

Enhancement of an Optical Burst Switch with Shared Electronic Buffers

Pierre Delesques^{*†}, Thomas Bonald[†], Gwillerm Froc^{*}, Philippe Ciblat[†] and Cédric Ware[†]

^{*}Mitsubishi Electric R&D Centre Europe, Rennes, France

[†]Institut Mines-Télécom / Télécom ParisTech / CNRS LTCI, Paris, France,

Email: cedric.ware@telecom-paristech.fr

Abstract—Future data networks face an energy consumption challenge: traffic grows exponentially, but the energy cost per bit in electronic routers and switches does not decrease so fast. All-optical switching techniques have not delivered a solution to this problem: despite their requiring fewer energetically-costly optical-to-electronic conversions, they suffer from poor contention handling even at low network loads, thus needing heavy overprovisioning, which negates the energy savings achieved in the first place. This contention issue largely stems from the lack of sufficiently-mature optical buffers. Thus a proposition of hybrid switch architecture supplementing optical switching with an electronic buffer.

We analyze such a hybrid switch in terms of loss probability and sustainable load. Simulations and an Engset-type analytical model both find significant performance improvements for relatively few electronic ports to/from the buffer. The highest gains are shown when few channels are available per destination.

Moreover, we note that traffic re-emitted from the buffer is a major cause of unnecessary buffering and secondary collisions. An adjustment to the re-emission policy is found to mitigate such collisions and offer slight gains on the sustainable load.

Index Terms—optical burst switching, optical packet switching, contention resolution.

I. INTRODUCTION

One of the major issues for future data networks is their ballooning energy consumption, especially in routers and switches: the traffic they carry keeps increasing exponentially, but their energy requirements per bit do not decrease so fast [1]. Routing and switching is currently performed electronically, despite the fact that transmission is now preferably optical for virtually all non-wireless data links beyond a few meters' distance. This entails multiple energetically and financially costly optical-to-electronic (O-E) conversions.

Although all-optical switching solutions have been proposed and are still being considered [2], such as Optical Packet Switching (OPS) and Optical Burst Switching (OBS), recent studies are rather pessimistic as to how much savings they could effect in practice [3]. While they do require much fewer O-E conversions, their major shortcoming is contention handling: ingress packets bound to a busy destination cannot easily be buffered, as all-optical buffering solutions are still impractical. This leads to an unacceptably high loss probability even at low load conditions [4]; for instance, the sustainable load is often under 0.2 for a target loss probability of 10^{-7} .

Consequently, networks relying on all-optical switches would have to be heavily overprovisioned; especially, increasing the number of channels per destination could help

considerably [5]. Unfortunately, using Wavelength-Division Multiplexing (WDM) channels for this purpose would require sufficient wavelength converters, which themselves consume power, largely negating the achieved energy savings [3]. Alternatively, non-wavelength-specific channels could be implemented by multiple systems running over parallel optical fibers or even cores in a multi-core fiber; but the number of available channels would then necessarily be more limited.

Therefore, as in electronic switches, buffers seem to be the key. Partial solutions have been proposed using optical delays: recirculating loops [6] or slow-light effects [7]. But they do not allow random access to stored packets/bursts, and have other drawbacks, such as memory lifetimes limited by signal degradation over propagation in a loop, or fixed bandwidth-delay products for slow-light devices.

Another possibility was recently proposed [8] and demonstrated [9]: *hybrid* switches, using OPS or OBS supplemented by an electronic buffer. This would appear to combine the best of both worlds, bringing optical switches to acceptable levels of loss probabilities; nevertheless, this approach still requires O-E conversions at the electronic input and output ports of the buffer, although fewer than a comparable all-electronic switch.

In this communication, we shall focus on the performance enhancement brought to OBS by electronic buffers, in terms of loss probability, in several configurations: number of destinations, number of channels per destination, and number of electronic ports. Note that OPS is also covered in that it is formally equivalent to OBS in our analysis.

Given a switch architecture and traffic model (Sec. II), we quantify its loss probability as a function of the system load, and the gain of sustainable load at a specific loss probability. The comparison of simulated performance (Sec. III) with closed-form expressions obtained through an Engset-type model (Sec. IV) shows that the gain in terms of sustainable load can be further improved by mitigating collisions between ingress packets/bursts and re-emitted packets/bursts; some insights on collision management are derived in Sec. V. This additional gain is a function of the re-emitting policy and has a direct impact on buffer sizing, as shown in Sec. VI.

