

DYNAMIC DWDM EXPLOITATION IN CONNECTION-ORIENTED OPTICAL PACKET SWITCHES

Franco Callegati, Walter Cerroni, Carla Raffaelli, Paolo Zaffoni
D.E.I.S. - University of Bologna, Viale Risorgimento 2- 40136 Bologna – Italy
{fcallegati, wcerroni, craffaelli, pzaffoni}@deis.unibo.it

Abstract: This paper deals with an MPLS over DWDM network scenario and with the related problem of Label Switched Path (LSP) allocation to wavelengths. Static and dynamic allocation strategies within an optical packet switch are considered and compared. In particular dynamic allocation algorithms can be used to improve the overall system performance by exploiting queuing behaviors related to the connection-oriented nature of the traffic. As a consequence packet loss probability is obtained in an acceptable range with very limited queuing resources as required by implementation limits.

Key words: MPLS, DWDM, Optical Packet Switching, Optical Buffer, Dynamic Wavelength Assignment

1. INTRODUCTION

Optical packet networks have been recently considered as a powerful solution to tackle the rapidly increasing bandwidth demand of the global Internet [1][2][3]. At the same time the introduction of the Multi-Protocol Label Switching (MPLS) paradigm in TCP-IP based networks promises for effective network management and traffic engineering in the future [4]. On the transmission side the recent availability of the Dense Wavelength Division Multiplexing (DWDM) technology provides the platform to exploit the huge capacity of optical fibers. So the integration of MPLS with the all-optical networks is a widely discussed issue, and proposals such as, for instance, MPλS and Generalized MPLS (GMPLS) are emerging [5].

Assuming the availability of all-optical packet switching technology, this paper focuses on the problem of integration of MPLS and DWDM.

Most of the work developed on optical packet switching is related to the case of fixed-length synchronous packets, that provides easier switching matrix design but do not easily match with IP. For this reason more recently the case of asynchronous, variable-length packets has also been studied [6] [7], showing that several new problems arise but also that acceptable performance may be achieved. The work here presented considers the case of MPLS traffic and therefore assumes variable-length packets.

In the case of all-optical packet switching, congestion resolution may be achieved in the time domain by means of queuing and in the wavelength domain by means of suitable wavelength multiplexing.

Queuing is achieved by delay lines (coils of fibers) that are used to delay packets in relation to output queue occupancy then retrieved for transmission at the right time. In particular, buffering variable-length packets with delay lines is a very critical task and the performance depends strongly on the so-called buffer time scale [8]. This is the unit of delay introduced by the delay lines that, depending on their length, delay a packet of a multiple of such unit. In particular the ratio between delay unit and average packet length is a key parameter and for this reason most of the results presented in the following are plotted against this quantity.

This work extends these ideas in the case of a connection-oriented network, such as an MPLS network, and provides some insights into the problem related to the application of dynamic wavelength allocation to a connection-oriented environment. A previous work on dynamic wavelength assignment was developed for fixed-length, synchronous optical packets [9].

The paper is structured as follows. In section 2 the MPLS over DWDM scenario is described and the suitability of the MPLS concept for optical packet networks is outlined. In section 3 exploitation of Fiber Delay Line (FDL) buffers with DWDM links is discussed in a connectionless environment. In section 4 the algorithms for dynamic Label Switched Paths (LSPs) mapping into DWDM buffers are introduced and their performance presented in section 5. Conclusions of the work are presented in section 6.

2. MPLS OVER DWDM SCENARIO

MPLS is a connection-oriented protocol setting up unidirectional Label Switched Paths (LSPs) identified by an additional label added to the IP datagrams [4]. With MPLS the network layer functions are partitioned into two basic components: *control* and *forwarding*. The control component uses standard routing protocols to build up and maintain the forwarding table.

The forwarding component examines the headers of incoming packets and takes the forwarding decisions. Packets are classified into a finite number of subsets, called Forwarding Equivalence Classes (FECs). Packets belonging to the same FEC are, from a forwarding point of view, indistinguishable and are forwarded from source to destination along the same LSP.

MPLS brings concepts such as traffic engineering, explicit routing, alternative path routing, quality of service management etc. into the IP-layer. In present proposals for optical networking, such as MPLS, the LSPs are mapped into wavelengths, in order to realize end-to-end wavelength switched paths. These may not be the most efficient solutions for multiplexing, since the incoming flow of information is bursty by nature. To achieve maximum flexibility in terms of bandwidth allocation and sharing of capacity the challenge is to combine the advantages of DWDM with emerging all-optical packet switching capabilities to yield optical routers able to support label switching of packets.

