

Performance Study of Various Packet Scheduling Algorithms for Variable-Packet-Length Feedback Type WDM Optical Packet Switches

Chih-How Chang, Shou-Kuo Shao, Malla R. Perati, and Jingshown Wu, *Fellow, IEEE*

Abstract—In this paper, we investigate the switching performance of feedback (FB) type wavelength division multiplexing (WDM) optical packet switches (OPSeS) with various packet scheduling algorithms that determine the buffer entrance sequence of packets arriving at the input end of the switches. Five different packet scheduling algorithms have been applied to the FB type WDM OPSeS. It is observed that OPSeS incorporated with shortest packets first (SF) and oldest packets first (OF) algorithms will yield better switching performance. Next, we observe that packet scheduling algorithms have stronger influence on switching performance when they are applied to the primary output buffers than to the re-circulated buffers. Moreover, we demonstrate that by appropriately employing the suitable packet scheduling algorithms in the primary output buffers and re-circulated buffers, the benefits of packet scheduling algorithms could be significant. Hence, the switching performance of FB type WDM OPSeS could be greatly improved simply by employing the appropriate packet scheduling algorithms.

Index Terms—Packet scheduling, feedback type WDM optical packet switch, self-similar traffic, variable packet length.

I. INTRODUCTION

WDM technology combined with optical packet switching can manage the rapidly increasing bandwidth demands on the network nodes, which is the bottleneck of current Internet traffic issues. Hence, WDM OPS is the most promising solution to the next generation Internet. On the other hand, recent measurements and observations have revealed that current network traffic tends to be bursty over many or all time scales in both Ethernet and wide area networks (WANs) and can be characterized mathematically as self-similar or long-range dependent (LRD) traffic [1], [2].

OPSeS can be classified by the way of deployment of the optical buffers, namely, one is the feed-forward (FF) type with only primary output buffers [3], and the other is the feedback (FB) type with both primary output buffers and commonly shared re-circulated buffers [4], [5]. From [6], it is known that the FF type WDM OPSeS using wavelength conversion techniques obtain good performance under short

range dependent (SRD) traffic, since the multi-server queues could interleave the packets among the available wavelength channels. However, there is a limitation on the performance improvement by employing wavelength conversion for FF type WDM OPSeS under LRD traffic [7]. The re-circulated buffers which can provide deeper buffer space and better traffic smoothing capability, can improve the switching performance further. Therefore, we mainly investigate the performance of FB type WDM OPSeS under variable-packet-length self-similar traffic [8].

When WDM OPSeS operate under self-similar variable-packet-length traffic, two different buffer control mechanisms could be adopted to reduce the packet loss probability (PLP). One is the wavelength allocation strategy which focuses on the entry order of the virtual wavelength queues [9], [10], [11]. It is well known that the wavelength allocation strategy of “choosing the virtual wavelength queue with the minimum queue length” provides the best packet loss performance among the wavelength allocation strategies, since it can utilize the benefits of WDM technology. The other mechanism is the packet scheduling algorithm which determines the buffer entry sequence of the packets. These two buffer control mechanisms can be utilized simultaneously in the OPSeS to improve the packet loss performance further. Therefore, it is worth to analyze which packet scheduling algorithm combined with the wavelength allocation strategy of “choosing the virtual wavelength queue with the minimum queue length” could obtain the optimal switching performance for each differently dimensioned FB type WDM OPSeS.

This paper is organized as follows. We first overview the architecture of FB type WDM OPSeS in Section II. In Section III, we introduce the operation of various packet scheduling algorithms. In Section IV, we analyze the performance of different packet scheduling algorithms by simulation. Finally, this paper is concluded in Section V.

II. ARCHITECTURE OF FB TYPE WDM OPSeS

The general block diagram of the FB type WDM OPSeS is depicted in Fig. 1. There are N input/output ports and R re-circulated ports with n wavelength channels in each port. When packets arrive at the switch, they will be scheduled into the suitable wavelength channels in the fiber delay line (FDL) optical buffers in order to be switched out successfully. The collided packets which cannot enter the primary output buffers will be sent to the commonly shared re-circulated buffers.

