The Design and Performance Analysis of QoS-Aware Edge-Router for High-Speed IP Optical Networks

E. Kozlovski, M. Düser, R. I. Killey, and P. Bayvel

Department of Electronic and Electrical Engineering, University College London

Abstract: *The design of the class-of-service capable edge-router for WDM IP optical backbone is considered. The effect of traffic burstiness is investigated by the comparison of Poisson and Pareto traffic models. The performance of the edge-router with respect to both packet loss rate (PLR) and average packet delay are studied by simulation under both traffic models. A trade-off between input buffering and output bandwidth provisioning for input traffic is proposed. It is shown that the combination of both techniques can significantly improve the performance of the edge-router.*

1. Introduction.

The optical burst-switched network model is considered here as introduced in [1]. Edge-routers are located at the ingress of the core optical network and each edge-router is connected to an optical core passive router. Electronic pre-sorting and scheduling of the incoming packets according to their class of service (CoS) are provided at the ingress in the electronic input buffers and flows of packets are then aggregated in electronic output buffers and sent off through the core network without intermediate optical processing. This way, each network node performs two main functions: (1) an edge-router carries out electronic processing of the packets arriving from the access layer and (2) an optical passive router directly routes traffic arriving from the other nodes (Fig. 1 (a)). Traffic arriving at the edge-router shares the same optical bandwidth with transit traffic in the core network. Because the latter type of traffic is multiplexed from different sources, it arrives with much higher bit-rate than that of the added traffic and is able of using the entire optical bandwidth available at any node. Therefore, the problem of sufficient bandwidth provisioning for the added traffic arises (Fig. 1 (b)). Whilst packet queuing delays can be slightly improved by applying different scheduling techniques [2], current investigation is focused on the relationship between the edge-router performance and output bandwidth provisioning as the latter seems to be the only way to improve the performance in the presence of burstiness intrinsic to IP traffic (see Section 3.1 for explanation). In the proposed architecture the performance is influenced both by the input buffering and flow aggregation in the output buffers. Current investigation is concerned with minimising the PLR and the delays at the stage of the input pre-sorting and scheduling.

2. Edge-Router Architecture and Traffic Model.

The edge-router carries out the following functions: *packet pre-sorting*, *packet electronic input buffering* and *scheduling*, and *flow aggregation*. The latter is performed when packets enter the output buffers so that aggregated bursts assembled out of the incoming traffic are routed through the core network over a wavelength which is assigned according to dynamic wavelength routing and assignment algorithm (DWRA) (Fig. 1 (c)). It is assumed that, whilst a scheduling algorithm resides in the control unit of each edge-router, DWRA is performed by means of centralised network-level control whose main functions are: (1) CoS-scheduling of the aggregated bursts arriving from each edge-router and (2) dynamic lightpath allocation to the bursts scheduled at the step (1).

In the simulation carried out in this work, it was assumed that each input buffer is connected to a traffic source that generates packets of three classes of service: premium