Optical routers require a further partitioning of the forwarding component into *forwarding algorithm* and *switching*. The former is the routing tables look up to determine the next hop destination and the latter is the physical action of transferring a datagram to the proper output interface. The main goal of this separation is to limit the electro-optical conversion to the minimum to achieve better interfacing with optical WDM transmission systems:

- the header is converted from optical to electrical and the execution of the forwarding algorithm is performed in electronics;
- the datagram payload is optically switched without conversion to the electrical domain.

This approach suits very well the MPLS paradigm and exploits the best of both electronic and optical technology. Electronics is used for processing matters in the routing component and forwarding algorithm, while optics is used in switching and transmission where high data rates are required.

This paper focuses on this scenario, assuming the availability of an all-optical switching matrix able to switch variable-length packets, for instance like the one described in [7]. The problem of congestion due to temporary overloads is addressed by means of queuing in a fiber delay line buffer and by means of wavelength multiplexing.

3. BUFFERING WITH FIBER DELAY LINES AND DWDM

The behavior of the FDL buffers with variable-length, asynchronous packets has been analyzed in detail in previous papers such as [8]. The main

problem is that, in the assumption of delay lines that are multiple of the same delay unit, delays are available in a discrete set (generally small) while the queuing of asynchronous packets requires delay values from a continuous set, that depends on the arrival pattern and packet lengths.

The resulting effect is that the FDL buffer creates artificial gaps between packets, resulting in an under-utilization of the output links. This phenomenon can be considered equivalent to an artificial increase in the average packet length and, as such, it has been described and analyzed in [8]. One of the main conclusions of the analysis is that the performance of the buffer is very sensitive to the value of the delay unit and that, in the case of single wavelength fibers, many delays are required to reach acceptable performance.

In a DWDM environment both time and wavelength multiplexing can be used to solve congestion. With the term wavelength multiplexing we address the strategy used to choose the wavelength to transmit the packets on the DWDM links. The alternatives are mainly the following ones [10]:

- *Wavelength Circuit* (WC) in which the elementary path within the network for a given LSP is designated by the wavelength, therefore packets belonging to the same MPLS connection will all be transmitted on the same wavelength;
- *Wavelength Packet* (WP) in which the wavelengths can be used as a shared resource, the traffic load is spread over the whole wavelength set on an availability basis and packets belonging to same MPLS connection can be spread over more wavelengths.

Wavelength Circuit is easier to manage, while Wavelength Packet is more complex to control in general but provides higher performance.

In the case of a connectionless packet switched network, buffering with delay lines and WP multiplexing may be used as combined tools. When a packet arrives the forwarding algorithm, by means of header processing, decides the packet destination in terms of output fiber. Then, since all the wavelengths on that fiber are equivalent for forwarding purposes, a wavelength selection algorithm chooses the wavelength on that fiber to send the packet on. Several wavelength selection algorithms may be designed. The *void filling* algorithm presented in [11] is probably the most performing even though very demanding from the computational point of view. According to its name, void filling aims at filling all the gaps in the delay line buffer. Two simpler algorithms have been proposed in [12] called *minimum length* and *minimum gap*. They implement FIFO queuing, and are simpler since the only information needed is the time when the various wavelengths on the output fiber will be free. The minimum length algorithm chooses the first wavelength that will be available while the minimum gap looks for the wavelength where the queuing of the packet will cause the

smallest gap. A very simple example of what wavelength selection means is presented in Fig. 1 for the case of 4 delays and 4 wavelengths per fiber, where D is the delay unit. The example makes clear what is the meaning of discrete delays. When a new packet arrives at time t_0 and finds all wavelengths busy it can be delayed through delay line 1 to time $t_0 + D$, or through delay line 2 to time $t_0 + 2D$, etc. up to $t_0 + 4D$. In this example the minimum length algorithm will delay the packet of $2D$ and send it to λ_4 , that is the first wavelength to become available for transmission. The minimum gap algorithm will delay the packet of $3D$ and send it to λ_1 where the gap between the previous and the present packet is minimized.

From this example intuition may suggest that the minimum length algorithm is the best. In Fig. 2 results provided in [12] prove that this is not the case and the minimum gap algorithm actually performs better both in terms of packet loss probability and in terms of queuing delay.

The improvement in packet loss probability that can be achieved with WP over WC may be of several orders of magnitudes (depending on the algorithm chosen). Therefore the conclusion is that, in a connectionless environment, the combination of time and wavelength multiplexing is a very powerful congestion resolution strategy. The results presented in literature show that good performance is possible with a very limited number of delay lines (in the order of the tenth), if the number of wavelength per fiber is in the order of 16 or 32. These numbers are compatible with foreseen optical technology for this kind of applications.

