Assessment of Packet Loss for an Optical Feedback Buffer Node using Slotted Variable-Length Packets and Heavy-Tailed Traffic

Franco Callegati[†], **Davide Careglio**[§], **Walter Cerroni**[†], **Josep Solè-Pareta**[§]
[†] *University of Bologna - D.E.I.S. viale Risorgimento 2 - 40136 Bologna, Italy*
{fcallegati, wcerroni}@deis.unibo.it
[§] *Universitat Politècnica de Catalunya - Advanced Broadband Communications Lab. – CCABA*
c/ Jordi Girona 1-3, Modul D6 - 08034 Barcelona, Spain
{careglio, pareta}@ac.upc.es

ABSTRACT

The new telecommunication networks scenario is characterised by a very rapid increase of the volume of traffic, and also very variable and unpredictable traffic demands. In this context, all-optical packet switching is emerging as the most promising technology for covering such new requirements. In this paper, we deal with the performance analysis of an optical packet switch with a feedback delay line buffers configuration. The network mode operation is slotted and the case of realistic traffic model (i.e., variable-length packets modelled by self-similar traffic) fitted into trains of fixed-size slots is considered. Results from contention resolution performed by using both the time and the wavelength domains show that they must be exploited as combined tools in order to obtain good node performance.

Keywords: optical packet switching, time-wavelength exploitation, train of slots, heavy tailed traffic.

1. INTRODUCTION

Telecommunication networks currently experience the convergence of traditional voice and newer data and multimedia traffic, carried over IP and coming from Internet applications. This new scenario is characterised by a very rapid increase of the volume of traffic, and also very variable and unpredictable traffic demands.

As the electronic circuits are achieving relatively moderate advancement in speed, the optical technologies appear to be the inevitable network trend when more and more functionalities will be implemented in the optical layer. In this context, all-optical packet switching is emerging as the most promising technology for covering these new requirements [1]. An example is the ACTS KEOPS (Keys to Optical Packet Switching) project [2] that proposed a node architecture based on broadcast and select switch to carry out all-optical packet.

In this paper, we focus on the node architecture based on a space switch allowing to alter the switching configuration on a slot-by-slot basis with full wavelength conversion and optical feedback buffers (see Fig. 1). The optical packet switch node operates in a packet synchronous mode, i.e., the input interface is in charge of re-synchronising the incoming optical packets. Consequently, data packets are inserted into synchronous time slots of fixed size T.

In this work, we present the performance analysis of the optical packet switching node in terms of packet loss rate using a modern, realistic traffic model, i.e., by means of variable-length packets generated by self-similar sources to be fitted into an integer number of slots. We assume that the slots carrying information belonging to the same datagram are switched altogether as a train of packets and we call this choice Slotted Variable-Length Packet (SVLP) approach.

The remainder of the paper is organised as follows. In Section 2, we give an overview of the optical packet switching node architecture with focus on the switch operation and buffer scheduling. In Section 3, we describe the simulation scenario and we present some simulation results to assess the performance of our proposed scheme. We conclude the paper in Section 4, where we also describe future research directions.

2. OPTICAL PACKET SWITCHING NODE ARCHITECTURE

In this section we briefly describe the optical packet switching node architecture with particular attention on describing the switch operation and buffer scheduling adopted within it for contention resolution.

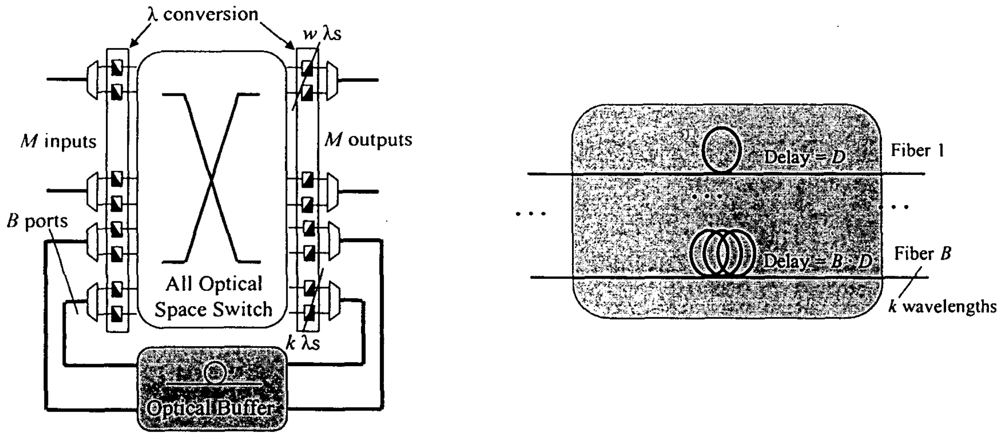
The node architecture, shown in Fig. 1, is based on all-optical space switch interconnecting M input/output fibres operated in WDM with w wavelengths per fibre. Some optical buffering capability is realised by means of B feedback ports accessing a set of fibre delay lines. The optical buffer implementation chosen is shown in Fig. 2: any re-circulating fibre has a specific length that is a multiple of the basic delay D. Therefore, by choosing the output port the switch control logic automatically selects the delay as well. Any fibre of the buffer carries k wavelengths, therefore in a given instant at most k packets may experience the same delay. The maximum delay

