

GUARANTEEING SEAMLESS END-TO-END QOS IN OBS NETWORKS

Maurizio Casoni,¹ Maria Luisa Merani¹, Alessio Giorgetti², Luca Valcarenghi²,
Piero Castoldi²

¹*Department of Information Engineering, University of Modena and Reggio Emilia - Via Vignolese, 905 - Modena - Italy*

²*Sant'Anna School of University Studies and Doctoral Research, Center of Excellence for Communications Networks Engineering (CEIRC) - via Cisanello, 145 - Pisa - Italy*

Abstract:

In this paper the authors propose a method to guarantee end-to-end QoS to multiple traffic classes in optical burst switched (OBS) scenario, even in case of network congestion. The OBS network utilizes a core node architecture with no fiber delay lines and a limited set of wavelength converters. Traffic class performance differentiation is achieved by allowing high class traffic to utilize more node functionalities than low class traffic. To improve the likelihood of finding a route, even upon network congestion, bursts are allowed either to be deflected from their default route or to choose from a set of preplanned end-to-end paths. Network performance evaluation is focused on burst blocking probability and end-to-end TCP throughput. Performance is determined under the assumption of exponentially distributed burst interarrival times and arbitrarily distributed burst durations. Numerical results show that the proposed approach is able to guarantee different end-to-end TCP throughput performance to each traffic class. Moreover the proposed burst routing policies allows to decrease, with respect to a shortest path routing policy, the burst blocking probability. Thus the end-to-end TCP throughput is seamlessly guaranteed even in case of network congestion.

1. INTRODUCTION

Optical networking ultimate goal is the development of a full optical Internet, where signals carried within the network never leave the optical domain [1]. A first important step in this direction is to have optical networks transparent at least for data, with the control part converted and processed in

electronics. In Optical Burst Switching (OBS) [2,3] the key idea is to dynamically set up an optical path, i.e. a lightpath, whenever a *large* data flow is identified and needs to traverse the network. In OBS data never leave the optical domain but for each data burst assembled at the network edge, a reservation request is sent as a separate control packet, well in advance, and processed within the electronic domain. The control packet, carrying relevant forwarding and routing information, precedes each burst by a basic offset time. The offset time is set to accommodate the non-zero electronic processing time of the control packets inside the network nodes. OBS network nodes can be classified as either edge or core routers. The main task of edge nodes is the burst assembly function: as they represent the border between "traditional" electrical LAN/MAN IP networks and a high speed optical transport network, they must collect incoming IP datagrams and assembly them into bursts according to suitable algorithms. Core routers, on the other hand, deal with data bursts and the related control packets; they have to set up internal optical paths on the fly for switching bursts and take them hop-by-hop to their final destination. The control packet offset time allows the core switches to be buffer-less, thus avoiding the utilization of optical memories, e.g. fibre delay lines, required on the other hand by optical packet switching [4].

In a previous paper [5] some of the authors have evaluated, in terms of burst loss and delay, the performance of OBS networks that utilize the Just Enough Time (JET) reservation mechanism and mechanisms for service differentiation. In this paper the authors propose a scheme for differentiating the blocking probability experienced by three different burst classes and for making the end-to-end throughput unaffected by network congestion. The considered OBS network is built of core nodes equipped with a limited set of full range optical wavelength converters. Burst class differentiation is implemented by allowing different class bursts to utilize different node functionalities. In addition selected class bursts are allowed either to be deflected from their default shortest path route or to choose among a set of preplanned end-to-end paths. Thus during the optical path routing different burst classes experience a different blocking probability. Furthermore the round trip time experienced by the bursts can be either unpredictable, in case of deflection routing, or increase of a bounded amount, in case of alternate end-to-end re-routing.

The metrics utilized to evaluate the performance of the proposed scheme are the burst blocking probability, i.e. the probability that a burst is blocked because the request of bandwidth at the traversed optical node cannot be fulfilled, and the end-to-end throughput, obtained from the burst blocking probability and the round trip time (RTT) through a widely-accepted TCP model. Numerical results shows that assigning to different classes the right of utilizing different node functionalities allows to differentiate burst classes blocking probability. In addition the possibility for the optical paths to be rerouted al-

lows bursts to overcome network congestion. Moreover while deflection routing unpredictably affects the RTT, network congestion avoidance through alternate end-to-end re-routing allows to limit the burst blocking probability and only slightly affects RTT and therefore does not downgrade the end-to-end TCP throughput.

The rest of the paper is organized as follows. In Section 2 the core network with the related routing algorithms is shown. Section 3 provides the framework for the end-to-end performance evaluation when a transport protocol like TCP is employed. Numerical results are then collected in Section 4 while concluding remarks are in Section 5.