II. ARCHITECTURE AND SYSTEM MODEL

The general switch architecture is presented in Fig. 1. Our OBS switch works in asynchronous mode: bursts can arrive at any instant. It is directly connected to n_a remote switches by as

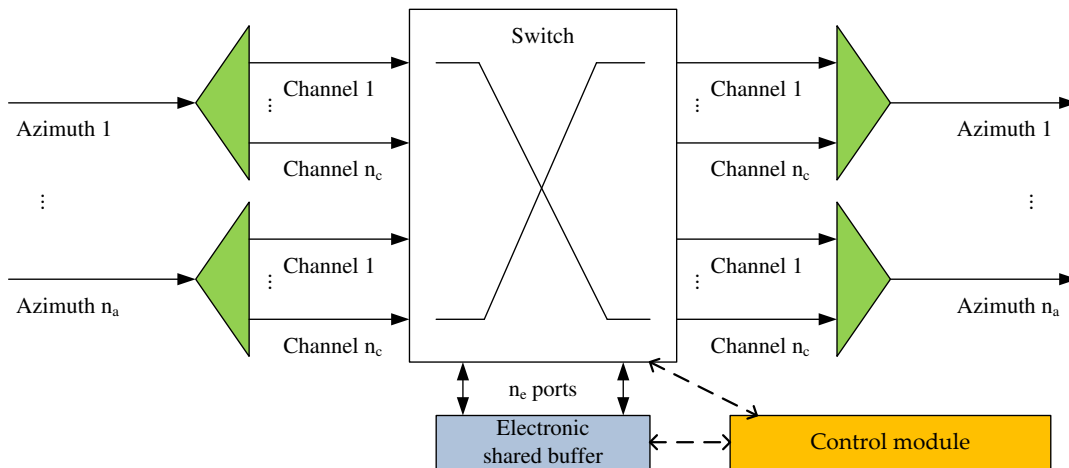


Fig. 1. Architecture of the hybrid switch.

many azimuths. Each azimuth is supposed to be bidirectional and to support n_c independent channels in each direction. We also assume that an ingress burst can indifferently use any channel of its egress azimuth. As explained in the introduction, such channels could be the cores of a multi-core fiber or the different fibers of the same cable; they may also represent WDM channels of the same fiber, at the expense of higher energy consumption through wavelength converters.

In addition, the considered OBS has a shared electronic buffer, earning it the name of hybrid switch. This overflow system permits to temporarily store bursts, if no channel is available on their egress azimuth. It has n_e input ports as well as n_e output ports.

The switch operates in the following way: when a burst arrives, its egress azimuth is obtained by the control module (*e.g.* by reading the header from its control channel). If a channel is available, the burst is directly sent on its way over it. Otherwise, if an electronic input port is available on the buffer, the burst is sent there. Otherwise the burst is dropped.

We assume that buffered bursts have priority when a channel for its destination is released. This rule minimizes the latency and the memory consumption. However, it also entails a higher use of electronic ports, which leads to dropping bursts that could have been buffered if the re-emission had been delayed. This phenomenon of so-called secondary collisions is analysed in section V, where the impact of delaying the re-transmission of buffered burst is considered.

In our analysis, we model each ingress channel of the node as an on-off process: the “on” periods correspond to the burst transmission, and the “off” periods correspond to the idle times between bursts. These idle times are assumed to be independent, exponentially distributed with mean τ . Two models are considered for the busy times:

- Pure Chance Traffic (PCT) where the burst durations are independent, exponentially distributed with mean σ .
- Traffic Shaping (TS) where burst duration is fixed to σ . Packets are grouped at each source in order to form bursts of equal duration, which are then sent over the network.

This shaping reduces the randomness of the input traffic. Furthermore, the destination of each ingress burst is chosen uniformly between all azimuths other than the ingress one. We also set $\sigma = 10 \mu\text{s}$. It represents about 100 kbit for standard 10 Gbit/s systems, and may correspond to a jumbo Ethernet frame, or an aggregation of several IP packets.