---

This work is partially funded by the Commission of the European Union under the COST 266 Action and the IST DAVID Project (IST-1999 11742) and by the Spanish Ministry of Education under contract CICYT TIC99-0572-C02-02.

0-7803-7375-8/02/$17.00 ©2002 IEEE

achievable is $D_M = B{\cdot}D$ that is the number of fibres times the basic delay unit. This is only one of the possible solutions to implement optical packet buffering, as discussed in [3].



(a) (b)

Figure 1. Optical packet switching node architecture (a) and scheme of the optical buffer implementation (b).

Contention resolution is necessary in the optical packet switch because of congestion due to the contemporary arrivals of trains of packets directed to the same output port. Both wavelength and time multiplexing are used to solve such a congestion. In particular, the proposed architecture makes an intensive use of wavelength multiplexing, which is used to solve contention on the output port, when several packets must be forwarded to the same output fibre. In this case as many as w packets may be transmitted at a time, multiplexed on different wavelengths. To this end wavelength converters at the input and at the output of the switching fabric are required. Furthermore, wavelength multiplexing is also used to physically share the feedback fibres that are used as a delay line buffer.

When contention resolution by wavelength multiplexing is not enough on the output ports, meaning that more than w packets are addressed to the same output fibre at a given time, time multiplexing is required. This is available by means of the delay line buffer that is used to implement a conventional output queuing with first-in-first-out scheduling. When a new train of slots arrives, according to the routing table it is forwarded to a given output fibre. Then the switch control logic is in charge of the wavelength allocation, which is made with a first availability policy: if a wavelength is immediately available, the packet is transmitted on that wavelength without sending it to the buffer, otherwise the packet is delayed until the first wavelength is available (i.e. it is sent to the shortest queue in the delay line buffer). Of course, in case the delay required is greater than D_M, the packet is dropped.

By means of such buffer policy, the feedback fibres are a fully shared buffer from the physical implementation point of view, but are used as a set of separate logical queues. This choice simplify the control logic implementation, an important issue in the perspective of limiting as much as possible the processing load due to forwarding decisions.

3. THE SLOTTED VARIABLE-LENGTH PACKET FORMAT

When coming to the packet format the choice made is to insert packets into synchronous time slots of fixed size T. The choice is motivated by a simpler design of the optical switching fabric counterbalancing the additional hardware complexity due to synchronization units [4]. Moreover, it is well known that deterministic (i.e. slotted) server queues provide best performance, an important issue in the optical environment where queuing is not easy to implement [5].

Data from legacy networks are gathered in bursts at the edges of the optical network and bursts are fitted into packets inserted into slots. Some guard bands are added between slots to let room for set up of optical devices in network nodes and the slots from the same data burst are switched altogether as a train of slots. We call this a Slotted Variable-Length Packet (SVLP) approach. As in most cases the SVLP has pros and cons with respect to various network architecture and performance issues. Let us discuss these issues.

Processing load

It is well known that header processing and forwarding decisions are critical issues in an optical packet router because of the very high speed links and the related very high packet arrival rate [6]. In a network scenario with variable length packets, this claims for a lower bound on the size of the packets in order to avoid overloads in the control logic of the switch. With fixed length packets the problem is easier to control because the maximum packet arrival rate is known "a priori", but still a lower bound on the packet size is necessary. With SVLP this limit is slightly relaxed because the processing in the optical nodes is reduced to the minimum, since the routing information is inserted only in the first slot of the train and the whole train is then processed as a whole. The mean load on the control logic of the switch is thus reduced by a factor that is roughly proportional to the average number of packets per train.

