A Scheduling Algorithm Minimizing Voids Generated by Arriving Bursts in Optical Burst Switched WDM Network

Masanori Iizuka, Makoto Sakuta, Yoshiyuki Nishino and Iwao Sasase
Department of Information and Computer Science, Keio University
3-14-1 Hiyoshi, Kohoku, Yokohama, 223-8522, JAPAN
E-mail: masanori@mmm-keio.net

Abstract — **We propose a scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network. In the proposed scheduling algorithm, when the burst which has arrived at optical core router at a certain time can be transmitted in some data channels by using the unused data channel capacity(it is called void), our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum. We compare the performance of our proposed scheduling algorithm with that of the conventional one with respect to burst loss ratio by computer simulations. It is shown that our proposed algorithm can improve burst loss ratio.**

I. INTRODUCTION

In recent years, due to the growth in Internet traffic, there are a lot of research activities in devising new high-speed transmission and switching technologies. Wavelength division multiplexing(WDM) technique is focused on as a core technology for realizing the next-generation IP backbone network. The bandwidth requirement of Internet traffic can be filled by using WDM technology for the IP backbone networks. However, to realize an IP backbone network with WDM technology, switching capacity of electronic routers becomes a bottleneck because processing speed of routers does not catch up with transmission capacity of an optical fiber.

A number of switching technologies are available for transport of IP traffic over WDM. The first approach is to use the traditional circuit switching, that is, when a connection setup request occurs, a lightpath is set between the ingress-egress router pair and wavelength routing is carried out in an optical domain. However, that approach is justified only if there is a large volume of traffic between the ingress-egress router pair, since wavelength is a scarce resource. In fact, the majority of resource allocation in the wavelength routing approach is coarse. The second approach is to use additional electronic layers(e.g SONET/SDH,ATM) between IP layer and optical layer. Presently, WDM is mainly deployed in the backbone of major long distance carriers as point-to-point links with a synchronous optical network(SONET) as a standard interface to higher layers in the protocol stack. That approach necessitates optical-to-electrical(O/E) and electrical-to-optical(E/O) conversions at every node, and hence, it fails to take advantage of the wavelength routing capability provided by WDM technology. The third approach is to have all-optical backbone using optical packet switching technology. This approach will provide the most flexible and efficient use of raw bandwidth available at the optical layer. However, optical buffering technology using fiber delay line(FDL) is limited in respect of random access and storage capacity, and because of synchronization requirements, optical packet switches are typically designed for fixed-size packet. On the other hand, the concept of optical burst switched(OBS) network is focused on as a new data transmission scheme to realize IP over WDM[1]-[4]. In the OBS network, several IP packets are assembled into a burst, and are forwarded through the network as one entity at optical domain. Different from electrical burst switching, the OBS uses separate wavelengths/channels to transmit the burst payload and its header. The header packet and the corresponding data burst are launched at the source at time instants separated by an offset time.

In the OBS network, when two or more data burst arrive at a same output link, output data channel contention occurs, and burst loss ratio deteriorates. In order to avoid burst loss due to output link contention, some data channel scheduling algorithms have been proposed [2],[3]. Data channel scheduling algorithms can be classified into two categories: without

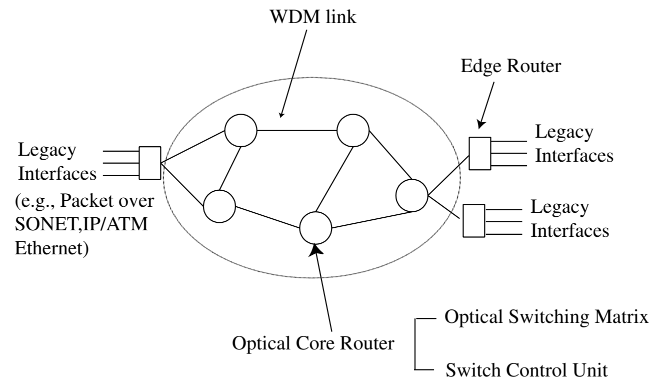


Figure 1. An OBS network.

void filling(VF)[3], and with VF[2]. When a burst which has arrived at optical core router at a certain time can be transmitted using unused channel capacity(it is called void), VF scheduling algorithm can fill voids by newly arriving data bursts. Here, void is the gap between the two data bursts which is scheduled in data channel. Among them, LAUC-VF(Latest Available Unscheduled Channel-VF) scheduling algorithm [2] can reduce burst loss ratio rather than FF-VF(First Fit VF)[2] and without VF scheduling algorithm, because LAUC-VF scheduling algorithm selects the latest available unused data channel for each arriving data burst. However, in the LAUC-VF scheduling algorithm[2], even though voids are available, the latest available unused data channel is selected for transmitting data burst. The latest available unused data channel is not necessarily the data channel in which void is available, and thus, voids are not necessarily assigned to the data burst. Therefore, the number of voids which are not assigned to the data burst increases, and data channel utilization and burst loss ratio degrade.