The system load ρ can be written as a function of the mean burst duration σ and the mean idle time per source τ :

$$\rho = \frac{\sigma}{\tau + \sigma}. \quad (1)$$

III. SIMULATION RESULTS

In order to evaluate the influence of the electronic buffer on performance, we perform simulations of the proposed hybrid switch in two cases: $n_a = 10$ azimuths (Fig. 2(a), 2(b), 2(c)) and $n_a = 5$ azimuths (Fig. 2(d), 2(e), 2(f)). We consider two main performance criteria: the burst loss probability; and the sustainable system load at a target loss probability of 10^{-7} .

Figures 2(a) and 2(d) show the burst loss probability versus the switch load, both for $n_c = 10$, in the cases of: full OBS ($n_e = 0$), and $n_e = 5, 10, 20$ ports on the overflow system, and for the two traffic models. First, whatever the number of electronic ports, the curves obtained under the different traffic assumptions are approximately the same. This is not surprising given the insensitivity property of the Engset model [10]. PCT is thus a good model for performance evaluation. Then, the results show that the buffer brings a significant gain in terms of achievable system load: *e.g.* in the case $n_a = 10$, the sustainable load at 10^{-7} reaches 0.6 for $n_e = 20$ electronic ports; in the case $n_a = 5$, a similar gain is observed. Furthermore, this gain even exists for a small number of ports on the buffer: *e.g.* going from full OBS to $n_e = 5$ more than doubles the sustainable load for $n_c = 10$ channels in both cases on n_a .

Figures 2(b) and 2(e) refine the previous assertion by plotting the sustainable load at a loss probability of 10^{-7} versus the number of electronic ports, for different numbers of channels per azimuth: $n_c = 1, 5, 10, 20, 50, 100$. Notice that