Manuscript received December 20, 2005. This work was supported in part by National Science Council, Taiwan, R.O.C. under Grant NSC 93-2213-E-002-019.

Chih-How Chang, Malla R. Perati, and Jingshown Wu are with the Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan 106, R.O.C. (phone: 886-2-23635251 ext.:519; fax:886-2-23638247; e-mail: wu-jsh@cc.ee.ntu.edu.tw)

Shou-Kuo Shao is with the Network Operation Supporting Technology Lab., Chunghwa Telecommunication Laboratories, Chunghwa Telecom Co., Ltd., Yang-Mei, Taoyuan 326, Taiwan. (e-mail: skshao@cht.com.tw)

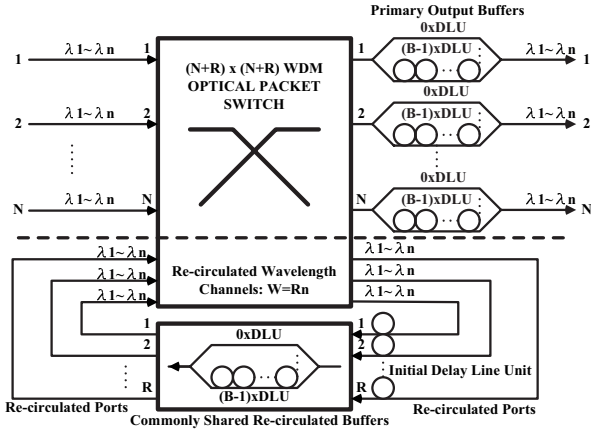


Fig. 1. General block diagram of FB type WDM OPSes.

These packets will keep re-circulating inside the switch until they can be forwarded out. We use B FDLs to construct the FDL optical buffer in a degenerate fashion [12], i.e., each FDL is incremented by one delay line unit (DLU) time, thus, B is addressed as the buffer depth. From [13], the buffer depth for any practical implementation is limited within a few tens, otherwise the switch will be too bulky to be implemented. Hence, we limit the B to be less or equal to 20 in this paper. When the packets re-circulate inside the switch, the signal of packets needs to be amplified to maintain distinguishable signal strength. Hence, optical components such as semiconductor laser amplifiers (SLAs) are required in the switch. However, the imperfection of these components will induce accumulated amplified spontaneous emission noise (ASE) [14], which degrades the signal-to-noise ratio (SNR) of the packets. Therefore, we set a parameter, re-circulation limit, b , to discard the packets whose recirculations inside the switch exceed this limit. It is found in [14] that with SLA gates and FDL buffers, the SNR of the packets are sufficiently high even after 40 circulations. Thus, it is reasonable to assume that b could be ≥ 30 . Further, the system operation slot, T , denotes the basic time unit of the operation of the WDM OPSes under slotted synchronous mode. Table I lists the symbol and the meaning of each parameter used in this paper [8].

TABLE I
SYSTEM PARAMETERS OF FB TYPE WDM OPSSES

Symbol	Meaning of Parameter
N	Number of input/output ports
n	Number of wavelength channels per port
R	Number of re-circulated ports
b	Re-circulation limit
B	Buffer depth (number of fiber delay lines)
DLU	Basic delay time unit in the FDL optical buffers (buffer granularity)
H	Hurst parameter of input traffic
L	Traffic load of input traffic

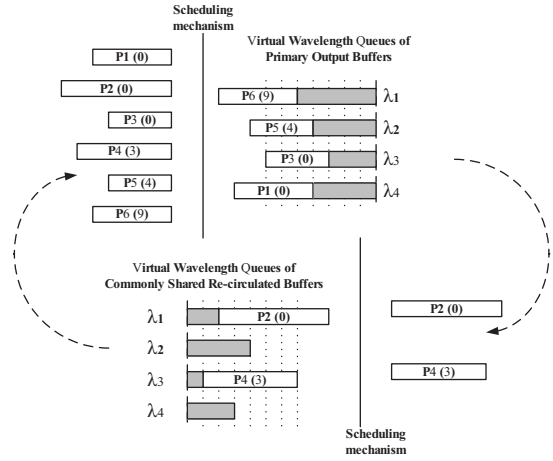


Fig. 2. The operation of “(SF, SF)” packet scheduling algorithm.