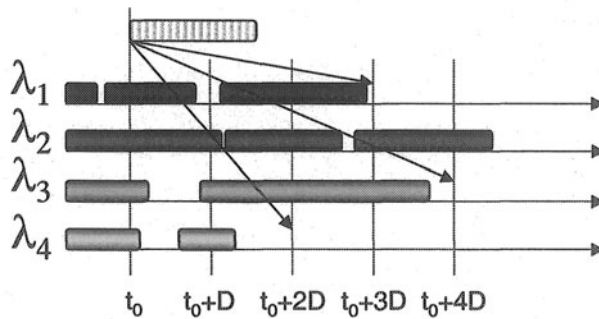


Figure 1. Example of wavelength selection in the case of Wavelength Packet multiplexing for 4 wavelengths and 4 delays, multiple of the unit D .

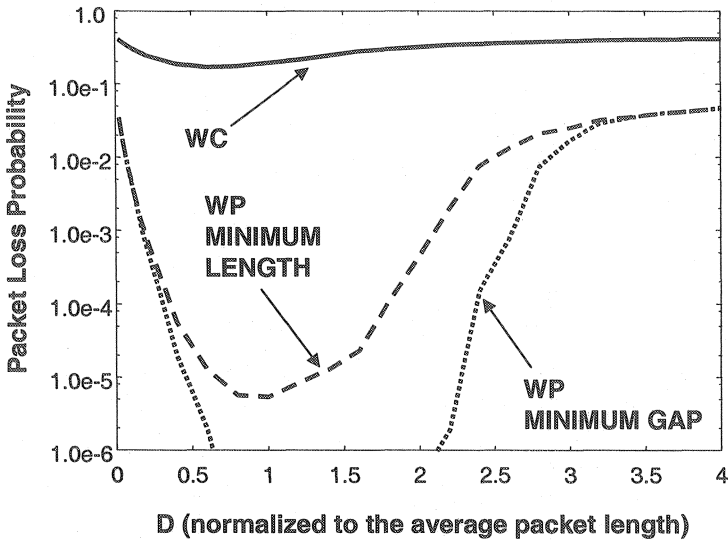


Figure 2. Packet loss probability as a function of the delay unit D for a 4 fibers \times 4 fibers switching matrix with 8 delay lines, 16 wavelengths per fiber, uniform traffic and load 0.8.

4. DYNAMIC LSP TO WAVELENGTH ASSIGNMENT

In a connection-oriented scenario, such as MPLS, a WC approach should be preferable, but then the performance issue arises. A low packet loss probability can be achieved only with a large number of delay lines. Therefore WP is necessary, but taking into account the existence of the LSPs. Research in this area has been started very recently. The idea is to design dynamic algorithms that, while following the same principles of those previously explained, exploit the connection-oriented nature of the MPLS traffic.

In this paper we compare the performance of three algorithms:

- Round Robin Wavelength Selection (RRWS) algorithm;
- Grouping Wavelength Selection (GWS) algorithm;
- Empty Queue Wavelength Selection (EQWS) algorithm.

The algorithms are designed and evaluated taking into account two performance measures:

1. congestion resolution capability;
2. forwarding transparency.

Defining the *congestion* as the condition in which a delay line queue is full and a packet addressed to that queue would be lost, the effectiveness of the algorithm with respect to congestion resolution is the ability to reduce the occurrence of these events, and therefore reduce the overall packet loss probability.

In a WP scenario, the hopping of packets belonging to the same LSP from wavelength to wavelength will cause additional updates of the forwarding table and possible out of order arrivals. The former issue means an overload for the control function of the optical packet switch and the latter more complex interfaces at the edges of the optical network for re-sequencing. A wavelength allocation algorithm that limits these phenomena as much as possible is more transparent with respect to the forwarding actions of the optical packet switches. For this reason we call *forwarding transparency* this measure of performance.

The RRWS is the simplest algorithm that could be imagined and was originally proposed in [13]. When a packet belonging to an LSP arrives and has to be delivered to a wavelength that is congested, the algorithm starts a round robin search for another not congested wavelength of the same fiber. If such a wavelength is found, the packet is sent and the forwarding table is updated to keep the same destination for all subsequent packets belonging to the same LSP.