Overhead

Obviously some overhead is due to guard bands. Moreover some padding overhead is due to the mapping of variable-length bursts into multiples of a fixed-length slot. Often the client traffic is natively made of variable-length packets, as it is the case with IP, that is by far the most significant client of a backbone network, at least for what foreseeable today. IP datagrams may be transmitted on the optical network as they are or, more likely, groomed into longer bursts, but the basic data block to be transmitted is still of variable-length. In order to make such variable-length bursts fit into slots, the optical backbone network must round them up to an integer number of slots. Both guard bands and padding are basically independent of the approach used in switching the slots and not affected by the SVLP choice. On the other hand the overhead due to header information is still reduced by a factor that is roughly proportional to the average number of packets per train, since just the first slot of the train carries the header.

Queuing performance

In slotted networks in general it is well known that queuing performance are worse in the case of bursty traffic. This is an issue that has been widely studied in the ATM case [7]. The SVLP makes the traffic more bursty since a correlation is introduced among slots that have to be switched together when belonging to a train. Therefore we must expect more congestion in a network with SVLP than in a network where packets are switched independently. The longer the trains the worse the queuing performance.

From the discussion above it should be clear that T plays a major role in determining the convenience of a SVLP approach. A slot as short as possible will give the maximum advantage to SVLP in terms of saving in overhead and processing load. On the other hand this is in contrast with the requirements of optical technology that asks for packet slots significantly longer than the switching times of the optical devices, thus limiting the overhead introduced by guard bands. Moreover the shorter the slot, the longer the train size in terms of packets and the more bursty the resulting traffic profile. In this paper we will provide numerical results that try to provide some guidelines to choose the best trade-off between these issues once the basic network parameters are known.

4. SIMULATION RESULTS

In order to assess the performance of the optical packet switch with the SVLP approach, an ad-hoc event-driven simulator has been built and tested. As for the traffic model used to generate the bursts incoming from legacy networks, each input wavelength is modelled with 32 multiplexed point arrival processes having Pareto distributed interarrival times that simulated self-similar sources, as illustrated in [8]. For the burst size we consider an exponential distribution with average burst length L. Each burst is then inserted into a train of slots. The simulation scenario presented in Table 1 has been adopted.

Table 1. Network scenario parameters

Parameter	Symbol	Value
No. of I/O fibres	M	6
No. of wavelengths per fibre	w	32
Average load per wavelength	ρ	0.8
Average burst length	L	500 bytes
Slot size normalised to L	T	0.1, 0.2
Buffer basic delay unit normalised to L	D	T, $2T$

The slot size has been chosen to be 10%-20% of the average burst length in order to keep the overhead due to padding limited. With reference to the asynchronous variable-length case, it has been shown that the ratio between the buffer delay unit D normalised to the average burst length L is crucial to determine the performance of the buffer [9]. In particular, there is an optimal choice for D, which should neither be too small nor too large in order to get the best trade-off between time resolution and queuing capacity of the delay buffer. A similar relation also holds for the case of SVLP. In particular it is required that D is a multiple of T. This choice reduces as much as possible the gaps that may happen among queued packets, as explained in [9]. In studies about delay line optical buffering with fixed length packets, it is usually assumed that the best choice of the delay line time unit is equal to the length of packet. This may not be the case with SVLP and, for this reason, in this paper simulations have been performed for D equal to T and $2T$.

According to the optical buffer architecture considered in section 2, being the delay line buffer a shared buffer, at least from the physical resources point of view, the packet loss probability is due to a twofold reason: lack of delays and lack of wavelengths inside the delay lines.

The dependence of performance on the number of delay lines is studied in Fig. 2. Here the curves show the packet loss probability as a function of the number of fibre delay lines in the optical buffer for different slot and delay unit sizes normalised to the average burst length L. The number of wavelengths within the buffer is set to infinite, in order to characterise the losses due only to the lack of delays. The figure shows that choosing $D=2T$ gives better performance than $D=T$, at least for the slot sizes considered. This result means that the choice $D=T$ is not necessary the best one. The matching between T and D is not effective in the regions where the packet loss probability is very bad because of the poor performance of the buffer with respect to the variable size of the trains, i.e. where D is too small.

