

In this paper we introduced a new scheduling algorithm for collision free WDM star networks. The new scheme offers a better utilization of the available channels of the network and brings an improvement in channels utilization and network throughput by changing the order of the processing of each node based on the largest request.

REFERENCES

[1] G. I. Papadimitriou, Ch. Papazoglou, and A. S. Pompotrsis,"Optical Switching : Switch Fabrics, Techniques, and Architectures", IEEE/OSA Journal of Lightwave Technology, vol. 21, no. 2, 2003, pp. 384-405.

[2] K. M. Sivalingam, K. Bogineni, and P. W. Dowd, "Pre-allocation media access control protocols for multiple access WDM photonic networks", ACM SIGCOMM Computer Communication, vol. 22, no. 4, 1992, pp. 235-246.

[3] C.A. Brackett, "Dense wavelength division multiplexing network: Principles and applications", IEEE J. Selected Areas Commun., vol. 8, 1990, pp. 948-964.

[4] G. I. Papadimitriou, P. A. Tsimoulas, M. S. Obaidat and A. S. Pomportsis: Multiwavelength Optical LANs, Wiley, 2003.

[5] K. M. Sivalingam, J. Wang, J. Wu and M. Mishra, An interval-based scheduling algorithm for optical WDM star networks, Photonic Network Communications, vol. 4, no. 1, (January 2002), pp. 73-87.

[6] E. Johnson, M. Mishra, and K. M. Sivalingam, Scheduling in optical WDM networks using hidden Markov chain based traffic prediction, Photonic Network Communications, vol. 3, no. 3, (July 2001), pp. 271-286.