This algorithm is very simple to implement but not very intelligent. In particular the choice of the new wavelength is driven by the round robin search and therefore cannot be optimal. Nevertheless it was shown in [13] that it is able to achieve a non negligible improvement in performance with respect to a WC approach. The price to pay is a large number of forwarding table updates, meaning a more complex switching control and many packets out of sequence at the edges of the optical network.

While analyzing the behavior of the RRWS algorithm, an effect related to the connection-oriented nature of the traffic became evident. Fig. 3 shows an example with two wavelengths per fiber, that refers to the input of a switching matrix. On wavelength λ_2 three LSPs labeled L1, L2 and L3 are active, while only LSP labeled L4 is active on λ_1 . By observing the figure it is trivial to understand that, whatever the output destination, packets from L1, L2 and L3 will never overlap, while they may overlap with packets from L4. Since congestion in the output queues is due to overlapping arrivals, only packets from LSPs incoming from different input wavelengths will cause congestion. As a consequence if L1, L2 and L3 are the only LSPs forwarded to the same output wavelength, congestion will never arise on that wavelength. This is graphically explained in Fig. 4, where not only the input, but also the output wavelengths, with related queues, are plotted. Here the queue to access wavelength λ_2 of output fiber 2 is always empty and,

obviously, no packets from L1, L2 or L3 are lost. This situation will be called *optimal allocation* in the following. The figures focus on the optimal allocation case, but it is obvious that intermediate configurations are also possible. The example shows that, in the connection-oriented environment, the performance depends on the forwarding table and may vary in time following the changes in the forwarding table. This is because the configuration of the forwarding table influences the behavior of the output queues.

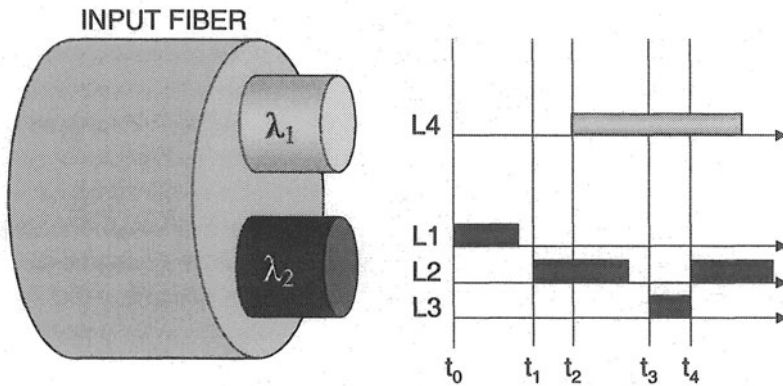


Figure 3. Packets belonging to different LSPs incoming from the same input wavelength do not overlap in time because of the serial nature of the transmission line.

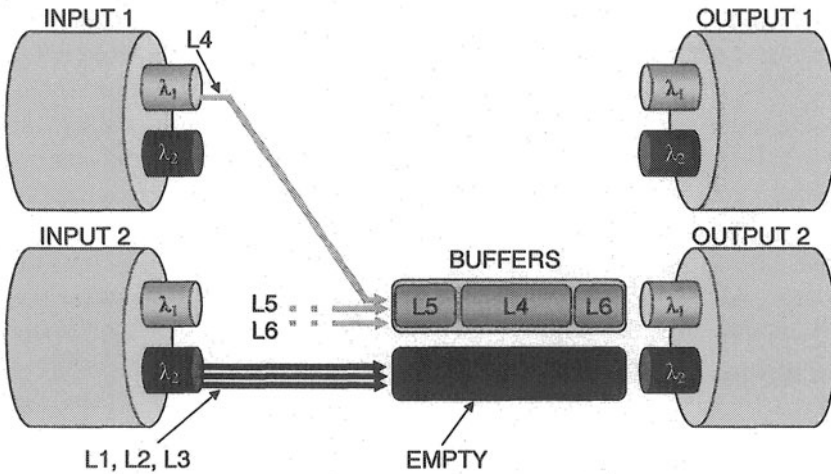


Figure 4. Example of optimal allocation; the delay line queue for λ_2 is never used and always empty because arrivals never overlap.

In [14] a method was defined that can be used to quantify the amount of LSPs which are grouped at the input and at the output. Such method was used to define the GWS algorithm, designed with the aim of maximizing the number of LSPs from the same input wavelength forwarded to the same output wavelength. In practice when a packet belonging to an LSP arrives and has to be delivered to a wavelength that is congested, the algorithm starts a search for another not congested wavelength of the same fiber. The new wavelength is chosen in order to maximize the number of LSPs that come from the same input wavelength and are forwarded to the same output wavelength. In case of optimal allocation, like the one of λ_2 in Fig. 4, the wavelength is considered unavailable for dynamic allocation and congestion will never occur on that wavelength. The GWS does not improve significantly the packet loss probability with respect to the RRWS algorithm but is able to reduce the number of forwarding table updates of about one half. This is because it achieves an overall configuration of the forwarding table which is better, from the performance point of view, than the one achieved by RRWS. A better configuration of the forwarding table means less congestion events and therefore less dynamic re-allocation of LSPs.