III. OPERATION OF PACKET SCHEDULING ALGORITHMS

Both the buffer control mechanisms, packet scheduling algorithms and wavelength allocation strategies, are critical issues in the design of OPSes. The wavelength allocation strategies determine the specific queue among the available virtual wavelength queues of the output ports for a packet to join in. In this paper, we employ the “choosing the virtual wavelength queue with the minimum queue length” strategy, which is known to yield the best PLP performance among the wavelength allocation strategies [9], [10]. For a packet destined to some output port, the queue with the minimum queue length will be chosen among n virtual wavelength queues of this output port. On the other hand, packet scheduling algorithms determine the buffer entry sequence of the packets. For FB type WDM OPSes, the packets can be sorted by either the lengths or re-circulation times. Thus, packet scheduling algorithms could be categorized into the following classes according to their sorting criteria: (1) when no sorting algorithm is applied, the packets enter the buffers in a round-robin fashion, which is denoted as RR algorithm; (2) sorted by packet length: shortest packet first (SF) and longest packet first (LF) algorithms; (3) sorted by re-circulation times: oldest packet first (OF) and newest packet first (NF) algorithms. Fig. 2 exemplifies the operation of packet scheduling algorithms. In this figure, the notation, “packet number (re-circulation times)”, e.g., “P4 (3)”, indicates that Packet P4 has re-circulated three times. Fig. 2 demonstrates the operation of “(SF, SF)” algorithm, which indicates SF algorithm is employed on both primary output buffers and the re-circulated buffers. At the beginning of some arbitrary time slot, six packets P1–P6 are arriving at the input of the switch. Packets P1, P2, and P3 are the newly arriving ones, and packets P4, P5, and P6 having re-circulated for 3, 4, and 9 times, respectively, are from the re-circulated buffers. Further, we assume that there are 4 virtual wavelength queues with queue lengths $5T$, $4T$, $3T$, and $4T$ in the primary output buffers. Since SF algorithm and “choosing the virtual wavelength queue with the minimum queue length” wavelength allocation strategy are applied simultaneously, the shortest Packet P3 is sent to the

shortest queue λ_3 first, and then Packets P5, P1, and P6 are entering the buffers in the same way. Since there's no space for Packets P2 and P4, they are sent to the re-circulated buffers. The shorter packet P4 enters the shortest queue λ_3 in the re-circulated buffers first, and then P2 enters queue λ_1 according to the SF algorithm.

IV. PERFORMANCE ANALYSIS – SIMULATION RESULTS

We program a discrete event simulator (DES) in C++ to perform the analysis by simulation. The self-similar traffic can be generated by aggregating many ON/OFF sources modulated with Pareto distribution. The cumulative distribution function (CDF) of Pareto distribution is $P[X < x] = 1 - (\theta/x)^\alpha$, where α represents the heaviness of the tails of Pareto distribution. When $\alpha \in (1, 2)$, the Pareto distribution has infinite variance, and the degree of the self-similarity (Hurst parameter, H) is given by $H = (3 - \alpha)/2$, thus, $H \in (0.5, 1)$ [15]. The ON periods and the OFF periods represent the variable packet lengths and inter-arrival times, respectively. The location parameter θ is set to be $1T$, which means the minimum packet length equals $1T$ in our simulations. Under slotted synchronous operation mode, the variable packet lengths are integer multiples of T . The destinations of the packets are assumed uniformly distributed among N output ports.

