

A FRAMEWORK FOR THE ANALYSIS OF DELAY JITTER IN OPTICAL PACKET SWITCHED NETWORKS

F. Callegati,¹ W. Cerroni,² G. Muretto,¹ C. Raffaelli,¹ P. Zaffoni¹

¹*DEIS - University of Bologna*

Viale Risorgimento, 2 - 40136 Bologna - ITALY

{fcallegati,gmuretto,craffaelli,pzaffoni}@deis.unibo.it

²*CNIT - Research Unit at the University of Bologna*

Viale Risorgimento, 2 - 40136 Bologna - ITALY

walter.cerroni@cnit.it

Abstract: Out of order delivery and/or delay jitter are typical phenomena occurring in networks adopting a packet transfer mode and may have a relevant impact on the behavior of higher layer protocols. In this paper an original framework is proposed to analyze these phenomena in optical packet-switched networks that employ the wavelength and the time dimensions to solve congestion. This framework is the first step toward a deeper understanding of the interaction between transport networks and higher layers behaviors.

Key words: Optical networks, Optical packet switching, Delay jitter, TCP

1. INTRODUCTION

Optical Packet Switching (OPS) is a networking technology promising a huge breakthrough in terms of available capacity [1] and is considered the best candidate to provide the long-term solution for future, very high-speed networks.

Many fundamental networking problems related to OPS are still open: the interworking with higher layer functions, such as end-to-end flow control mechanisms, is one of the most significant. Some works on this issue recently appeared mainly focusing on the effects of aggregation of IP packets into longer optical bursts or packets [2, 3]. In general not many results are available regarding the effects that congestion, and the consequent congestion resolution

mechanisms, in the OPS network may have on the behavior of transport protocols such as TCP.

Congestion resolution in OPS can be implemented in the time domain, using some form of delay lines, in the wavelength domain, sending contending packets to different wavelengths on the same fiber, and in the space domain, by means of deflection and alternate routing. Deflection routing is the only congestion resolution scheme considered in [3], while here we focus on the combined use of time and wavelength, that have been shown to realize a good trade-off between network control complexity and performance [4, 5].

The congestion resolution schemes affect the packet stream by modifying the time framework of the packet stream (delay jitter), changing the order of the packet transmission (out-of-order) and dropping some packets because of lack of resources (packet loss). In this paper we focus on delay jitter and out-of-order delivery, aiming at understanding how to evaluate the effects of the congestion resolution mechanisms on these phenomena.

We assume a network architecture consisting of OPS facilities exploiting a DWDM transmission infrastructure and capable to transport IP traffic by means of integration with a GMPLS control plane [6]. The network operation is therefore connection-oriented and the switching granularity of the OPS nodes is at the Label Switched Path (LSP) level. Each LSP represents a top-level, explicitly routed path formed by an aggregation of lower-level connections including several traffic flows (an Internet Draft proposes how to implement LSP hierarchies in GMPLS [7]).

We assume the availability of an all-optical switching matrix able to switch variable-length packets. Implementation issues are beyond the scopes of this work and therefore a general OPS node with full connectivity and wavelength conversion capabilities is considered. The node may also delay packets by means of buffers realized with Fiber Delay Lines (FDLs) [8].

The content of the paper is organized as follows. Section 2 provides a brief review of the congestion resolution issue in OPS. In section 3 the problem of out-of-sequence packet delivery and its influence on end-to-end protocol performance is discussed. In section 4 the proposed methodology to measure delay jitter and evaluate the degree of out-of-sequence events is presented. Conclusions are drawn in section 5.

2. CONGESTION RESOLUTION IN THE OPS NODE

The basic assumption that makes possible the use of the wavelength domain for congestion resolution is that network paths are associated with fibers and not with wavelengths. This is motivated by the fact that all the wavelengths of the same fiber can be seen as a set of parallel links toward the same destination. Within a node, the forwarding component decides to which path (i.e. fiber)

a packet must be sent and then the Switch Control Logic (SCL) decides the detailed scheduling and sets up the switching devices. The objective of the SCL is to exploit at its best the resource usage by performing two major tasks:

- choose which wavelength of the output fiber will be used to transmit the packet;
- decide whether the packet has to be delayed by using the FDL buffer or it has to be dropped.

This is called the Wavelength and Delay Selection (WDS) problem because the choices of wavelength and delay are actually correlated, being the need to delay a packet related to the availability of the wavelength selected. The WDS algorithm to solve this problem can be implemented by following different policies:

- **Static** - the LSP is assigned to a wavelength at LSP set-up and this assignment is kept constant all over the LSP lifetime. This approach requires little complexity due to processing at LSP set-up only.
- **Connectionless-like** - the wavelength is selected on a per-packet basis. This approach provides maximum flexibility in resource allocation but also requires per-packet processing, therefore it is fairly demanding in terms of load on the SCL.
- **Dynamic** - the LSP-to-wavelength assignment is executed only when congestion arises, i.e. when the time domain is not enough to solve contention due to the lack of buffering space.