The EQWS algorithm is proposed for the first time in this paper and follows a strategy that is opposite to that of the GWS algorithm. GWS avoids to allocate an LSP to a wavelength in case of optimal allocation, it prevents congestion to arise, but also wastes all the queuing space devoted to the wavelengths in this special condition. On the opposite EQWS algorithm aims at a full exploitation of the queuing space available in the switching matrix and behaves as follows: when a packet belonging to an LSP arrives and has to be delivered to a wavelength that is congested, the algorithm searches for a wavelength in optimal allocation, i.e. an empty queue. The LSP is assigned to that wavelength as long as needed, that is until congestion arises also in the new wavelength or congestion disappears from the previous one. In this case the LSP is switched back to the original wavelength. An example of the behavior of the EQWS is plotted in Fig. 5, where L4 is switched to λ_2 which is in the optimal allocation state. Because of L4, overlapping of packets forwarded to λ_2 is now possible and some queuing may be necessary. It may also happen that λ_2 becomes congested and the queue fills up as in Fig. 5. In this case L4 is switched back to the original wavelength and the queue at λ_2 will progressively become empty again.

The EQWS aims at a full exploitation of the queuing space available, following the idea that, if some queuing space is not used, it should be used in the case of congestion.

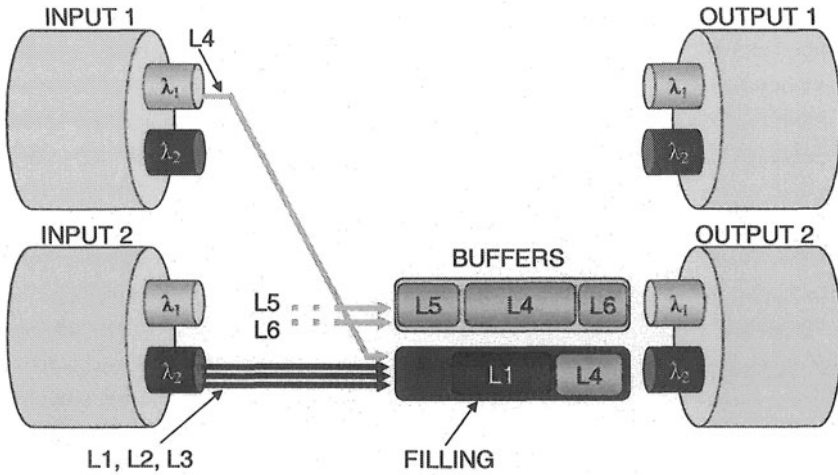


Figure 5. The queue for λ_2 fills up because packets from L4 may overlap with packets from L1, L2 and L3.

5. NUMERICAL RESULTS

In this section numerical results are shown in order to measure and compare the performance of the algorithms presented in section 4. The measures have been obtained by means of an ad-hoc, event-driven node simulator, implementing the behavior of a 4×4 switch with 16 wavelengths per fiber, which results in a 64×64 optical switching matrix. Each input wavelength is supposed to carry 3 different LSPs randomly directed to the output fibers, for a total of 192 incoming LSPs.

In particular for each input fiber, one wavelength carries LSPs which are directed to the same output, 9 wavelengths transport only 2 LSPs directed to the same output (the third LSP being directed to a different one), while the remaining 6 wavelengths have the 3 LSPs directed to 3 different outputs. As described in [14] such a distribution of the outputs is the average distribution in case of uniform traffic, i.e. when no hot-spots are considered and all the outputs have the same probability to be addressed, as supposed here. As for the initial forwarding table, the LSPs are assigned to the output wavelengths by keeping as much optimal allocations as possible.

The input traffic is random, generated by a Poisson process and the packet size distribution is exponential with average and minimum lengths of 500 and 32 bytes respectively. The input load on each wavelength is 0.8, equally divided among the LSPs. Performance evaluation is carried out with reference to the packet loss probability (PLP), being the delay a less

interesting parameter due to the small size of the buffer adopted (only a few delay lines).