A. Performance of various packet scheduling algorithms applied to both primary and re-circulated buffers

In this subsection, we will first analyze the performance of various packet scheduling algorithms applied to both primary and re-circulated buffers. Figs. 3(a) and 3(b) demonstrate the PLP vs. buffer depth, B , for the half bandwidth feedback ($R = 8$) and full bandwidth feedback ($R = 16$) cases, respectively. First, we observe that OPSes employing SF and OF algorithms will obtain better performance than that of RR algorithm. The reason is that employing SF algorithm could allow more packets entering the limited buffer space, and thus the buffer space can be used more efficiently. On the other hand, OF algorithm makes the older packets enter the buffer prior to the newer packets. Thus, the older packets will have higher probability to be successfully switched out within the limited re-circulation times. It is well known that there are two causes behind the packet loss in FB type WDM OPSes: (1) lack of buffer space, and (2) exceeding the re-circulation limit. SF and OF algorithms could alleviate the said causes, respectively, hence the better PLP performance is achieved. In the case of LF algorithm, longer packets will occupy more buffer spaces and hinder the subsequent packets from accessing the buffers. Hence, most of the packets are discarded due to insufficient buffer space. On the other hand, when NF algorithm is adopted, the older packets have lower priority to enter the buffers and will exceed the re-circulation limit more easily. Thus, we could observe LF and NF algorithms lead to the worse PLP performance than RR algorithm does.

We also compare the PLP performance of FB type WDM OPSes vs. re-circulation limit, b , to understand the performance of SF and OF algorithms in Figs. 4(a) and 4(b). From these figures, we could observe that SF and OF algorithms still

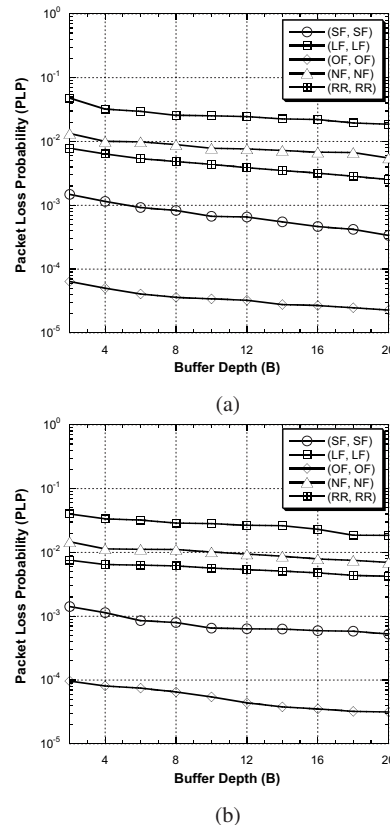
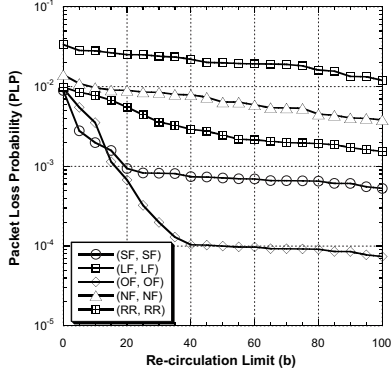
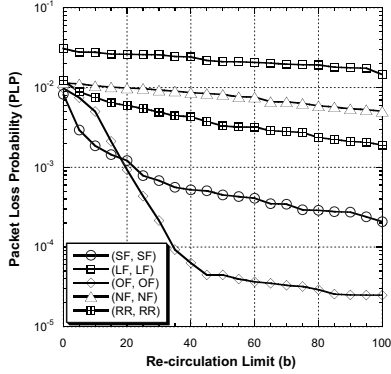


Fig. 3. PLP vs. B of 16x16 FB type WDM OPSes applying the same packet scheduling algorithms to both primary output buffers and re-circulated buffers, when $b = 30$, $DLU = 10T$, $L = 0.8$, and $H = 0.9$, (a) $R = 8$; (b) $R = 16$.