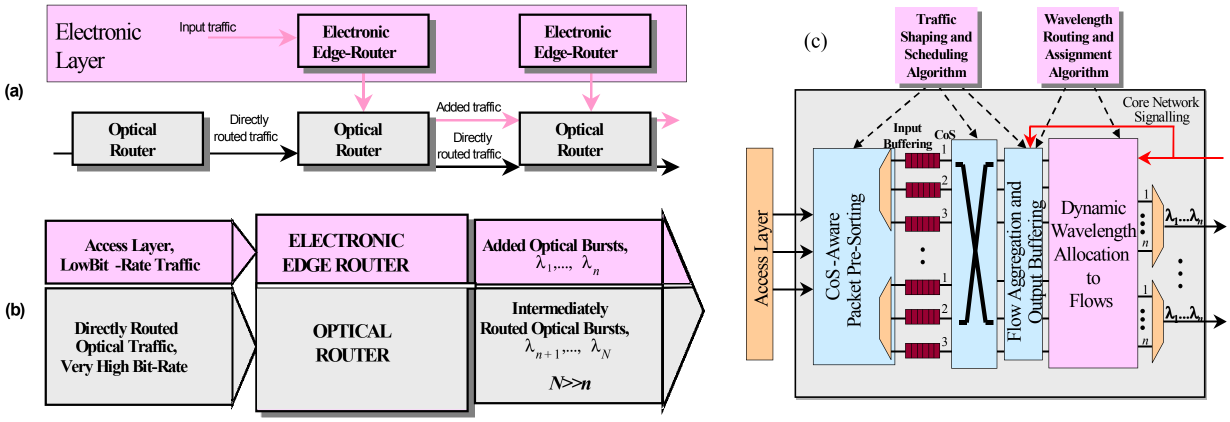

Fig. 1. ***System Architecture.***
(a) Electronic processing of input traffic in the edge-routers . (b) Bandwidth sharing between input traffic and intermediately routed traffic, (c) Edge-router schematic

traffic (the most delay-sensitive and packet-loss sensitive traffic which is given 90% of the router capacity), high-quality traffic (which is given 7% of the router capacity) and best-effort traffic. The uniform distribution was used to generate packet CoS and destination, assuming equal probability for any of CoS and any of output ports to be assigned to a packet.

In order to evaluate the performance of the router under different traffic conditions, the Poisson and the Pareto traffic models were compared. The effect of burstiness was studied using packet train model as described in [2]. Each traffic source alternatively produces a burst of active packets (ON-period) followed by a sequence of empty timeslots (OFF-period). The number of packets in the ON-period and the number of timeslots in the OFF-period are generated using the Pareto distribution with the PDF given by $f(X) = \alpha X_0^{\alpha} / X^{\alpha+1}$, where X_0 is the minimum number of X occurring in the simulation, and $1 \le \alpha \le 2, X_0 \ge 1$ are the conditions for traffic to exhibit self-similarity [3]. Using the above formula, the ON and OFF periods can be generated as follows: $T_{\mathrm{ON}}(U) = T_{\mathrm{OFF}}(U) = \lfloor X_0 / U^{1/\alpha} \rfloor$, where U is a random variable uniformly distributed between 0 and 1. Throughout the simulation, α was set to 1.5 and X_0 was set to 1. This way, the minimum burst length was just one packet.

The scheduling technique was based on the FIFO input scheduling as described in [2]. The following assumptions were made for the scheduling algorithm: (1) The packets entering the input buffer are of fixed size, and (2) the input parameters for the scheduling algorithm are destination field and CoS field of the packet.

3. Results and Discussion.

The performance of the edge-router with respect to the PLR and average packet latency was evaluated by simulation. In order to simulate PLR of 10^{-n}, as many as 10^{n+1} packets have to be generated in order to achieve the level of confidence of 95% and an interval of confidence of 2×10^{-n} [4]. The value of n was set to 5 throughout the simulations.

3.1. Effect of burstiness.

First, influence of traffic self-similarity on the performance of the edge-router was investigated compared to the Poisson traffic model under the different traffic loads. A system with a maximum input bit-rate of 10 Gbit/s and input buffer size of 25 Mbit was considered. The traffic load was defined as ratio between average bit-rate and maximum