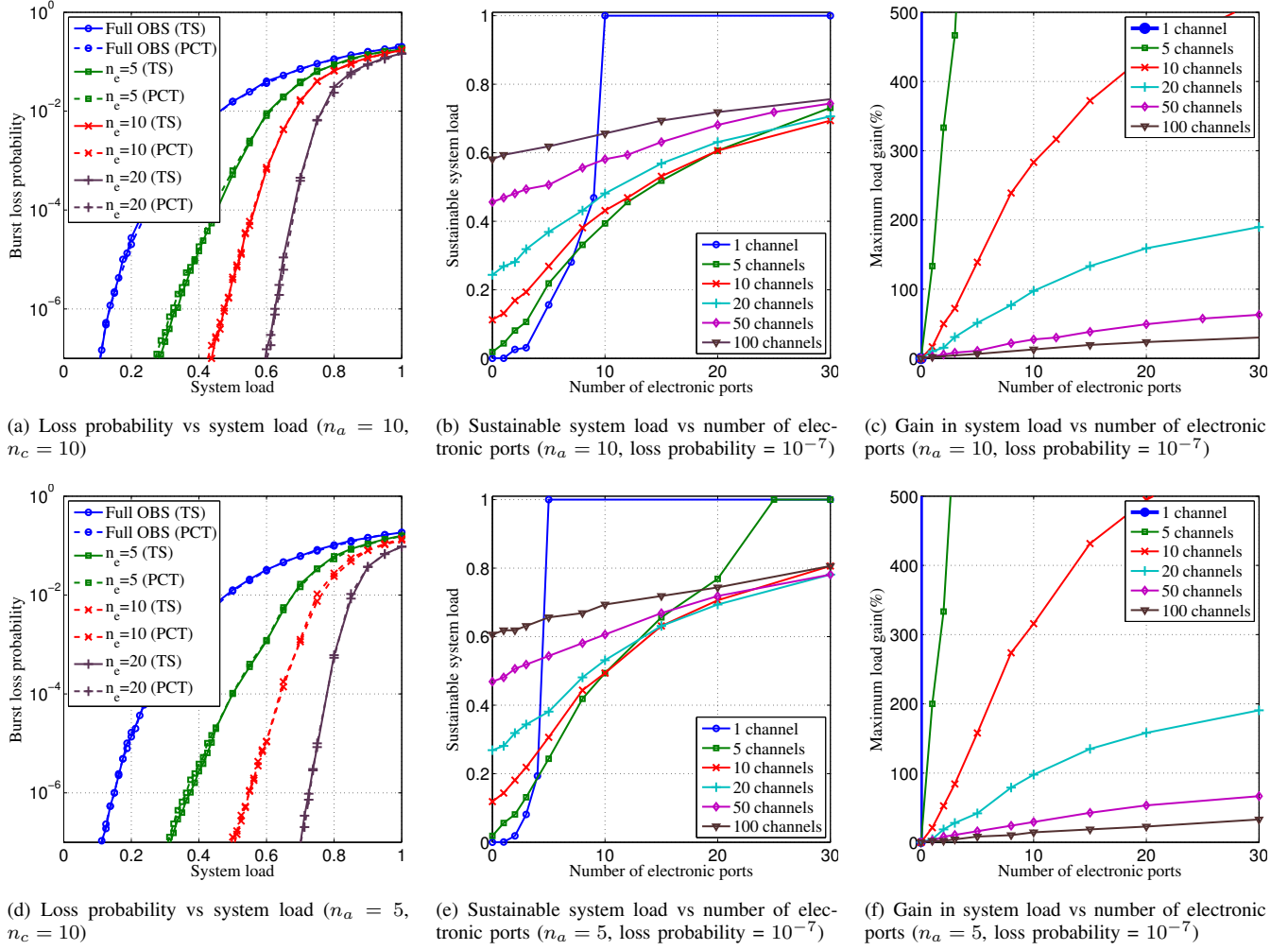


Fig. 2. Performance enhancement by addition of an electronic buffer.

the curves proposed here are for the TS model, but are almost identical to the PCT one. At low values of $n_e < 3$, the curves are rather similar for both cases on n_a ; then, as expected, the sustainable load increases with n_e , and reaches 1 for $n_e = n_a n_c$. Indeed, in this case, an ingress burst can always be collected by the buffer if it cannot be directly switched.

Choosing, as a minimum acceptable operating point, a sustainable load of 0.6, we observe that $n_e = 20$ ports on the buffer is sufficient in the case $n_a = 10$ azimuths, and that $n_e = 15$ ports on the buffer is sufficient in the case $n_a = 5$.

The influence of the electronic buffer on performance can also be measured in terms of load gain. It is defined as the percentage increase of sustainable load compared to the bufferless case. Figures 2(c) and 2(f) present this gain versus the number of electronic ports, for $n_c = 1, 5, 10, 20, 50, 100$ channels per azimuth and the two considered values of n_a . This gain is very high for small numbers of channels per azimuth: e.g. over 4000% for $n_c = 5$ and $n_e = 15$ overflow ports in both cases. In contrast, it is much more limited for high numbers of channels per azimuth. This is not surprising since the number of channels per azimuth was clearly identified as a

key parameter to mitigate the burst loss issue. Therefore, this hybrid architecture would be especially interesting when extra channels are expensive.

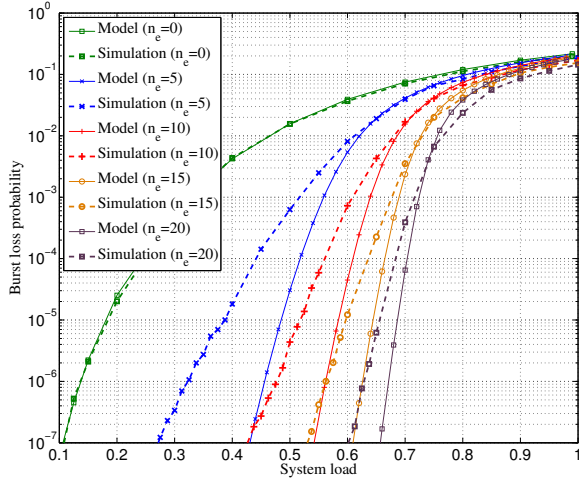
IV. ANALYTICAL APPROACH

In the previous section, the performance of the hybrid switch is analyzed by simulation. We present here an analytical approach based on a fixed-point approximation which enables a better understanding of the overall problem, and the sensitivity of the results to each parameter.