2. OBS NETWORK MODEL

This work assumes that OBS nodes support JET and that traffic is mapped into three (but in general N_{QoScl}) QoS classes. Each class is characterized by a different statistical traffic description and features different QoS requirements determined by specifying an upper bound of burst blocking probability and/or end-to-end delay. Traffic incoming into edge nodes is supposed to be *M/Pareto*. If, for instance, \bar{x}_1 , \bar{x}_2 and \bar{x}_3 are the mean duration times of the ON periods due to aggregation of short, medium and long-sized datagrams, respectively, on each incoming link the offered load ρ is the sum of three distinct contributions:

$$\rho = \sum_{i=1}^3 \rho_i = \sum_{i=1}^3 p_i \lambda \bar{x}_i, \quad (1)$$

where p_1 , p_2 and p_3 represent the occurrence probabilities of the three classes and λ is the mean arrival rate.

The core optical routers are assumed to be equipped with $M \times M$ optical interfaces capable of supporting N wavelengths each and with a limited pool of wavelength converters wc . The optical nodes are buffer-less, i.e., no fiber delay lines (FDLs) are present in order to resolve contention for an output fiber, output wavelength (Figure 1). It is however reasonable to assume the use of a set of input FDLs [6], whose exclusive task is to re-align the OBS data burst and its control packet, so as to guarantee a minimal offset time at any intermediate node, as depicted in Figure 1.

Bursts are created according to three class of service, class 1 carrying time-sensitive data, class 2 and 3 loss-sensitive data. The bursts default route is the shortest hop path from source to destination. However routing can be modified in order to better meet the performance required by each class; in particular no time-sensitive bursts are allowed to be re-routed along longer paths to decrease their blocking probability. Two alternative re-route schemes, deflection routing at intermediate nodes and end-to-end re-routing, are compared for traffic class

2 and 3 in terms of achievable burst blocking probability and TCP throughput. Deflection simply deflects bursts toward another output fiber if not any wavelength is available in the desired output fiber; while to minimize the blocking probability end-to-end re-route can be performed as presented in [7] by avoiding the contention among different optical paths.

The overall strategy for service differentiation is as follows: class 1 bursts are given the highest priority through an additional extra-offset and the use of wavelength converters in the core nodes; class 2 bursts have medium priority by just using the converters; class 3 bursts have low priority since they have no extra offset and cannot exploit wavelength converters; on the other hand, class 2 and 3 can use alternative sub-optimal variable delay paths through deflection or end-to-end re-routing.

In this paper the network considered covers most European countries (Figure 2). London, Oslo and Stockholm are sources of information flows and operate as edge routers, whereas Madrid, Rome and Athens are the possible destinations. All the other nodes work as core routers. Moreover, no flow is supposed to enter or leave the network at intermediate steps.

Section 4 will show the performance, in terms of overall edge-to-edge burst blocking probability for the three classes of bursts in the OBS network, when each node implements the JET reservation mechanism with a limited set of wavelength converters available.

3. END TO END PERFORMANCE

The throughput is here studied on end-to-end basis when TCP is assumed as transport layer protocol. In fact, in wide area data networks like Internet congestion control mechanisms have a fundamental role for the global functioning. TCP is a reliable window-based acknowledgment-clocked flow control protocol, thought to avoid to overload the network and to react to a possible congestion at network level. In the system under investigation TCP Reno is assumed to be employed by hosts. TCP Reno is modeled following the approach detailed in [8] where the throughput (bit/s) is approximated by:

$$Thr_{TCP} = \frac{MSS}{RTT \sqrt{\frac{2bp}{3}} + T_0 \min\left(1, 3\sqrt{\frac{3bp}{8}}\right) p(1 + 32p^2)} \quad (2)$$

being MSS the maximum segment size expressed in bits, RTT the round trip time, p the segment loss probability, T_0 the time out and b the number of packets acknowledged by ACKs.

The performance of TCP over OBS networks have been studied in some previous works [9] [10] but it can still be considered an open issue and thus a challenge for the research community. Here we want to add some thoughts to the worldwide discussion. Let us assume, as classified in [9], to have only slow

TCP sources which emit at most one segment during the interval $(0, T_{max})$. It also means that at most one segment for each connection is contained in a burst generated by edge nodes and injected into the OBS network. Therefore, even if approximated, for this type of sources the segment loss probability p can be assumed equal to the burst blocking probability.