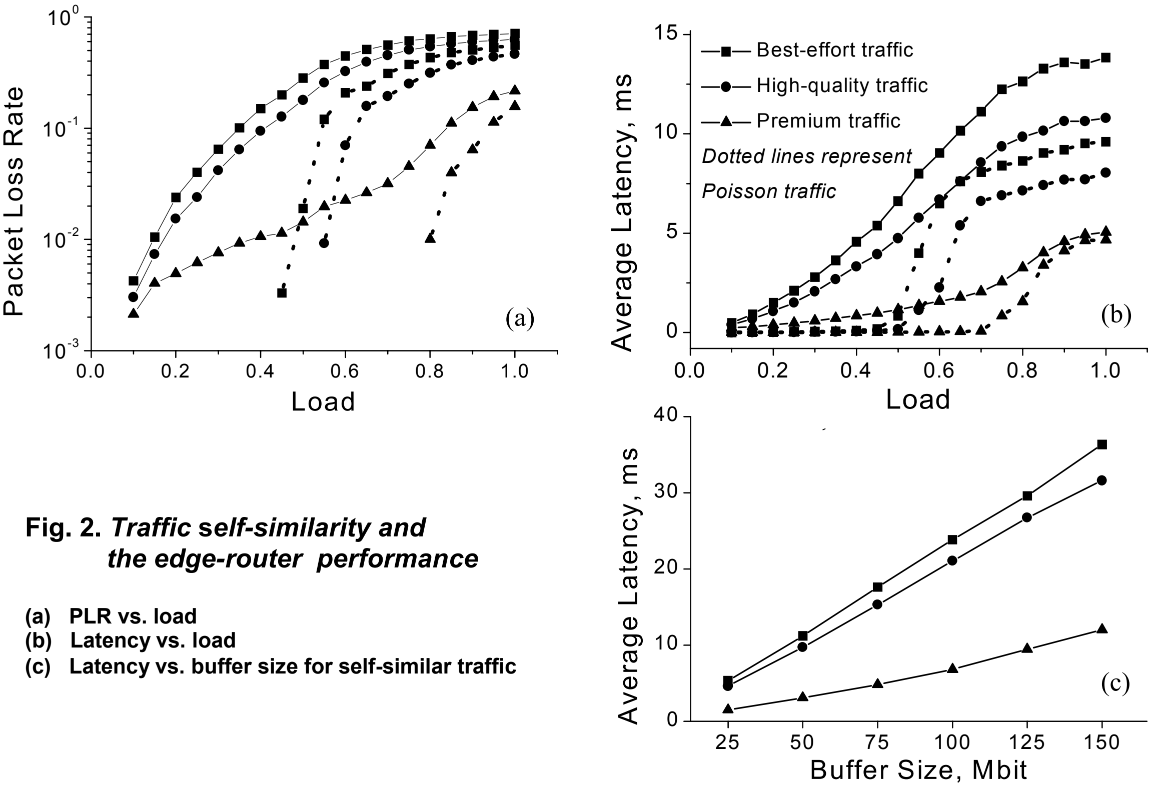

Fig. 2. ***Traffic self-similarity and the edge-router performance***

(a) PLR vs. load
(b) Latency vs. load
(c) Latency vs. buffer size for self-similar traffic

bit-rate for an input of the edge-router. The router was also assumed to have 4 input ports and 4 output ports and the number of wavelengths per each output port was assumed to be 3.

Figs. 2 (a)-(b) clearly show that the presence of self-similarity dramatically degrades the performance of the router. With the Poisson model, the PLR of less than 10^{-4} can be achieved under the load of 0.5 for both best effort and premium traffic. At the same time, with the Pareto model, under the load of 0.5 the PLR is of 10^{-1} for best-effort traffic and it is of 10^{-2} for premium traffic (see Fig. 2 (a)). As for the average latency, under the load of 0.5 the results show negligible latency for all three types of traffic with the Poisson model, whilst with the Pareto model under the same load, the latency of best-effort traffic is 7 ms, and the latency of premium traffic is about 1 ms (see Fig. 2 (b)).

Because bursts of infinite length can theoretically arrive at the router, it is virtually impossible to significantly improve the values of the PLR and the latency under bursty traffic conditions only by means of the improvement of the scheduling algorithm. The bursts can not be accommodated unless the buffers have an appropriate length. However, increase in the buffer length adds up to the packet delay and thus can not be considered either. Fig. 2 (c) shows how the increase of the buffer length affects average packet latency under the load of 0.8. It can be observed that the buffers of 150 Mbit lead to a dramatic increase of the delay up to 40 ms for best-effort traffic, and up to 12 ms for premium traffic.

3.2. Provisioning of the output bandwidth for input traffic.

Let the number of input ports be N, the bandwidth of the input port be B_{in} and the bandwidth of the output fibre be B_{out}. Then, the condition $B^*_{\text{out}} = N B_{\text{in}}$ guarantees collision-free traffic transport in the core network. However, this condition is unlikely to be met, because traffic directly routed through each node arrives with much higher bit-rate and occupies a significant portion of B_{out} which can not be reallocated to the input traffic (Fig. 1, (b)) and thus, $B_{\text{out}} << B^*_{\text{out}}$.