The more efficient utilization of the buffering space available in the switch by EQWS dramatically improves the performance, as shown in Fig. 6 where the PLP produced by the three algorithms is plotted as a function of the buffer delay unit D with 8 delay lines. In the figure RRWS and GWS show the same performance according to [14], while EQWS reduces the PLP by more than three orders of magnitude.

In order to have a benchmark to compare such an improvement to, we consider as a reference case a pure connectionless environment where the wavelength assignment is performed packet by packet, as explained in section 3. In Fig. 7 the PLP versus the maximum available delay is depicted for RRWS and EQWS in the connection-oriented (CO) case and the minimum length and minimum gap algorithms for the connectionless (CL) case. The value of the buffer delay unit is chosen for each algorithm with relation to the minimum of the PLP. The curves show that the buffer utilization by CO EQWS is comparable to that by CL minimum length, i.e. it is a very good utilization trying to exploit all the buffering space available in case of congestion. In other words, both in CO and CL environments, the average amount of inactivity gap between consequent buffered packets will be the same, and so the performance.

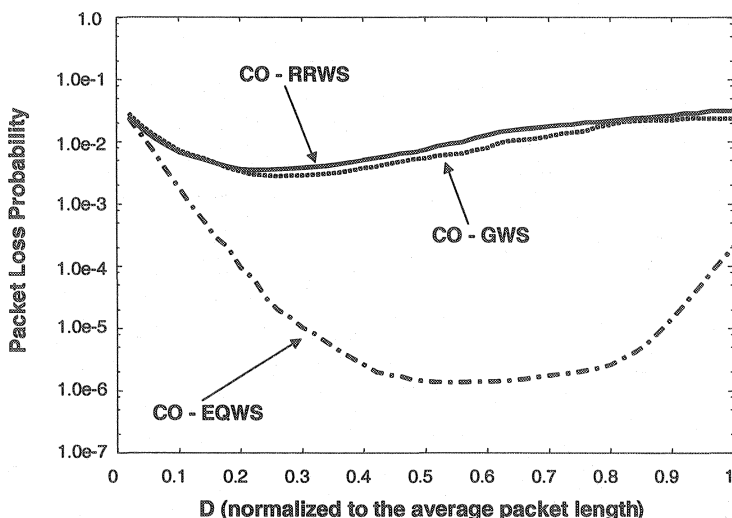


Figure 6. Packet loss probability as a function of the delay unit D for 4x4 switch with 16 wavelengths per fiber, 8 delay lines, uniform traffic and load 0.8.

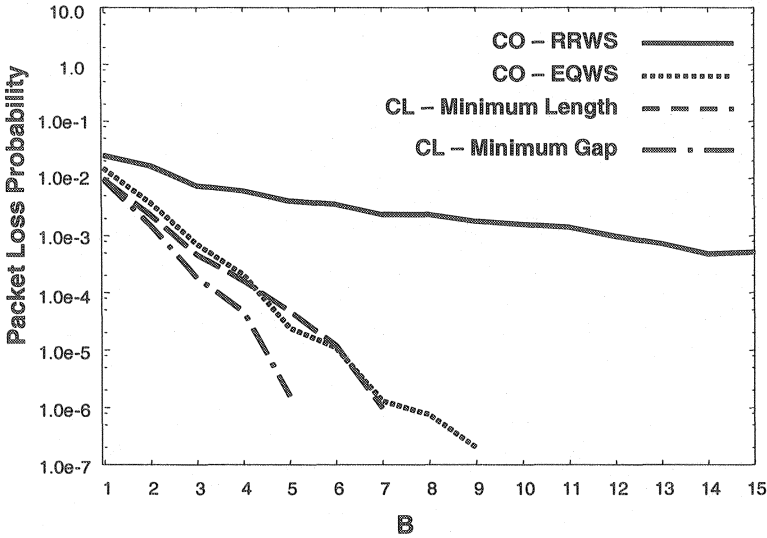


Figure 7. Packet loss probability as a function of the number of delay lines, for the delay unit D that minimizes the packet loss, a 4×4 switch with 16 wavelengths per fiber, uniform traffic and load 0.8.

Nevertheless, the CL minimum gap performs even better because it aims at reducing such gaps and it does it packet by packet, while this is not possible in the CO case.

The former results focus on PLP but we already outlined that, besides PLP, it is also very important to understand how the connection-oriented nature of each LSP is kept by the algorithms presented. In Fig. 8 the percentage of LSP-to-wavelength reassignments as a function of D for the three CO algorithms is shown, as an evaluation of the forwarding transparency of the algorithms. The figure shows that with only 8 delay lines EQWS is best from this point of view, with an LSP-to-wavelength reassignment each seventh packet. Also the reduction of wavelength re-allocations with respect to RRWS and GWS is very large.