Now, considering the reference network topology (Figure 2) an average link length of 800 Km can be assumed and the number of hops, N_{hops} , is in the range [3 – 5]. Since the light propagation speed in the fibers is roughly 70% the speed of the light in the vacuum, the propagation delay for each hop, T_{hop} , is 4 ms. Therefore the one-way delay edge-to-edge, from entering the ingress edge to leaving the egress edge, $T_{e2e1way}$ in the OBS network is:

$$T_{e2e1way} = T_{assembly} + N_{hops} \times T_{hop} + T_{disass} \quad (3)$$

which can be bounded to 30 ms. If in addition the network has a symmetric behavior, RTT is approximately 60 ms. Actually, RTT has also to consider the delays given by the access networks before entering the OBS network: this means that 60 ms can be considered as a kind of lower bound.

In the next Section the behavior of the throughput of TCP will be shown referring to the above assumptions for p and RTT .

4. NUMERICAL RESULTS

The numerical results are obtained for the reference network, in which the performance figures are the burst loss and the end-to-end TCP throughput computed by exploiting the aforementioned analytical model of TCP. The performance of a complete burst switching network have been determined by simulation through an *ad-hoc* event-driven C++ object oriented simulator. Regarding core routers, it is set $wc = 20$, $M = 2$ with $N = 16$ wavelengths per fiber. The following values for p_1 , p_2 and p_3 are considered, (0.5, 0.2, 0.3). Incoming traffic into edge nodes is supposed to be *M/Pareto* with $\alpha_{on} = 1.2$, $\bar{x}_1 = 218$ bytes, $\bar{x}_2 = 576$ bytes $\bar{x}_3 = 1500$ kbytes. The extra offset for class 1 bursts is set at $6\mu s$. It is worth reminding that only class 1 and 2 bursts exploit a set of 20 wavelength converters; however, class 2 and 3 can be deflected in case of unavailability of wavelengths on the outgoing fibre or end-to-end re-routed along an alternate path. Figure 3 shows the total burst blocking probability for the three burst classes having Oslo as source and Rome as destination, when deflection routing is utilized for traffic class 2 and 3. In order to have loss values in the range of 10^{-3} for class 1, 1% for class 2 and 50% for class 3, the overall load ρ must be less than 0.4.

Let us now discuss the consequences of the above values of burst blocking probability on end-to-end performance. Figure 4 shows the throughput of TCP Reno given by (2) as a function of RTT for $T_0 = 1.0$ s, $b = 2$, $MSS = 1500$

bytes and $p = 1\%$. As mentioned in Section 3, for slow TCP sources p can be approximated with the burst blocking probability. Also, the RTT values of interest fall reasonably in the $[60 - 200]$ ms range, depending on the type of access network. This figure says that even if a very high speed core network is employed the best we can get is a throughput of 1.3 Mbps and it remarkably decreases at 750 kbps as soon as the RTT doubles, or at 500 kbps when the RTT becomes three times, i.e. 180 ms.

Figure 5 shows the improvement in burst blocking probability when end-to-end re-route is utilized with only one preplanned alternative path for each source-destination pair. It is important to notice that the proposed end-to-end re-route scheme is able to strongly decrease the burst blocking probability guaranteed from deflection routing. Therefore, as shown in figure 6, edge-to-edge TCP throughput can significantly increase even though the RTT increases because of the utilization of a longer path.

5. CONCLUSIONS

This work investigates burst blocking probability and overall performance of an OBS network, with buffer-less nodes, adopting JET reservation mechanism.

The authors numerically demonstrated that the optical node architecture under examination can provide acceptable values of burst blocking probability and end-to-end TCP throughput, once the offered load is properly limited. Moreover the proposed burst routing policy avoids that network congestion deteriorate the end-to-end TCP throughput.

ACKNOWLEDGMENTS

This work has been partially supported by MIUR within the framework of the national project "INTREPIDO: Traffic Engineering and Protection for IP over DWDM".

REFERENCES

- [1] B. Bostica, F. Callegati, M. Casoni, C. Raffaelli, "Packet Optical Networks for High Speed TCP-IP Backbones" -*IEEE Communications Magazine*, January 1999, pp.124-129.
- [2] C. Qiao, M. Yoo, "Optical Burst Switching (OBS) - a New Paradigm for an Optical Internet," *Journal of High Speed Networks*, No.8, pp.69-84, 1999.
- [3] M. Yoo, C. Qiao, S. Dixit, "QoS Performance of Optical Burst Switching in IP-Over-WDM Networks," *IEEE Journal on Selected Areas in Communications*, Vol.18, No.10, pp.2062-2071, October 2000.
- [4] M. Casoni, C. Raffaelli, "Tandem Architecture for Photonic Packet Switches" *Journal of Communications and Networks*, Vol.1, No. 3, September 1999, pp.145-152.
- [5] M. Casoni, M.L. Merani, "Resource Management in Optical Burst Switched Networks: Performance Evaluation of a European Network", Proc. of 1st International Workshop on

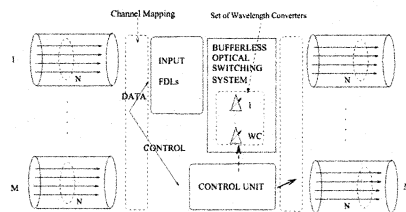


Figure 1. General core router bufferless architecture.