In this paper, we propose a scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network. In the proposed scheduling algorithm, when the burst which has arrived at optical core router at a certain time can be transmitted in some data channels by being assigned to void, our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum. Therefore our proposed scheduling algorithm can effectively utilize the data channel. We compare the performance of our proposed scheduling algorithm with that of the conventional one with respect to burst loss ratio by computer simulations. As a result, we show that our proposed scheduling algorithm can improve the burst loss ratio compared to a conventional one.

II. OBS NETWORK

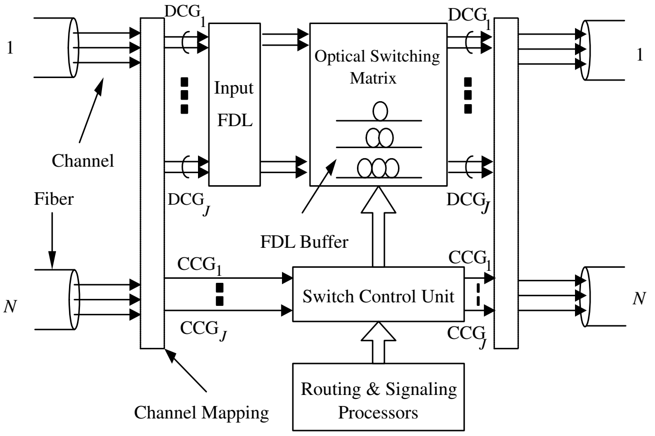


Figure 2. A structure of an $N \times N$ optical core router.

In the OBS network, in order to avoid bottleneck of electronic processing in traditional IP router, data packets having the same network egress address or some common attributes are assembled into a burst and are forwarded through the network as one entity at optical domain.

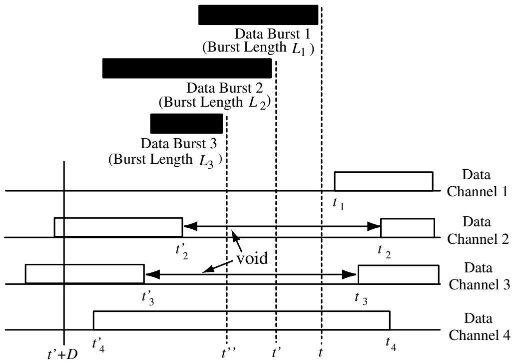
Figure 1 shows an OBS network. The network consists of optical core routers and electronic edge routers connected by WDM links. Edge routers provide burst assembly/disassembly functions and legacy interfaces(e.g, gigabit Ethernet, IP/ATM etc.). A core router is mainly composed of an optical switching matrix(OSM) and a switch control unit(SCU). A burst consists of a burst header and a burst payload. The burst payload and the burst header are called data burst and burst header packet(BHP), respectively. In the OBS network, a data burst and its header are transmitted separately on different wavelength/channels. Data channel group(DCG) is used to transmit data burst and control channel group(CCG) is used to transmit BHP, respectively. This feature facilitates the electronic processing of BHP and provides ingress-egress transparent optical paths in the OBS network for transporting data burst. Each BHP contains the necessary routing information to be used by core routers to route the associated burst in such a way that the burst simply flies through each OSM in the core routers. A BHP will be sent no later in time than the burst with a non-negative offset time τ. The offset time allows the SCU at each hop along the path to have enough time to process the BHP before its associated data burst arrives.

Figure 2 shows a structure of an $N \times N$ optical core router(OCR). OCR is composed of input FDLs, optical switching matrix(OSM), switch control unit(SCU), and routing and signaling processors. Data channels are connected to the OSM and control channel are terminated at the SCU. The input FDLs, if provided, are used to delay the arriving data burst, thus allowing the SCU to have enough time to process the associated BHPs[2]. The major function of SCU includes BHP processing, routing table look-up to determine output links of data bursts, scheduling data bursts on output channels, managing limited FDL buffers to avoid transmission conflicts, and regenerating BHPs. OSM is composed of a nonblocking spatial switch, a wavelength converter and FDL buffers. The spatial switch is able to switch a data burst from any incoming channel to any FDL as long as it does not overlap with other data bursts. The optical buffer consists of B FDLs with ith FDL being able to delay Q_i time, $1 \leq i \leq B$. For FDL=0, its delay time $Q_0 = 0$. And we assume $Q_i = i \cdot D$, where D is a given time unit of a FDL.