yield better PLP performance whereas LF and NF algorithms degrade the performance. However, we find that when b is small, OPSes with SF algorithm will obtain better performance than that with OF algorithm. The reason is that the advantage of OF algorithm is not eminent when the re-circulation limit is small, because the packets exceed the re-circulation limit easily and are discarded inevitably even OF algorithm is utilized. Hence, we should employ OF algorithm on FB type WDM OPSes only when b is large enough, e.g., OF algorithm outperforms SF algorithm when $b \geq 20$ in our simulation. Moreover, we could observe that when b is small, the OPSes with $R = 8$ has better performance than that with $R = 16$. This implies that when b is not large enough, i.e., the signal quality cannot be maintained through re-circulations, incrementing the re-circulated ports may lead to worse PLP performance. This is because that the traffic load seen in the primary output buffer becomes heavier as both newly arriving packets and re-circulated packets contend for the limited buffer space. Hence, the packets need more re-circulating times to be switched out. From the above observations, we could conclude that the function of the re-circulated buffers and the benefits of OF algorithm could be more significant when more re-circulation times are allowed.



(a)

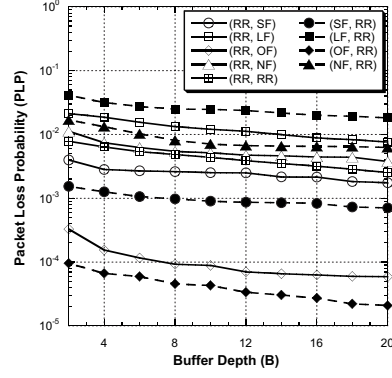


(b)

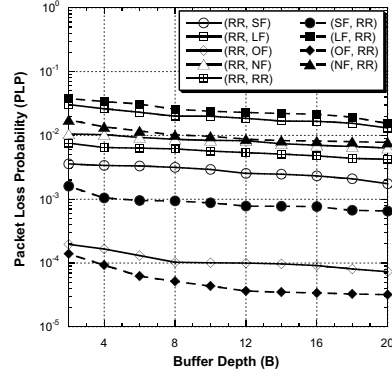
Fig. 4. PLP vs. b of 16x16 FB type WDM OPSes applying the same packet scheduling algorithms to both primary output buffers and re-circulated buffers, when $B = 10$, $DLU = 10T$, $L = 0.8$, and $H = 0.9$, (a) $R = 8$; (b) $R = 16$.

B. Performance of packet scheduling algorithms applied to primary output buffers only and to re-circulated buffers only

Next, we investigate the performance of various packet scheduling algorithms applied to the primary output buffers only and to the re-circulated buffers only. Figs. 5(a) and 5(b) demonstrate the PLP versus buffer depth when $b = 30$ in $R = 8$ and $R = 16$ cases, respectively. We compare the following two different cases: (1) applying packet scheduling algorithms only to the primary output buffers, and (2) applying packet scheduling algorithms only to the re-circulated buffers. The performance of “(RR, RR)” case is also shown there for comparison. We then observe that SF and OF algorithms still could improve the PLP performance, while LF and NF algorithms will degrade the PLP performance. In addition, we observe that FB type WDM OPSes employing “(SF, RR)” and “(OF, RR)” algorithms will have superior performance over “(RR, SF)” and “(RR, OF)” cases at both $R = 8$ and $R = 16$. Namely, employing SF and OF algorithms on the primary output buffers is more effective than applying the same ones to the re-circulated buffers. That means even if effective algorithm is employed on the re-circulated buffers, the benefits will be nullified when no strategy is applied to the primary output buffers, e.g., longer newly arriving packets may enter the buffer first when RR algorithm is used in the



(a)



(b)

Fig. 5. PLP vs. B of 16x16 FB type WDM OPSes applying packet scheduling algorithms to primary output buffers only and to re-circulated buffers only, when $b = 30$, $DLU = 10T$, $L = 0.8$, and $H = 0.9$, (a) $R = 8$; (b) $R = 16$.

primary output buffers. However, if we apply effective packet scheduling algorithm only to the primary output buffers, then the function of re-circulated buffers is merely to smooth the bursty traffic. The benefits of the SF and OF algorithms can still be eminent in the primary output buffers. On the other hand, we find that OPSes with “(LF, RR)” and “(NF, RR)” algorithms have inferior performance to that with “(RR, LF)” and “(RR, NF)” algorithms at both $R = 8$ and $R = 16$ cases. Hence, we could infer that the various packet scheduling algorithms have greater impacts on the PLP performance when they are applied to the primary output buffers.