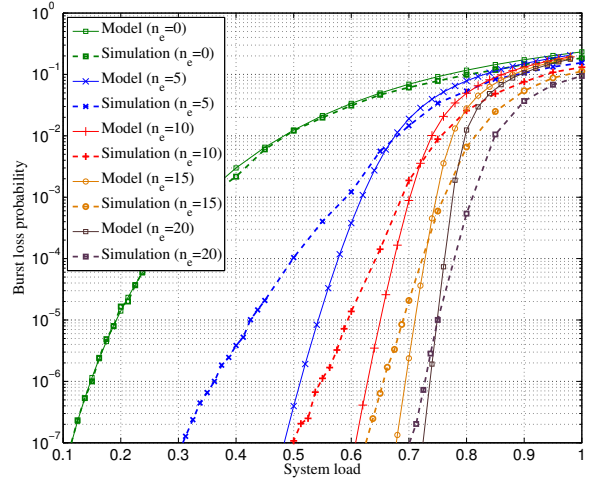
A. Fixed-point approximation

The Engset model was derived for evaluating the call loss probability of telephone network nodes as a function of: m , the number of circuits in charge of switching incoming calls (similar to our n_c); n , the number of sources connected to the system; and their calling rate a . Within this framework, the loss probability is:

$$E(n, a, m) = \frac{\binom{n-1}{m} \left(\frac{a}{1-a}\right)^m}{\sum_{k=0}^m \binom{n-1}{k} \left(\frac{a}{1-a}\right)^k} \quad (2)$$



(a) Case $n_a = 10$, $n_c = 10$



(b) Case $n_a = 5$, $n_c = 10$

Fig. 3. Comparison of closed-form expressions with simulations.

where $\binom{n}{k}$ stands for the binomial coefficient [11]. This result is insensitive to the traffic statistics beyond the traffic intensity per source a [10].

Let us first focus on the optical part of the hybrid switch. The traffic it handles has two parts: the total primary traffic $A = n_a n_c \rho$ (in Erlangs), coming from the input ports; and the total secondary traffic A_e re-emitted from the buffer, yet unknown. If bursts stored in the buffer are not retransmitted immediately, the secondary traffic can be considered independent of the primary traffic. Under this hypothesis, and assuming equal traffic intensities per source including secondary traffic, the Engset model applies¹. Since the burst destinations are uniformly chosen, each egress azimuth can be modeled as an independent Engset system with n_c channels carrying bursts. As the n_a egress azimuths are symmetrical, we thus obtain n_a equivalent systems with a traffic of $(A + A_e)/n_a$. Each system has $n = (n_a - 1)n_c + n_e$ sources (not $n_a n_c + n_e$, because of the no U-turn rule). Then the blocking probability for each system is

$$B_{\text{opt}} = E\left(n, \frac{A + A_e}{n_a n}, n_c\right). \quad (3)$$

The total traffic arriving at the electronic buffer's input ports is then $(A + A_e)B_{\text{opt}}$. Furthermore, these n_e electronic ports can be globally modeled as an Engset system too, with $n_a n_c$ sources and n_e channels. We deduce its blocking probability:

$$B_e = E\left(n_a n_c, \frac{(A + A_e)B_{\text{opt}}}{n_a n_c}, n_e\right). \quad (4)$$

Conversely, the traffic going through the electronic ports is $(A + A_e)B_{\text{opt}}(1 - B_e)$. This must be equal to the secondary traffic, so that:

$$A_e = (A + A_e)B_{\text{opt}}(1 - B_e). \quad (5)$$

¹Note that the Engset model is classical in the performance analysis of optical networks, see for instance [5], [12].

Equations (3), (4), (5) constitute the fixed-point approximation, which allows us to determine A_e . The total fraction of lost traffic (that is, blocked at the electronic ports), is then:

$$P_{\text{loss}} = \frac{(A + A_e)B_{\text{opt}}B_e}{A}. \quad (6)$$

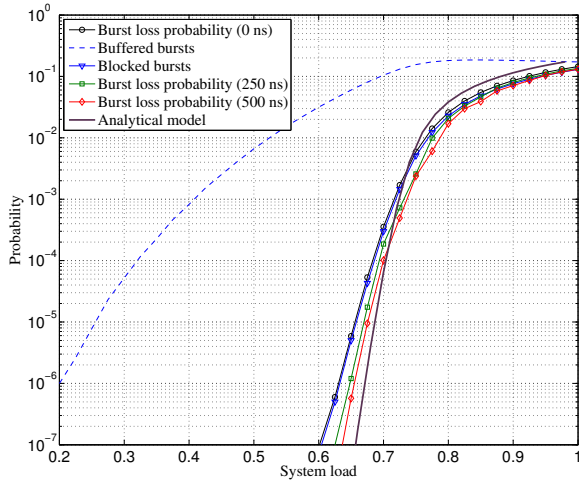
B. Numerical illustrations

In order to evaluate the accuracy of the approach, we plot on the same graph analytical and simulated curves of the burst loss probability versus the system load for different numbers of electronic ports $n_e = 0, 5, 10, 15, 20$, and in the two cases: $n_a = 10$ azimuths with $n_c = 10$ channels each (Fig. 3(a)), and $n_a = 5$ azimuths with $n_c = 10$ channels each (Fig. 3(b)).

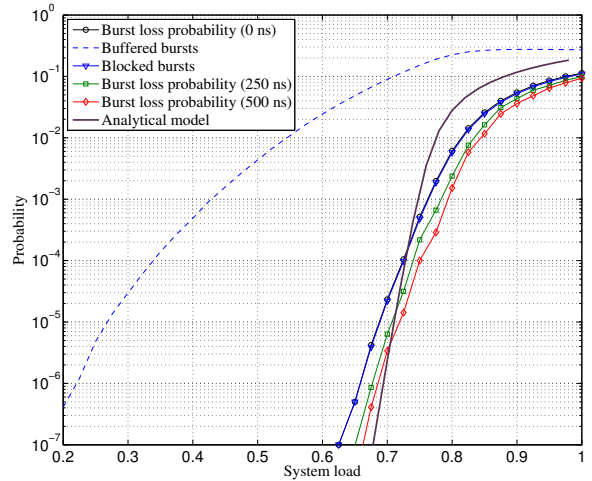
First, we observe that the bufferless switch ($n_e = 0$) behaves very closely to what the pure-Engset model predicts for both values of n_a ; in this case, Eqs. (5) and (6) reduce to $A_e = 0$ and $P_{\text{loss}} = B_{\text{opt}}$. The fit is not so good for $n_e > 0$, especially at lower loads, although the model seems fairly accurate at high loads for the larger number of azimuths.

The main reason is that the secondary traffic cannot be considered independent of the primary traffic. Indeed, those bursts re-emitted from the buffer have a destination azimuth that suffers congestion, otherwise said bursts would not have been buffered in the first place.

Another phenomenon must also be considered: so-called secondary collisions between primary and secondary traffic. In our simulations, a stored burst is re-emitted as soon as a channel for its destination is released. If another burst comes into the switch shortly thereafter, it will have to be stored—or worse, dropped. This could lead to burst losses that could have been avoided by *e.g.* delaying the re-emission; it would then be not surprising that our simulated loss probability be higher than predicted by the model. Section V will show that secondary collisions do indeed occur, and that a different re-emission policy improves system performance.



(a) Case $n_a = 10, n_c = 10, n_e = 20$



(b) Case $n_a = 5, n_c = 10, n_e = 15$

Fig. 4. Impact of the re-emitting policy on the burst loss probability.

V. MITIGATING SECONDARY COLLISIONS

As explained above, the re-emitting process creates so-called secondary collisions between the re-emitted bursts and the ingress bursts, possibly leading to a higher-than-necessary buffer usage and congestion in the overflow system, hence burst losses. Additionally, from an energy savings point of view, these collisions can be considered useless O-E conversions and their reduction can improve the energy efficiency of the architecture.

Fig. 4 proposes a number of results about secondary collisions versus system load in two cases: $n_a = 10$ azimuths with $n_c = 10$ channels each and $n_e = 20$ ports on the overflow system in Fig. 4(a), and $n_a = 5$ azimuths with $n_c = 10$ channels each and $n_e = 15$ ports on the overflow system in Fig. 4(b).

We first focus curves denoted “Buffered bursts” and “Blocked bursts”. The former is the probability that an ingress burst is buffered because the buffer is re-emitting bursts on its egress azimuth. The latter is the probability that an ingress burst is lost for the same reason.

The results show that an important part of the traffic goes through the buffer because of secondary collisions, even for low values of the system load: *e.g.* for a load of 0.4 this probability is on the order of 10^{-4} to 10^{-3} . Unsurprisingly, then, the “Blocked bursts” graph is very close to that of the burst loss probability, showing that almost all burst losses are in fact due to secondary collisions. Therefore, there could be a significant margin of improvement on performance by avoiding them. Failing that, even reducing their occurrences could have a notable impact on the number of O-E conversions required, hence the energy footprint of the hybrid switch.