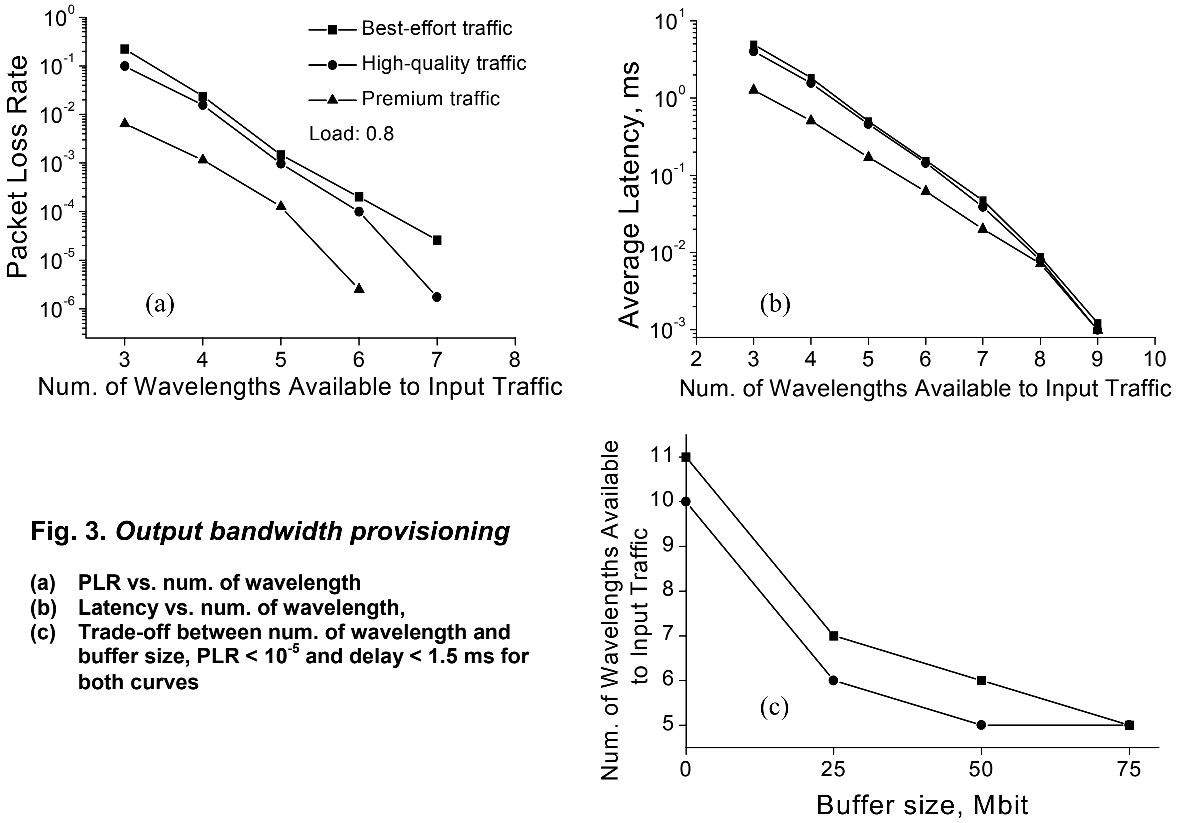

Fig. 3. ***Output bandwidth provisioning***

(a) PLR vs. num. of wavelength
(b) Latency vs. num. of wavelength,
(c) Trade-off between num. of wavelength and buffer size, PLR < 10^{-5} and delay < 1.5 ms for both curves

Therefore, the relationship between B_{out} and the edge-router performance was investigated. It is assumed that B_{out} can be increased by allocating additional wavelengths to the input traffic. Let W_{free} be the number of wavelengths available to input traffic. Then, by varying W_{free} one can analyse the value of B_{out} sufficient to achieve the desired edge-router performance.

The algorithm used for the wavelength assignment maintains an index table for the availability of free wavelengths in each output port and carries out the following procedures.

1. When a wavelength is occupied by transit traffic which directly goes through a given node, the value corresponding to the index of that wavelength is set to "occupied".
2. The search in the table of all the wavelengths free at a given time-slot is carried out. The set of free wavelengths represents current value of B_{out}, i.e. the bandwidth available for the added traffic at a given time-slot.
3. Once the scheduling algorithm has selected packets to be routed within the same time-slot, a free wavelength is randomly chosen for the assignment to these packets and also set to "occupied". The step is repeated until there are no more free wavelengths in the output.