Further, if we compare the PLP performance of Figs. 3(a) and 3(b) with Figs. 5(a) and 5(b), we observe that the performance of OPSes with “(OF, RR)” and “(SF, RR)” algorithms is close to that with “(OF, OF)” and “(SF, SF)” algorithms. Therefore, we confirm that the packet scheduling algorithms have stronger impact on PLP performance when employed on the primary output buffers. This also indicates that we can reduce the processing complexity by employing packet scheduling algorithms only on the primary output buffers and achieve the similar performance obtained by adopting the same algorithms on both primary and re-circulated buffers.

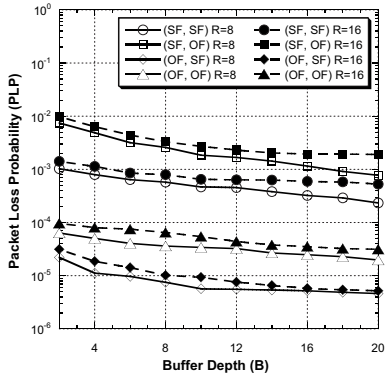


Fig. 6. PLP vs. B of 16x16 FB type WDM OPSes applying various strategies of SF and OF packet scheduling algorithms to both primary output buffers and re-circulated buffers with $R = 8$ and $R = 16$, when $b = 30$, $DLU = 10T$, $L = 0.8$, and $H = 0.9$.

C. The best strategy of applying packet scheduling algorithms

In this subsection, we are going to determine the best strategy of applying the packet scheduling algorithms by employing different packet scheduling algorithms on primary output buffers and re-circulated buffers at the same time. From the previous results, we observe that SF and OF algorithms are more efficient algorithms in terms of PLP performance, whereas SF algorithm can utilize the buffer space more efficiently, and OF algorithm can reduce the probability of exceeding the re-circulation limit. We then infer that there may be a specific strategy of applying the packet scheduling algorithms to FB type OPSes, that could lead to the optimal performance. Therefore, we demonstrate the PLP performance of different strategies of adopting SF and OF algorithms when $b = 30$ at $R = 8$ and $R = 16$ in Fig. 6. In this figure, we first observe that employing OF algorithm on the primary output buffers leads to better performance than using SF algorithm. On the other hand, SF algorithm is better when applied to the re-circulated buffers. The reason is that the influences of packet scheduling algorithms are more obvious when applied to the primary output buffers. The function of the re-circulated buffers is mainly on smoothing the traffic burstiness and providing extra spaces for storing the backlogged packets. If SF algorithm is used in the re-circulated buffers, more packets can re-circulate and then get more chances to be forwarded out. Hence, we can conclude that OPSes with “(OF, SF)” algorithm will yield the best performance.

V. CONCLUSIONS

In this paper, we have investigated the performance of various packet scheduling algorithms incorporated in FB type WDM OPSes. Five different packet scheduling algorithms, SF, LF, OF, NF, and RR algorithms, are discussed. First, we compare the individual functions of packet scheduling algorithms by employing them on both the primary output buffers and re-circulated buffers simultaneously. We observe that SF and OF algorithms could improve the switching performance as they can reduce the packet loss due to lack of buffer spaces and exceeding re-circulation limits respectively. Besides, we also

find that the improvement of SF and OF algorithms depends on the switch dimensions, i.e., OF algorithm is useful only when greater re-circulation limit is allowed. Next, we observe that the packet scheduling algorithms have stronger impact on switching performance when employed on the primary output buffers rather than on the re-circulated buffers. Since even if we employ efficient packet scheduling algorithms in the re-circulated buffers, the benefit will be nullified if no appropriate algorithm is employed on the primary output buffers. Moreover, we apply different strategies of packet scheduling algorithms to the primary output buffers and re-circulated buffers in order to find out the best strategy that could attain the best performance. Based on the results, we find that the strategy of using “(OF, SF)” algorithm, which could process the older packets to the primary output buffers first and accommodate more packets in the re-circulated buffers, will lead to the optimal performance.