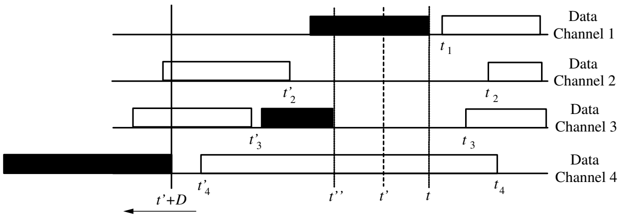
III. PROPOSED SCHEDULING ALGORITHM

In the OBS network when two or more data bursts arrive at a same output link, output link contention should be avoided using wavelength conversion and FDL buffering. Therefore, data channel scheduling is required in order to avoid output link contention. In[2], LAUC-VF scheduling algorithm has been proposed. That algorithm selects the latest available unused data channel for each arriving data burst and can fill voids by newly arriving data burst. In Figure 3(a), $t'_2 - t_2$ and $t'_3 - t_3$ are voids respectively. However, in the LAUC-VF scheduling algorithm[2], even though voids are available, the latest available unused data channel is selected for transmitting data burst. The latest available unused data channel is not necessarily the data channel in which void is available, and thus, voids are not necessarily assigned to the data burst. Therefore, the number of voids which are not assigned to the data burst increases, and data channel utilization and burst loss ratio degrade.

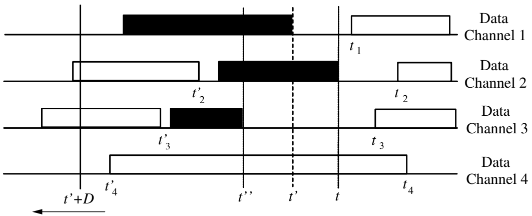
In the proposed scheduling algorithm, when the burst which has arrived at optical core router at a certain time can be transmitted in some data channels by using the void, our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum. Furthermore, when a void can not be used, the latest available unused data channel is selected to transmission of a data burst as well as LAUC-VF scheduling algorithm. Here, we explain the flow of the proposed scheduling algorithm. Given the arrival time t of a data burst with length L to the OSM, the scheduler first finds the output data channels in which voids are available for the time period of $(t, t + L)$. If there is at least one such data channel, the scheduler selects the data channel in which a void newly being generated after the burst transmission becomes minimum. When a void can not be used for the time period of $(t, t + L)$, the scheduler tries to find other data channels that are available for the time period of $(t, t + L)$. If there is at least one such data channel, the scheduler selects the latest available data channel, i.e., the channel having the smallest gap between t and the end of last data burst just before t as well as LAUC-VF scheduling algorithm. If all the data channels are ineligible at time t, the scheduler will then try to find the output data channels that are eligible at time $t + D$, and so on. If no data channels are found eligible up to time $t + B \cdot D$, the arriving data burst and the corresponding BHP are dropped. Note that $B \cdot D$ constitutes the longest time the data burst can be buffered.



(a) An example of an arrival of data bursts



(b) An example of LAUC-VF scheduling algorithm



(c) An example of proposed scheduling algorithm

Figure 3. An example of the conventional and proposed scheduling algorithm.

Figure 3 shows an example of the conventional and proposed scheduling algorithms. Figure 3(a) shows an example of an arrival of data bursts. In this figure, we give the arrival time t of data burst 1 with length L_1, the arrival time t' of data burst 2 with length L'_2,

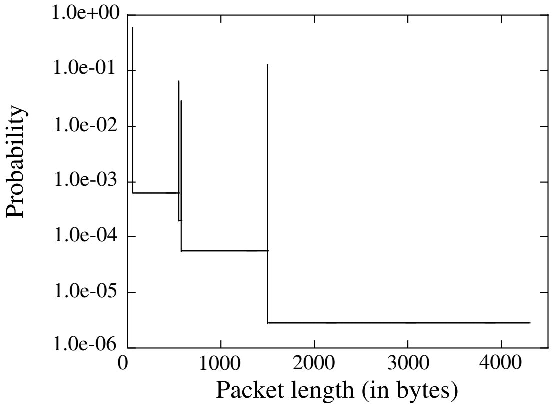


Figure 4. Probability distribution of packet length.