Several heuristic connectionless-like WDS algorithms have been studied in the past, showing that they may significantly change the performance [4, 5]. The price to pay for this performance improvement is a non-negligible processing effort.

On the other hand, it has been observed that with a static wavelength allocation the switch performance strongly depends on the configuration of the LSP forwarding table [9] and is very sensible to the distribution of the LSP destinations, providing performance that is not easy to control.

The dynamic case is somewhat in between, simpler than the connectionless alternative but also capable of providing a more uniform and effective resource utilization than the static case. Therefore this approach is preferable in a connection-oriented network.

The major drawback of the dynamic WDS algorithms is that they do not preserve, a priori, the sequence of the packet flow neither at a global switch level nor at the LSP level. This is mainly caused by the possibility of multiplexing packets following the same LSP on different wavelengths, allowing parallel transmission. Furthermore, the adoption of dynamic routing algorithms [10] may be an additional cause of unordered deliveries.

3. THE PROBLEM OF PACKET REORDERING

As already outlined packet loss as well as out-of-order packet delivery and delay variations affect end-to-end protocols behavior and may cause throughput impairments [11, 12].

When considering TCP-based traffic it is well known that these phenomena influence the typical congestion control mechanisms adopted by the protocol [13] and may result in a reduction of the transmission window size and consequently in bandwidth under-utilization. In particular the TCP congestion control is highly affected by loss or out-of-order delivery of bursts of segments. This is exactly what may happen in the OPS network where traffic is typically groomed and several IP datagrams (and therefore TCP segments) are multiplexed in a single optical packet, because optical packets must satisfy a minimum length requirement to guarantee a reasonable switching efficiency. Therefore out-of-order or delayed delivery of just one optical packet may result in out-of-order or delayed delivery of several TCP segments, causing multiple duplicate ACKs and/or expired timeouts and triggering congestion control mechanisms which cause unnecessary reduction in the window size.

Another example of how out-of-sequence packets may affect application performance is the case of delay-sensitive UDP-based traffic, such as real-time traffic. In fact unordered packets may arrive too late and/or the delay required to reorder several out-of-sequence packets may be too high with respect to the timing requirements of the application.

These brief and simple examples make evident the need to limit the number of unordered packets. In general out-of-order delivery is caused by the fact that packets belonging to the same flow of information can take different paths through the network and then can experience different delays. In traditional connection-oriented networks, packet reordering is not an issue since packets belonging to the same connection are supposed to follow the same virtual network path and therefore are delivered in the correct sequence, unless packet loss occurs. In an OPS network using the wavelength domain for congestion resolution, this may not be the case. Packets traveling along the same network path may use different wavelengths in order to exploit wavelength multiplexing for congestion resolution purposes. Therefore it may happen that packets of the same flow are delivered out of sequence, even though still following the same network path.

A possible solution could be to assume that this problem is solved at the egress edge-nodes that should take care of re-sequencing the various packet flows. This assumption in our view is not very realistic. It can be feasible for some flow of high-value traffic, but it is unlikely that this will happen for all the flows of best effort traffic, because of the amount of memory and processing effort that would be required. Therefore we argue that it is important

and necessary to control delay jitters and out-of-order delivery of packets directly in the OPS network nodes. This is what we will discuss in the following section, where we realize that, first of all, a clear definition of this term is necessary because of the difference between the OPS network and conventional networks.

4. OUT-OF-ORDER AND DELAY VARIATION IN OPS NETWORK WITH WDM

Some authors have been dealing with measuring the degree of packet reordering [14, 15] and the impact on higher layers performance [12] in a traditional Internet scenario. However in the scenario of an OPS network using WDM the problem is different. As already explained, packets belonging to the same LSP may be transmitted on different wavelengths according to the principles built in the WDS algorithm. This may cause overtaking and/or partial overlapping of packets belonging to the same connection in a number of different ways.

A full understanding of such phenomena requires:

- to understand which cases of out-of-order packets may happen and how this phenomena can be measured in some quantitative way;
- to evaluate which are the WDS algorithms that preserve at best the time framework of the packet flows;
- to understand the effects that a specific case of out-of-order delivery may have on higher layer protocols.

This paper addresses the first issue by defining, in this section, a framework to measure, at the single node level, the modification in the data packet flows caused by the WDS algorithm. Then the paper provides an example of application of this framework to two typical WDS algorithms, to show how this tool can be used.