REFERENCES

- [1] M. E. Crovella and A. Bestavros, “Self-similarity in world wide web traffic: Evidence and possible causes,” *IEEE/ACM Trans. Networking*, vol. 5, pp. 835–846, Dec. 1997.
- [2] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, “On the self-similar nature of ethernet traffic (extended version),” *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.
- [3] Soeren L. Danielsen, B. Mikkelsen, C. Joergensen, T. Durhuus, K. Stubkjaer, “WDM packet switch architecture and analysis of the influence of tunable wavelength converters on the performance”, *IEEE/OSA J. Lightwave Technol.*, vol. 15, pp. 219–227, Feb. 1997.
- [4] Lubo Tancevski, An Ge and Gerardo Castnon, “Optical packet switch with partially shared buffers: Design principles,” in *Proc. OFC’2001*, Anaheim, USA, Mar. 2001.
- [5] Juan Diao and Pak. L. Chu, “Analysis of partially shared buffering for WDM optical packet switching,” *IEEE/OSA J. Lightwave Technol.*, vol. 17, no. 12, pp. 2461–2469, Dec. 1999.
- [6] Soeren L. Danielsen, Peter B. Hansen and Kristian E. Stubkjaer, “Wavelength conversion in optical packet switching,” *IEEE/OSA J. Lightwave Technol.*, vol. 11, no. 12, pp. 2095–2108, Dec. 1998.
- [7] G. Shen, S. K. Bose, T. H. Cheng, C. Lu, T. Y. Chai, “Performance Study on a WDM packet switch with limited-range wavelength converters”, *IEEE Commun. Letters*, vol. 5, no. 10, pp. 432–434, Oct. 2001.
- [8] Shou-Kuo Shao, Meng-Guang Tsai, Hen-Wai Tsao, Paruvelli Sreedevi, Malla Reddy Perati, and Jingshown Wu, “Performance evaluation of feedback Type WDM optical routers under asynchronous and variable packet length self-similar traffic,” *IEICE Trans. Commun.*, vol. E88-B, Mar. 2005.
- [9] F. Callegati and W. Cerroni, “Wavelength allocation algorithms in optical buffers,” in *Proc. ICC’2001*, Helsinki, Finland, Jun. 2001, no. 1, pp. 499–503.
- [10] A. Ge, L. Tancevski, G. Castanon, and L. Tamil, “WDM fiber delay line buffer control for optical packet switching,” in *Proc. OptiComm’2000*, Dallas, Texas, Nov. 2001, pp. 247–257.
- [11] C. Develder, M. Pickavet, and P. Demeester, “Strategies for an FDL based feed-back buffer for an optical packet switch with QoS differentiation,” *Proc. COIN2002*, Cheju Island, Korea, Jul. 2002.
- [12] L. Tancevski, G. Castanon and L. Tamil, “Performance of an optical ip router using non-degenerate buffers,” *GLOBECOM’99*, pp. 1454–1459, 1999.
- [13] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen and I. Andonovic, “SLOB: A Switch with Large Optical Buffers for Packet Switching,” *IEEE/OSA J. Lightwave Technol.*, vol. 16, pp. 1725–1736, Oct. 1998.
- [14] R. Langenhorst, M. Eiselt, W. Pieper, G. Grosskopf, R. Ludwig, L. Kuller, E. Dietrich, and H. G. Weber, “Fiber loop optical buffer,” *IEEE/OSA J. Lightwave Technol.*, vol. 14, no. 3, pp. 324–335, Mar. 1996.
- [15] M. S. Taqqu, W. Willinger, and R. Sherman, “Proof of a fundamental result in self-similar traffic modeling,” *Computer Commun. Rev.* vol. 27, no. 2, pp. 5–23, Apr. 1997.