Fig. 3 shows that the edge-router performance can be significantly improved by applying the above approach. In the simulation, the load is 0.8, the buffer length is 25 Mbit, and B_{in} corresponds to 3 wavelengths, i.e. $B_{\text{in}} = B_{\text{out}}$ when $W_{\text{free}} = 3$. It can be observed that the PLR can be decreased to 10^{-5} with $W_{\text{free}} = 6$, i.e. with $B_{\text{out}} = 1/2\, B^{*}_{\text{out}}$ for the premium traffic, and the same PLR can be achieved with $W_{\text{free}} = 7$ for the best-effort traffic (Fig. 3 (a)). The average latency is reduced to 56 μs with $W_{\text{free}} = 6$ for the premier traffic, i.e. 15 times reduction is achieved compared to $W_{\text{free}} = 3$. It is also reduced to 0.92 ms for the best-effort traffic, i.e. 5 times reduction is achieved compared to $W_{\text{free}} = 3$ (Fig. 3

(b)). It should be noticed that the latency of all three types of traffic is reduced to 8 μs with W_{free} = 8 and it is reduced to 1.2 μs with W_{free} = 9.

Because it is difficult to allocate the sufficient output bandwidth, a trade-off between the buffer length and B_{out} was analysed (Fig. 3, (c)). The curves represent the values of W_{free} required to achieve the PLR = 10^{-5} and average latency of up to 0.2 ms with different buffer lengths under the traffic load of 0.8. It's shown that with no buffers optimal performance is achieved with W_{free} = 10 for the premium traffic and W_{free} = 11 for the best-effort traffic. At the same time, with the buffer length of only 75 Mbit, both types of traffic require W_{free} = 5 to achieve the desired performance. It should be noticed that further increase in the buffer length will lead to the increase in average latency (see Fig. 2 (c)). Therefore, for the considered configuration of the edge-router, a trade-off between the buffer length and output bandwidth can be achieved by implementing up to 75 Mbit input buffers and provisioning of at most $B_{\text{out}} = 1.7\ B_{\text{in}}$.

4. Conclusions.

The investigation showed that a trade-off between the input buffer size and the output bandwidth can be achieved with respect to the desired PLR and average packet latency. The results show the following.

1. Improvement of the edge-router performance can not be achieved by the increase of the buffer capacity because such an increase adds up greatly to the queuing delays. Scheduling can slightly improve the performance for premium traffic, however, this improvement is still insufficient in the presence of self-similarity.
2. The combination of the output bandwidth provisioning and the input buffering can significantly improve the performance of the router. The desired PLR and average latency can be achieved for all three CoS types of traffic if, whilst maintaining short input buffers, the output bandwidth which is near twice as much as the input bandwidth can be provided for the input traffic.
3. If the flows are assembled in the output buffers [1][5], there are two sources for the PLR: (1) packet scheduling at the input buffers and (2) flow aggregation at the output buffers. Therefore, provisioning of the output buffer capacity $B_{\text{out}} > 1.7\ B_{\text{out}}$ guarantees that for the parameters considered, the number of lost or delayed packets at the stage (1) will be negligible.

Acknowledgments.

The authors would like to thank Prof. J. Crowcroft (UCL), Dr. M. Searle (UCL), Dr. R.J. Gibbens (Cambridge) and Prof. S. Ferguson (Marconi Communications) for invaluable comments. The authors also gratefully acknowledge support from UCL Graduate School, ORS, Marconi Communications Ltd., EPSRC, and The Royal Society.

References.

[1] M. Düser, E. Kozlovski, R. I. Killey, P. Bayvel, "Distributed router architecture for packet-routed optical networks", Proc. *"4th Working Conference on ONDM 2000*", February 2000, pp. 202-221

[2] N. McKeown, T. Anderson, "A quantitative comparison of scheduling algorithms for input-queued switches", *Computer Networks and ISDN Systems*, Vol 30, No 24, December 1998, pp 2309-2326

[3] D. K. Hunter, I. Andonovic, M. C. Chia, "Optical buffers for multi-terabit IP routers", *37th Annual Allerton Conf.*, 1999, Monticello, IL

[4] M. Jeruchim, P. Balaban, K. Shanmugan, "Simulation of communication systems", Plenum Press, New York, 1992

[5] A. Ge, F. Callegati, L. S. Tamil, "On optical burst switching and self-similar traffic", *IEEE Comm. Letters*, vol.4, No. 3, March 2000, pp. 98-100