Hopefully, some secondary collisions can be mitigated by enhancing the re-emitting process. In practice, since the switching fabric needs time for its reconfiguration, the control informations of a burst are read in advance by adding a fiber

delay line (FDL) on each input. If this FDL is extended, it introduces an observation window; in other words, the control module can then anticipate what will happen on each ingress channel. The decision to re-emit a burst can thus be delayed in order to avoid a collision.

In Fig. 4, we also propose results obtained in this context. We choose two observation windows: 250 ns and 500 ns which respectively correspond to 50 m and 100 m of standard single-mode fiber at 10 Gbit/s.

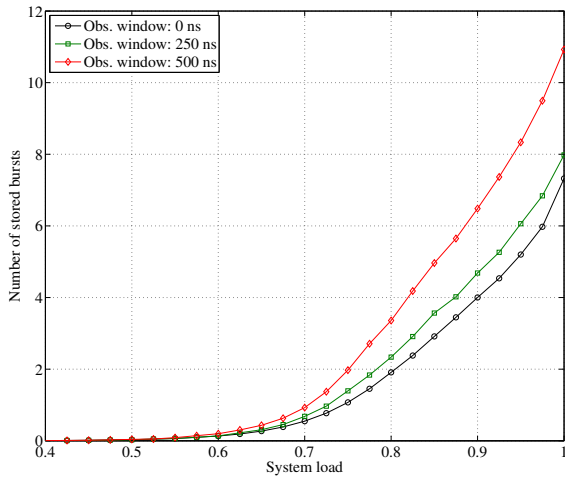
We observe a moderate performance improvement, especially for low values of the loss probability: the sustainable load at a loss probability of 10^{-7} rises by 0.03 for the biggest observation window in the first case, and by 0.05 in the second one. As expected, the longer the observation window, the greater the gain. Perhaps more significantly, simulated performance comes closer to that predicted by the model at low loads. The model may thus give an estimation of what margin of progression could be gained by optimizing re-emission policy to avoid secondary collisions.

However, by lengthening the time bursts spend in the buffer, this mechanism will directly impact the memory consumption of the architecture, as shown in the next section.

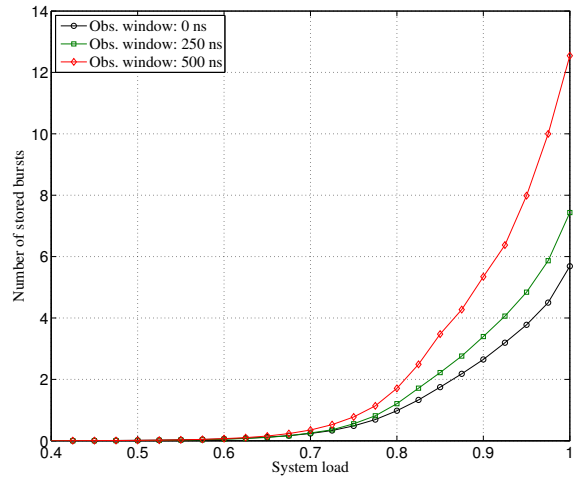
VI. SIZING THE ELECTRONIC BUFFER

The electronic buffer allows random access to stored bursts, and thus to optimize re-emission according to the output ports’ availability. However, this high-performance RAM has a non-negligible cost, especially in large amounts. In addition, the secondary collision mitigation mechanism proposed in the previous section will increase this requirement, since a burst may remain longer in the buffer.

We therefore study the mean number of bursts stored in the buffer. The obtained results are displayed in Fig. 5. Figures 5(a) and 5(b) respectively are obtained in the same context as previously: $n_a = 10, n_c = 10,$ and $n_e = 20$ for the first one, and $n_a = 5, n_c = 10,$ and $n_e = 15$ for the



(a) Case $n_a = 10$, $n_c = 10$



(b) Case $n_a = 5$, $n_c = 10$

Fig. 5. Impact of the re-emitting policy on the RAM occupation (mean number of stored bursts).

second one. Three observation windows are considered: no observation window, 250 ns and 500 ns. The curves show the mean number of stored bursts versus the system load.