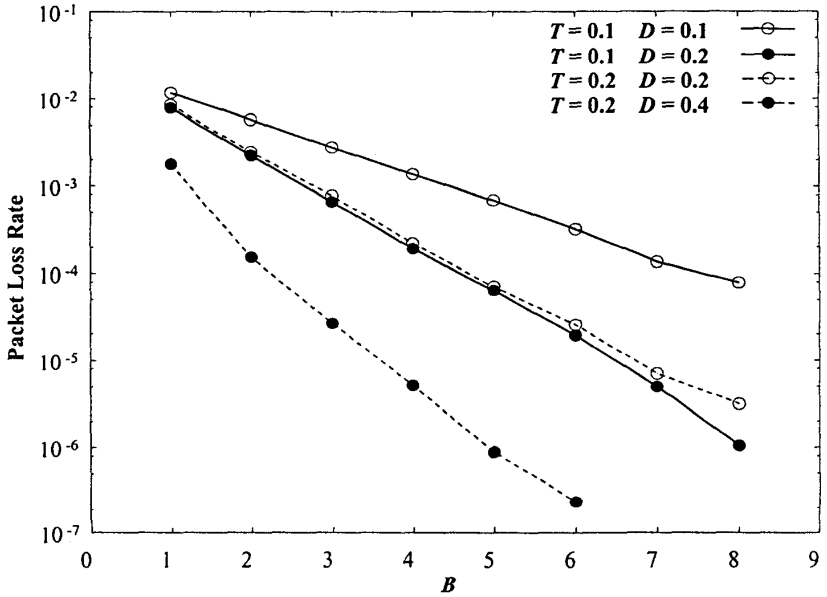


Figure 2. Packet loss rate vs. buffer depth for different slot and delay unit sizes, with infinite wavelengths per fibre delay line.

In Fig. 3 we analyse the role of k, number of wavelengths in the $B=4$ buffer fibres for different slot and delay unit sizes normalised to the average burst length L. The figure clearly shows that for small values of k the packet loss probability is mainly due to lack of wavelengths. Typically more than k packets require a given delay at a time and some are lost because all the wavelengths within that fibre delay line are busy. As long as k is large enough the packet loss probability flattens and does not change anymore, which means that in this case the loss is due to lack of delays. This figure is very important to understand that packet loss is the consequence of congestion on wavelengths or delays and that the buffer must be dimensioned taking into account both such causes to realize a suitable trade-off.

Finally, Fig. 4 shows the switch performance as a function of B and k with the constraint $B \times k = 256$. It is worth to notice that the buffer length versus number of wavelengths trade-off shows an optimum in terms of packet losses for $B = 4$ and $k = 64$. This means that wavelength multiplexing plays a major role compared to the contention resolution in the time domain. In other words, it is more effective to have a few delays contemporary available to a lot of packets than having a huge number of delays with a small number of wavelengths. The reason is that sharing several wavelengths on the output links relieves the long term congestion, but requires more resources to solve the short term congestion that still arises because of contemporary arrivals of trains.

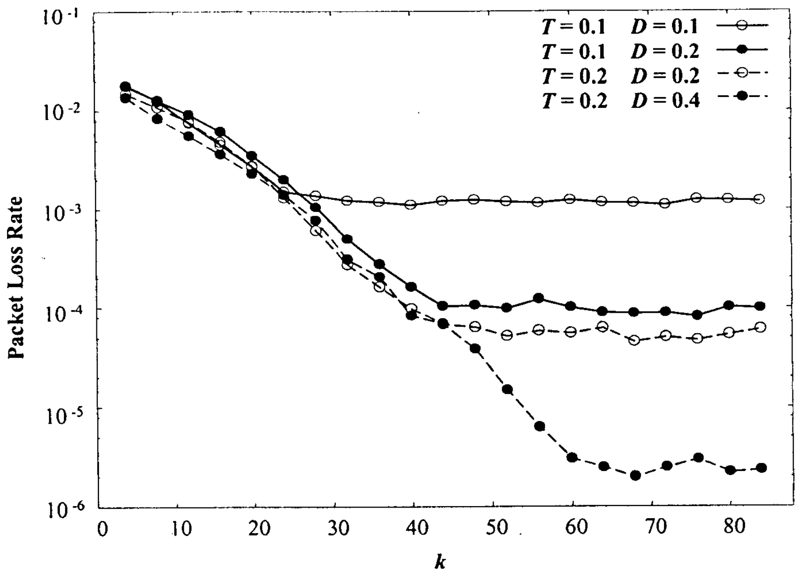


Figure 3. Packet loss rate vs. number of wavelengths per delay line, for different slot and delay unit sizes with B=4.