and the arrival time t'' of data burst 3 with length L_3', respectively. Now, we show the proposed scheduling algorithm and the conventional one, or LAUC-VF scheduling algorithm[2]. LAUC-VF scheduling algorithm shown in Figure 3(b) selects the latest available unused data channel. When the latest available unused data channel is selected to transmit data burst 1 at time t, since $t - t_1 < t - t_3 < t - t_2$, data channel 1 is selected to transmit data burst 1 as $t - t_1 < t - t_3 < t - t_2$. In that case, LAUC-VF scheduling algorithm can not select any data channels to transmit data burst 2 at time t'. Therefore, LAUC-VF scheduling algorithm needs to carry out buffering by using FDL, and data burst 2 is transmitted after time $t' + D$. In this figure, data channel 4 is selected to transmit data burst 2 as $t' + D - t_4' < t' + D - (t + L_1)$. Next, data channel 3 is selected so that data burst 3 is transmitted using void at time t''. Thus, LAUC-VF scheduling algorithm needs to carry out buffering by using FDL, and void is not effectively utilized. On the other hand, the proposed scheduling algorithm shown in Figure 3(c) selects the data channel in which voids are available from among available data channels at time t. In this figure, data channel 2 and data channel 3 can transmit data burst 1 using void at time t. Then, the scheduler selects the data channel in which a void newly being generated after the burst transmission becomes minimum. In this figure, data channel 2 is selected to transmit the data burst 1 as $t_2' - (t + L_1) < t_3' - (t + L_1)$. Next, data channel 1 is selected to transmit data burst 2 because void can not be used at time t'. And data channel 3 which can use void is selected to transmit data burst 3 at t''. Thus, when the burst can be transmitted in some data channels by using the void, our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum, and thus, it can effectively utilize the data channel.

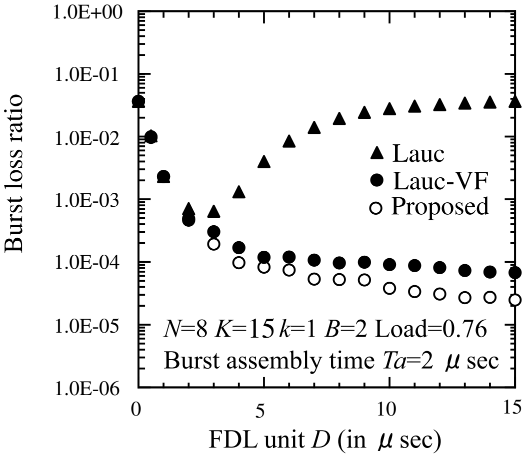


Figure 5. Burst loss ratio versus delay unit of fiber delay line. ($T_a = 2\mu$sec,load=0.76,$K = 15$,$k = 1$,$N = 8$,$B = 2$)

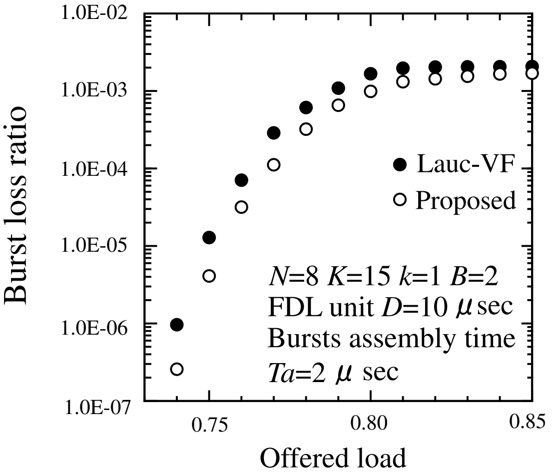


Figure 6. Burst loss ratio versus offered load to the edge router. ($T_a = 2\mu$sec,$K = 15$,$k = 1$,$N = 8$,$B = 2$,FDL unit $D = 10\mu$sec)

IV. PERFORMANCE EVALUATION

In this section, we evaluate burst loss ratio by computer simulations. Our focus here is on the first core router that connecting the edge routers. Core router size is $N \times N$ and eight edge routers are connected to the core router. Each edge router is connected to the fifteen IP routers. Assume that each link is composed of fifteen data channels and one control channel, and the transmission rate on each channel is 10Gbps. The packet stream traffic from an IP router to the edge router is random , which implies that the inter-packet time has an exponential distribution, length of each packet is independent and identically distributed. The packet length distribution used in our study is shown in Figure 4. The average packet length in Figure 4 is 389.5 bytes. Specifically, let l_p be the packet length in bytes. We have $\Pr[l_p = 44] = 0.5$, $\Pr[l_p = 552] = 0.05$,