4.1 A framework to evaluate the delay jitter

For a generic packet P_i crossing a given OPS node, let t_i be the arrival time at the node input, s_i the departure time from the node output and $d_i = d_p + k_i D$ the delay introduced by the node itself, due to the packet header processing time (d_p , fixed) and the possible delay inside the FDL buffer ($k_i D$, where $k_i = 0, 1, \dots, B$). Obviously $d_i = s_i - t_i$.

Let assume that two generic subsequent packets belonging to the same traffic flow P_n and P_{n+1} arrive in order, i.e. $t_{n+1} > t_n$. Let $\Delta t_n = t_{n+1} - t_n$ and $\Delta s_n = s_{n+1} - s_n$ be the relative packet offsets at the node input and output respectively. The *jitter* between packets P_n and P_{n+1} , representing the packet offset variation due to the node crossing, may be defined as

$$J_n = \Delta t_n - \Delta s_n \quad (1)$$

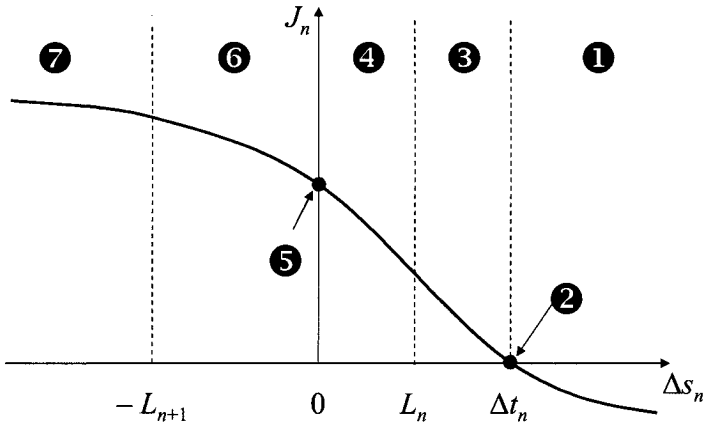


Figure 1. Behavior of the jitter depending on relative packet offset at node output

Equation (1) may also be written as

$$J_n = d_n - d_{n+1} = (k_n - k_{n+1})D = h_n D \quad (2)$$

where $-B \leq h_n \leq B$. The behavior of J_n for two particular packets P_n and P_{n+1} , with length L_n and L_{n+1} respectively, is shown in figure 1, where the x axis has been divided in seven different regions:

1. $\Delta s_n > \Delta t_n$ when the packet sequence is always guaranteed since P_{n+1} experiences more delay than P_n ($J_n < 0$);
2. $\Delta s_n = \Delta t_n$ when the node is transparent and P_n and P_{n+1} have the same offset at the input and output ($J_n = 0$);
3. $L_n \leq \Delta s_n < \Delta t_n$ when P_{n+1} experiences less delay than P_n ($J_n > 0$) but at the output it is still behind the tail of P_n (i.e. $s_{n+1} \geq s_n + L_n$);
4. $0 < \Delta s_n < L_n$ when the head of P_{n+1} partially overlaps the tail of P_n ;
5. $\Delta s_n = 0$ when P_{n+1} completely overlaps P_n ($J_n = \Delta t_n$);
6. $-L_{n+1} < \Delta s_n < 0$ when P_{n+1} has overtaken P_n but they are partially overlapping (i.e. $|\Delta s_n| < L_{n+1}$);
7. $\Delta s_n \leq -L_{n+1}$ when P_{n+1} has completely overtaken P_n (i.e. $s_n \geq s_{n+1} + L_{n+1}$).

4.2 A numerical example

The previous formalization allows to evaluate the delay jitter distribution as well as the amount of out-of-order packets, that depends on the specific definition of packet sequence. For instance, in case overlapping packets are not

considered in sequence, then the out-of-sequence regions will be 1, 2, and 3. If some overlapping is allowed, then out-of-sequence is guaranteed also in region 4. The same for region 5, in case packets arriving at the same time are not considered out of order.

As an example, figure 2 shows the jitter distribution over the different regions for a static and a connectionless-like WDS policy. The results for dynamic WDS are very similar to the connectionless-like case. As expected static WDS succeeds in maintaining the packet sequence, although it gives worst performance in terms of packet loss probability. On the other hand more dynamic policies cause some packets to get out of the node unordered, but the most frequent behavior is the one related to region 2, which means that congestion happens rarely and the packets are often transmitted transparently across the node.