In Fig. 9 the number of re-allocations for the EQWS algorithm is plotted as a function of D varying the number of delay lines as a parameter. As could be expected this number decreases with larger queues, because congestion is less frequent and re-allocation less needed. It is very important to outline that, for $B = 16$ delay lines the number of reallocations is very small (about 5%) and this is sufficient to provide a packet loss probability that is not measurable with 10 Billions of packets simulated, therefore in a range below 10^{-9} .

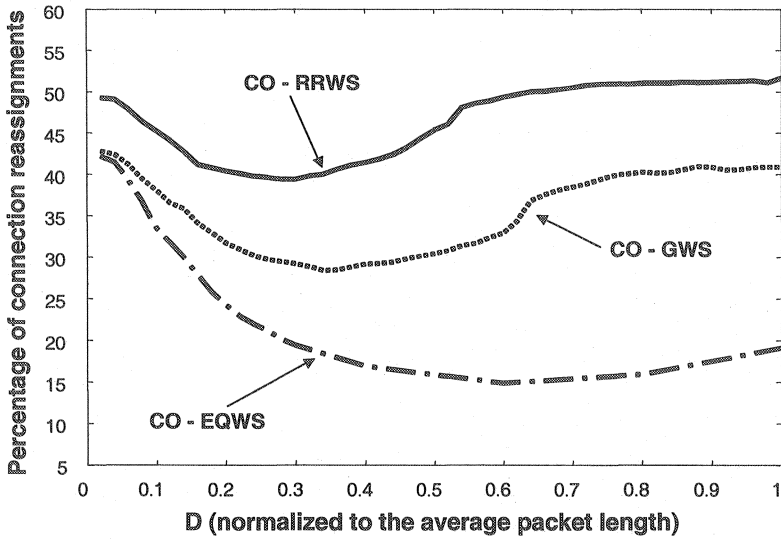


Figure 8. Comparison of number of LSP to wavelength re-allocations for the RRWS, GWS and EQWS as a function of D for a 4x4 switch with 16 wavelengths per fiber, 8 delay lines, uniform traffic and load 0.8.

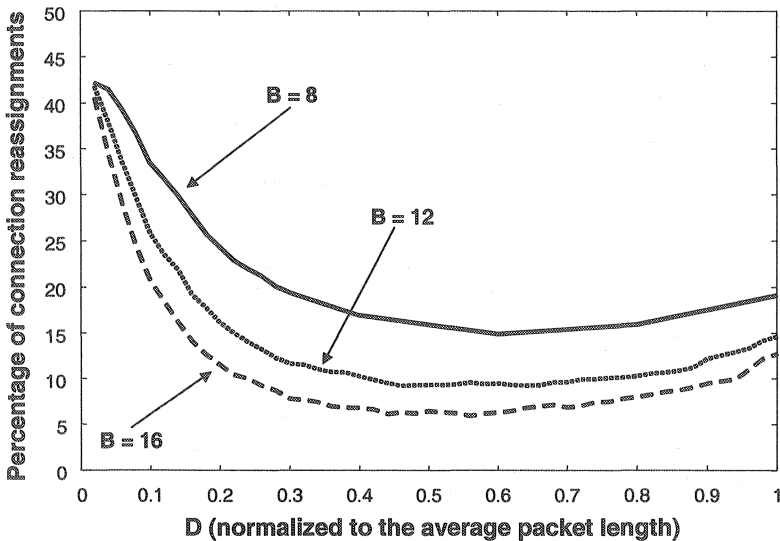


Figure 9. Comparison of number of LSP to wavelength re-allocations for the EQWS as a function of D for a 4x4 switch with 16 wavelengths per fiber, uniform traffic and load 0.8, varying the number B of delay lines in the buffer.