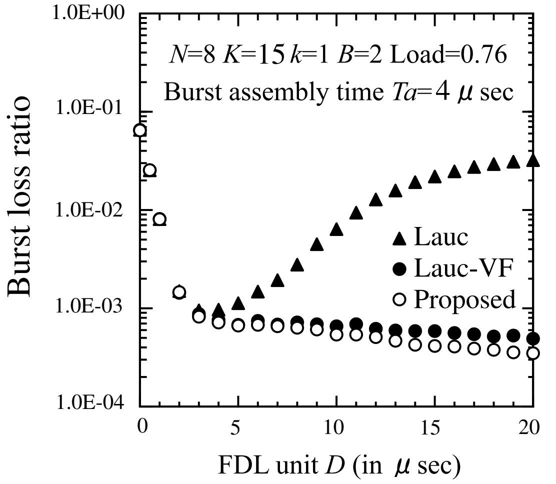


Figure 7. Burst loss ratio versus delay unit of fiber delay line. ($T_a = 2\mu$sec,load=0.76,$K = 15$,$k = 1$,$N = 8$,$B = 2$)

$\mathrm{Pr}[l_p = 576] = 0.03$, $\mathrm{Pr}[l_p = 1500] = 0.12$, $\mathrm{Pr}[45 \leq l_p \leq 551] = 0.25$, $\mathrm{Pr}[553 \leq l_p \leq 575] = 0.005$, $\mathrm{Pr}[577 \leq l_p \leq 1499] = 0.035$, and $\mathrm{Pr}[1501 \leq l_p \leq 4300] = 0.01$. In the edge router, IP packets which came from two or more IP routers are assembled into burst for every fixed interval. This interval is burst assembly time T_a in μsec. The burst length is the sum of lengths of packets collected in a given assembly time T_a. Through all results, the number of data channel $K = 15$, the number of control channel $k = 1$, the number of destination address $N = 8$ and the buffer size of FDL $B = 2$ are set[2]. And D is a time unit of FDL in μsec.

Figure 5 shows burst loss ratio versus delay unit of fiber delay line. The offered load to the edge router is 0.76, and $T_a = 2$ are set. It is shown that our proposed algorithm can improve burst loss ratio. This is because our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum, and can effectively utilize the data channel. And, it is shown that our proposed algorithm can improve burst loss ratio as delay unit D of FDL becomes large. This reason is as follows. As the increase of D, a void generated by delaying a data burst becomes large. Therefore, proposed scheduling algorithm which selects the data channel utilizing void from among some data channels becomes effective.

Figure 6 shows burst loss ratio versus offered load to the edge router. The burst assembly time $T_a = 2$, and $D = 10$ are set. It is shown that our scheduling algorithm can improve burst loss ratio. This reason is the same as Figure 6. Also, it is shown that improvement rate of the proposed algorithm decreases with the increase of offered load. This is because the number of IP packets which arrive at a edge router increases with the increase of offered load, then the burst length of a data burst which arrived at a core router becomes large, and thus the number of data burst which can be transmitted by utilizing void decreases.

Figure 7 shows burst loss ratio versus delay unit of fiber delay line. The offered load to the edge router is 0.76, and $T_a = 4$ are set. Compared with Figure 5, it is shown that the burst loss ratio of both LAUC-VF scheduling algorithm and proposed scheduling algorithm deteriorates as T_a becomes large. This is because the number of data burst which can utilize void decreases with the increase in burst length of a data burst which arrives at a core router.

V. CONCLUSION

We have proposed a scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network. Our proposed scheduling algorithm selects the data channel in which a void newly being generated after the burst transmission becomes minimum. By computer simulations, it is found that our proposed scheduling algorithm can improve burst loss ratio.

ACKNOWLEDGMENTS

This work is partly supported by Ministry of Education, the Telecommunications Advancement Foundation, and Fujitsu Laboratories.

References

[1] C.Qiao and M.Yoo, "Optical burst switching(OBS) - A new paradigm for an optical internet," *J. High Speed Networks*, vol.8, pp.69-84 ,1999.

[2] Yijun Xiong, Marc Vandenhoute and Hakki C. Cankaya, "Control Archtecutre in Optical Burst-Swtiched WDM Networks," *IEEE JSAC*, vol.18, no.10, pp.1838-1851, Oct. 2000.

[3] J.Turner, "Terabit burst switching," *J. High Speed Networks*, vol.8, pp.3-16 ,1999.

[4] Sanjeev Verma, Hemant Chaskar and Raradurgam Ravikanth, "Optical Burst Switching: A Viable Solution for Terabit IP Backbone," *IEEE Network*, pp.48-53, Nov/Dec. 2000.