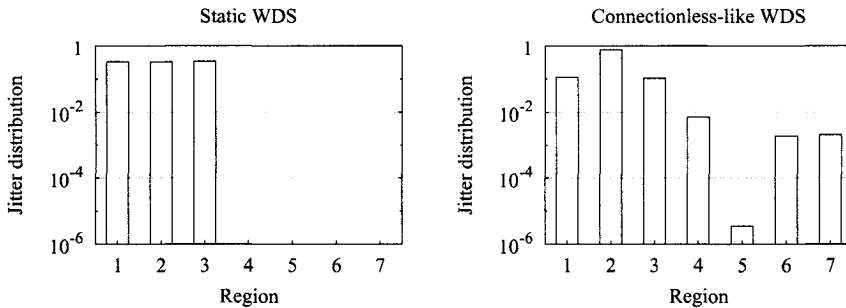


Figure 2. Jitter distributions for static and connectionless-like WDS policies

5. CONCLUSIONS

In this paper we have discussed the effect of scheduling algorithms on the packet sequence and time framework, in the scenario of an OPS network exploiting the time and wavelength domains for congestion resolution. We have proposed an original framework to quantitatively analyze delay jitter and out-of-order packets. This analysis can be used as a basis to compare different scheduling algorithms as shown in the numerical example provided. Moreover it represents the starting point to establish a link between the OPS network performance and the performance of higher layer protocols. This last issue is currently under investigation.

ACKNOWLEDGMENTS

This work is partially funded by the Italian Ministry of Education and University (MIUR) under the projects “INTREPIDO - End-to-end Traffic Engi-

neering and Protection for IP over DWDM Optical Networks” and “GRID.IT - Enabling platforms for high-performance computational grids oriented to scalable virtual organizations”.

REFERENCES

- [1] M.J. O’Mahony, D. Simeonidou, D.K. Hunter, A. Tzanakaki: The application of optical packet switching in future communication networks, *IEEE Communications Magazine*, Vol. 39, No. 3, March 2001, pp.128-135.
- [2] A. Detti, M. Listanti: Impact of segments aggregation on TCP Reno flows in optical burst switching networks, *Proc. of IEEE/ACM INFOCOM 2002*, Vol. 3, June 2002, pp. 1803-1812.
- [3] Jingyi He, H.G. Chan: TCP and UDP performance for Internet over optical packet-switched networks, *Proc. of IEEE ICC ’03*, Vol. 2, May 2003, pp. 1350-1354.
- [4] L. Tancevski et al.: Optical routing of asynchronous, variable length packets, *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 10, pp. 2084-2093, October 2000.
- [5] F. Callegati, W. Cerroni, G. Corazza: Optimization of wavelength allocation in WDM optical buffers, *Optical Networks Magazine*, Vol. 2, No. 6, November/December 2001, pp. 66-72.
- [6] L. Berger, Ed.: Generalized Multi-Protocol Label Switching (GMPLS) signaling functional description, IETF RFC 3471, January 2003.
- [7] K. Kompella, Y. Rekhter: LSP hierarchy with generalized MPLS TE, draft-ietf-mpls-lsp-hierarchy-08.txt, IETF draft, September 2002.
- [8] D.K. Hunter, M.C. Chia, I. Andonovic: Buffering in optical packet switches, *IEEE/OSA Journal of Lightwave Technology*, Vol. 16, No. 12, December 1998, pp. 2081-2094.
- [9] F. Callegati, W. Cerroni, C. Raffaelli, P. Zaffoni: Dynamic wavelength assignment in MPLS optical packet switches, *Optical Networks Magazine*, Vol. 4, No. 5, September 2003, pp. 41-51.
- [10] F Callegati, W Cerroni, G. Muretto, C. Raffaelli, P. Zaffoni: Adaptive routing in DWDM optical packet-switched networks, *Proc. of 8th IFIP ONDM 2004*, Gent, Belgium, February 2004, pp. 71-86.
- [11] J. C. R. Bennett, C. Patridge: Packet reordering is not a pathological network behavior, *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, December 1999, pp. 789-798.
- [12] M. Laor, L. Gendel: The effect of packet reordering in a backbone link on application throughput, *IEEE Network*, vol. 16, no. 5, September/October 2002, pp. 28-36.
- [13] M. Allman, V. Paxson, W. Stevens: TCP congestion control, IETF RFC 2581, April 1999.
- [14] T. Banka, A. A. Bare, A. P. Jayasumana: Metrics for degree of reordering in packet sequences, *Proc. of IEEE LCN 2002*, Tampa, FL, November 2002, pp. 333-342.
- [15] S. Jaiswal, G. Iannacone, C. Diot, J. Kurose, D. Towsley: Measurement and classification of out-of-sequence packets in a tier-1 IP backbone, *Proc. of INFOCOM 2003*, San Francisco, CA, March 2003, vol. 2, pp. 1199-1209.