We observe that the mean number of stored bursts is not so severe, even for high values of the load, around a tens of bursts. For load values where the loss probability is lower than 10^{-7} , this mean is lower than 1 in both cases.

VII. CONCLUSION

Aiming to find a middle ground between the high energy consumption of electronic switches and the poor contention handling of buffer-less OBS, we have studied a hybrid switch architecture that supplements OBS with a shared electronic buffer. In terms of loss probability and sustainable load, we find significant improvements for relatively few electronic ports on the buffer, both via numerical simulations and an analytical Engset-type approach. The highest gains are shown when the number of channels per azimuth is limited, which is often the case for financial and/or energetic reasons.

We also remark that even at low loads, a significant part of the ingress traffic has to be buffered because of traffic re-emitted from the buffer, incurring unnecessary O-E conversions. Accordingly, most remaining burst losses are in fact secondary collisions due to re-emission. Techniques which limit these secondary collisions, such as an observation window at the switch input, offer slight gains on the sustainable load.

The analysis of the mean number of stored bursts in the buffer indicates that the required RAM is limited, but the mitigation techniques of secondary collisions have a significant impact on this amount. In practice, a trade-off has to be found between the number of channels per azimuth, the number of ports on the buffer, and the size of the observation window.

The proposed hybrid switch is a promising technique for reducing the energy consumption of packet-switched networks. Future work is required to assess the exact energy savings

brought by this equipment, through the analysis of the fraction of traffic which must be handled electronically. A future, more accurate, analytical model will also provide network designers with a useful dimensioning tool.

REFERENCES

- [1] R. S. Tucker, "Green optical communications—part II: Energy limitations in networks," *IEEE J. Sel. Topics Quantum Electron.*, vol. 17, pp. 261–274, 2011.
- [2] S. J. Ben Yoo, "Energy efficiency in the future internet: the role of optical packet switching and optical-label switching," *IEEE J. Sel. Topics Quantum Electron.*, vol. 17, no. 2, pp. 406–418, 2011.
- [3] R. S. Tucker, "Optical packet switching: A reality check," *Optical Switching and Networking*, vol. 5, pp. 2–9, 2008.
- [4] Z. Rosberg, H. Le Vu, M. Zukerman, and J. White, "Performance analyses of optical burst-switching networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1187–1197, 2003.
- [5] J. Zhang, E. Wong, and M. Zukerman, "Modeling an OBS node under critical load and high utilization conditions," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 544–546, 2012.
- [6] Y. Liu, M. Hill, R. Geldenhuys, N. Calabretta, H. de Waardt, G. Khoe, and H. Doreen, "Demonstration of a variable optical delay for a recirculating buffer by using all-optical signal processing," *IEEE Photon. Technol. Lett.*, vol. 16, no. 7, pp. 1748–1750, 2004.
- [7] R. S. Tucker, P. Ku, and C. J. Chang-Hasnain, "Slow-Light Optical Buffers: Capabilities and Fundamental Limitations," *J. Lightw. Technol.*, vol. 23, pp. 4046–4066, 2005.
- [8] R. Takahashi, T. Nakahara, K. Takahata, H. Takenouchi, T. Yasui, N. Kondo, and H. Suzuki, "Ultrafast optoelectronic packet processing for asynchronous, optical-packet-switched networks [invited]," *J. Opt. Netw.*, vol. 3, no. 12, pp. 914–930, 2004.
- [9] D. Chiaroni, R. Urata, J. Gripp, J. E. Simsarian, G. Austin, S. Etienne, T. Segawa, Y. Pointurier, C. Simonneau, Y. Suzuki *et al.*, "Demonstration of the interconnection of two optical packet rings with a hybrid optoelectronic packet router," in *European Conference on Optical Communications (ECOC)*, no. PD3.5, 2010.
- [10] T. Bonald and M. Feuillet, *Network performance analysis*, Wiley ed., 2011.
- [11] T. Engset, "On the calculation of switches in an automatic telephone system," in *Tore Olaus Engset: The man behind the formula*, A. Myskja and O. Espvik, Eds., 1998.
- [12] H. Overby, "Performance modelling of optical packet switched networks with the Engset traffic model," *Optics Express*, vol. 13, no. 5, pp. 1685–1695, 2005.