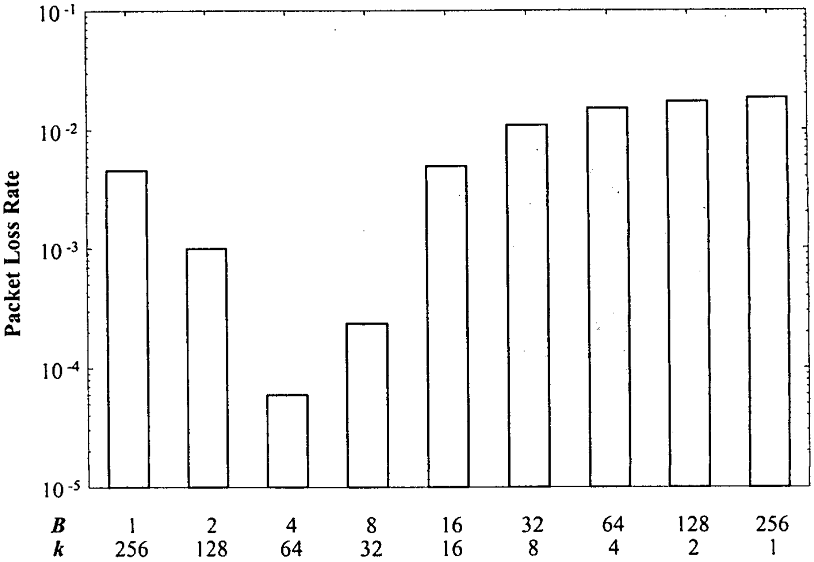


Figure 4. Packet loss rate vs. different choices for the optical buffer resources, with the constraint $B \times k = 256$.

In conclusion the results presented show that, with a given input traffic profile, a time slot T as large as possible is preferable in a performance perspective, because of the related reduction of the average length of the train. This must be counterbalanced with the overhead due to padding, that is going to increase with the size of the slot. The padding overhead has not been considered in the results here presented and a detailed evaluation of its influence on performance is the subject of further studies. Moreover the results show the importance of the role played by WDM in determining the performance of the feedback delay line buffer. In particular it is shown that, despite what could be intuitively imagined, an increase in the number of delays (meaning size of the buffer) is not always the best choice for performance improvement. As shown in Fig. 4 in order to be effective the feedback buffer must be dimensioned with a correct coupling of number of delay fibres and amount of wavelength multiplexing.

5. CONCLUSIONS

In this paper we presented the performance analysis of an optical packet router with feedback delay line buffer subject to slotted variable-length packets. We have analysed the node performance as a function of the buffer depth and degree of wavelength multiplexing. Simulation results suggest that both the time and wavelength domains are essential to solve congestion and reach acceptable performance when buffering trains of slots with delay lines. These resources can be used as combined tools to reduce the packet loss and, at the same time, to exploit the optical technology features. As shown, increasing the number of wavelengths per fibre delay line reduces packet loss up to a certain extent when the buffer length becomes the system bottleneck. Such bottleneck has to be solved by increasing the buffer size or by increasing the number of wavelength per input/output fibre, thus allowing a higher degree of wavelength multiplexing.

ACKNOWLEDGEMENTS

The authors wish to thank all the partners of the DAVID project for their contributions to the discussion on the subject of this paper.

REFERENCES

[1] Franco Callegati, Maurizio Casoni and Carla Raffaelli, "Packet Optical Networks for High-Speed TCP-IP Backbones", IEEE Communications Magazine, vol. 37, n. 1, pp 124-129, January 1999.

[2] C. Guillemot et al., "Transparent Optical Packet Switching: the European ACTS KEOPS project approach", IEEE/OSA Journal of Lightwave Technologies, vol. 16, n. 12 pp. 2117-34, December 1998.

[3] D.K.Hunter, M.C.Chia, and I.Andonovic, "Buffering in optical packet switches", IEEE/OSA Journal of Lightwave Technology, vol. 16, n.12, pp.2081-2094, December 1998.

[4] P. Gambini et al., "Transparent optical packet switching: network architecture and demonstrators in the KEOPS project", IEEE Journal on Selected Areas in Communications, Invited paper, Vol. 16, No. 7, pp. 1245-1259, September 1998.

[5] P.B. Hansen, S.L. Danielsen, and K.E. Stubkjaer, "Optical packet switching without packet alignment", Proc. of ECOC '98, Madrid, Spain, September 1998.

[6] F. Callegati, H.C. Cankaya, Y. Xiong, and M. Vandenhoute, "Design issues of optical IP routers for Internet backbone applications", IEEE Communications Magazine, Vol. 37, No. 12, pp. 124-128, December 1999.

[7] A. Baiocchi, N. Blefari Melazzi, and A. Roveri, "Buffer dimensioning criteria for an ATM multiplexer loaded with homogeneous on-off sources", Queuing performance and control (ITC 13), Elsevier Science Publisher, 1991.

[8] J. J. Gordon, "Long range correlation in multiplexed pareto traffic", Proceedings of the International IFIP-IEEE Conference on Broadband Communications, pp. 28-39, Montreal, Canada, April 1996.

[9] F. Callegati, "Optical buffers for variable length packets", IEEE Communications Letters, Vol. 4, No. 9, pp. 292-294, September 2000.