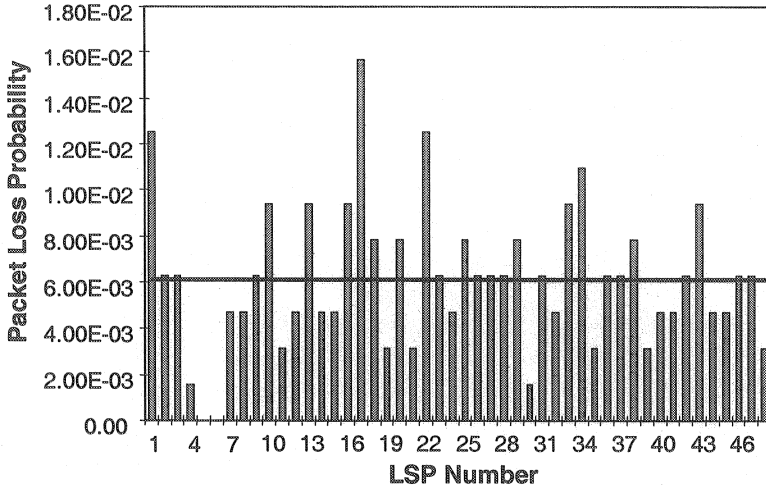


Figure 10. Packet loss probability per label switched path of the generic output link, for a 4x4 switch with 16 wavelengths per fiber, 4 delay lines, uniform traffic and load 0.8. The line represents the average value of the packet loss probability for the considered output link (=6.11E-03).

Finally, in Fig. 10 the PLP per LSP on a given output link is depicted together with the average PLP for that link in case of EQWS. The histogram gives a measure of the fairness of such algorithm, which is not the best due to some of the LSPs without losses against some others quite disadvantaged. In future studies we plan to improve the EQWS algorithm in order to reach better results also in terms of fairness.

6. CONCLUSIONS

This paper shows that the LSP to wavelength allocation in a WDM network carrying MPLS connections has a major impact on the network performance. We have shown that performance can be improved by means of dynamic wavelength allocation algorithms, and presented a comparison of three different algorithms.

The numerical results show that a dynamic algorithm that exploits the connection-oriented nature of the traffic, such as the Empty Queue Wavelength Selection here proposed, is able to realize the best improvement

both in terms of reduction of the packet loss probability and of limited overload on the network control functions.

ACKNOWLEDGMENTS

Work partially funded by the Italian Minister of Scientific Research, project "IP Optical Packet Networks - IPPO" and by the Commission of the European Community, Project IST-1999-11742 "DAVID - Data And Voice Integration over DWDM".

REFERENCES

- [1] F. Masetti et al., "High speed, high capacity ATM optical switches for future telecommunication transport networks", *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 5, pp. 979-998, June 1996.
- [2] 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.
- [3] L. Chlamtac et al., "CORD: contention resolution by delay lines", *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 5, pp. 1014-1029, June 1996.
- [4] E. Rosen, A. Viswanathan, R. Callon, "Multiprotocol Label Switching Architecture", IETF RFC 3031, January 2001.
- [5] E. Mannie et al., "Generalized Multiprotocol Label Switching (GMPLS) architecture", (draft-ietf-ccamp-gmpls-architecture-00.txt), IETF draft, June 2001.
- [6] L. Tančevski, S. Yegnanarayanan, G. Castañón, L. Tamil, F. Masetti, T. McDermott, "Optical routing of asynchronous, variable length packets", *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 10, pp. 2084-2093, October 2000.
- [7] F. Callegati, G. Corazza, C. Raffaelli, "Design of a WDM Optical Packet Switch for IP traffic", *Proc. of IEEE Globecom 2000*, San Francisco, USA, November 2000.
- [8] F. Callegati, "Optical Buffers for Variable Length Packets", *IEEE Communications Letters*, Vol. 4, No. 9, pp. 292-294, September 2000.
- [9] C. Raffaelli, "Architectures and Performance of Optical Packet Switching over WDM", *Photonic Network Communications*, Kluwer Academic Publishers, Vol. 3, No. 1/2, pp. 91-99, January 2001.
- [10] C. Guillemot et al., "Transparent Optical Packet Switching: the European ACTS KEOPS project approach", *IEEE/OSA Journal of Lightwave Technology*, Vol. 16, No. 12, pp. 2117-2134, December 1998.
- [11] L. Tančevski, A. Ge, G. Castañón, L. S. Tamil, "A New Scheduling Algorithm for Asynchronous, Variable Length IP Traffic Incorporating Void Filling", *Proc. of OFC '99*, San Diego, USA, February 1999.
- [12] F. Callegati, W. Cerroni, "Wavelength Selection Algorithms in Optical Buffers", *Proc. of IEEE ICC 2001*, Helsinki, Finland, June 2001.
- [13] F. Callegati, W. Cerroni, G. Corazza, C. Raffaelli, "Wavelength Multiplexing of MPLS Connections", *Proc. of ECOC 2001*, Amsterdam, The Netherlands, October 2001.

- [14] F. Callegati, W. Cerroni, G. Corazza, C. Raffaelli, "MPLS over Optical Packet Switching", Proc. of IWDC 2001, Taormina, Italy, September 2001.