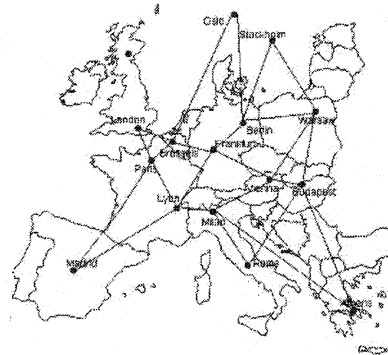


Figure 2. Reference network.

- Optical Burst Switching, October 16 2003, Dallas, (Texas, USA).
- [6] Y. Xiong, M. Vandenhoute, H.C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks", *IEEE Journal on Selected Areas in Communications*, Vol.18, No.10, October 2000, pp.1838-1851.
 - [7] A. Fumagalli and L. Valcarengi, "The Preplanned Weighted Restoration Scheme", Proc. of IEEE Workshop on High Performance Switching and Routing, 2001.
 - [8] J. Padhye, V. Firoiu, D. F. Towsley, J. F. Kurose, "Modeling TCP Reno Performance: A Simple Model and its Empirical Validation", *IEEE/ACM Trans. on Networking*, vol.8, no.2, pp.133-145, April 2000.
 - [9] A. Detti, M. Listanti, "Impact of Segment Aggregation on TCP Reno Flows in Optical Burst Switching Networks", Proc. of IEEE Infocom 2002, 23-27 June 2002, New York (U.S.A.).
 - [10] S. Gowda, R. K. Shenai, K. M. Sivalingam, H.C. Cankaya, "Performance Evaluation of TCP over Optical Burst-Switched (OBS) WDM Networks", Proc. of ICC 2003, 11-15 May 2003, Anchorage (U.S.A.).

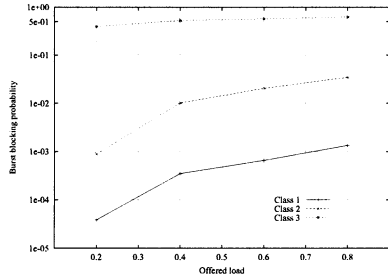


Figure 3. Total end-to-end burst blocking probability as a function of the offered load for the three burst classes, relative to (Oslo, Rome) pair.

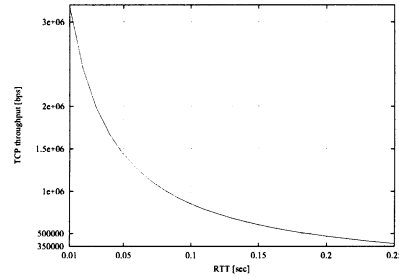


Figure 4. Throughput of TCP as a function of RTT for $MSS = 1500$ bytes and $p = 1\%$ achieved with shortest path routing.

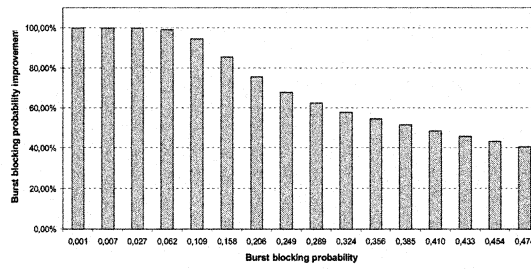


Figure 5. Burst blocking probability percentage improvement achievable utilizing one alternative path in function of bursts blocking probability p obtained with shortest path routing and deflection routing. If $p < 10\%$ the re-route scheme completely cancels block events.

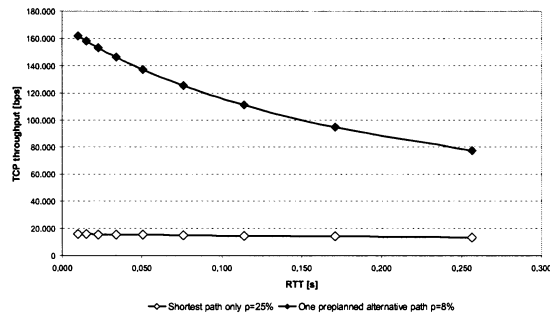


Figure 6. Throughput of TCP as a function of RTT for $MSS = 1500$ bytes and $p = 25\%$ achieved with shortest path routing, $p = 8\%$ achievable providing one alternative path.