

## Performance overview of the quasi-synchronous operation mode in optical burst switching (OBS) networks

O. Pedrola<sup>a,\*</sup>, S. Rumley<sup>b</sup>, D. Careglio<sup>a</sup>, M. Klinkowski<sup>c</sup>, C. Gaumier<sup>b</sup>, J. Solé-Pareta<sup>a</sup>

<sup>a</sup> CCABA, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain

<sup>b</sup> Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

<sup>c</sup> National Institute of Telecommunications (NIT), 04-894 Warsaw, Poland

### ARTICLE INFO

#### Article history:

Received 11 December 2009

Received in revised form 18 June 2010

Accepted 22 June 2010

Available online 30 June 2010

#### Keywords:

Optical burst switching (OBS)

Quasi-synchronous (QS)

Analytic model

Performance analysis

Deflection routing

### ABSTRACT

Optical burst switching (OBS) is able to operate under both asynchronous and synchronous burst departures. Although the asynchronous approach is generally assumed in OBS owing to its simplicity and low technical requirements, it has been shown that performance improvements, in terms of the overall burst loss probability, can be achieved under synchronous operation. Moreover, when effective contention resolution mechanisms are applied, such improvements are significant. Since perfect synchronization is not a viable solution, we propose a novel operation mode with relaxed technical requirements, the quasi-synchronous (QS) operation.

In this paper, the performance of the QS operation mode is evaluated by an exact analytic expression, assuming a Poisson burst arrival process and a single-wavelength scenario. This mathematical analysis is completed with numerical results that validate its accuracy. Furthermore, we tackle the issue of not having well-aligned clock information between network nodes by proposing a novel re-synchronization mechanism for OBS networks. Finally, a performance study of the QS operation under different deflection routing techniques is presented. The results show that our novel QS approach achieves significant performance benefits with respect to the asynchronous operation while avoiding the architectural complexity of the synchronous scheme.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction and motivation

Two main features distinguish optical burst switching (OBS) [1] from other optical switching technologies: the transmission of large data bursts, which are aggregated at the edge of the network, and the possibility to establish a path dynamically and on-the-fly (i.e., without acknowledgment of the availability of transmission resources). Because of the absence of optical buffering capabilities, the main challenge of OBS is to deal with high burst losses that arise due to the contention of bursts transmitted in the network.

To mitigate the burst contention problem, solutions have been proposed based on deflection (or alternative)

routing. All these methods allow re-routing contending bursts from primary to alternative routes and, by these means, alleviating congestion on bottleneck links and achieving dynamic load balancing in the network. In this paper, we consider the so-called offset time emulated OBS (E-OBS) network architecture [2], which facilitates the application of alternative routing since routing decision can be taken freely inside the network without constraints on the length of routing path.

In principle, the transmission of optical bursts is asynchronous in an OBS network. That means that bursts are not aligned with each other, and thus, they arrive at a core switching node in casual instances of time. Performance improvements can be achieved if synchronous operation is applied: in fact, in such a case, contention may only occur between entire data units, and better transmission

\* Corresponding author. Tel.: +34 93 401 7182.

E-mail address: [opedrola@ac.upc.edu](mailto:opedrola@ac.upc.edu) (O. Pedrola).

resource utilization can be obtained with simple contention resolution mechanisms [3]. Such synchronous operation was proposed in the past for optical packet switching (OPS) networks (see e.g., [4]), and although it has not been widely considered in OBS networks, some relevant studies can be found in the literature (see e.g., [5–9]). Besides, in [10], we verified that, with the utilization of effective deflection routing techniques, a performance gain of the synchronous operation with respect to the asynchronous one is brought about to a very motivating extent. However, all these studies consider that synchronization of data bursts is achieved by means of an input stage at each core node, which involves the use of additional hardware elements such as incremental fibre delay blocks and switching devices. All these components and the increased control complexity that they entail lead to a bulky and complex structure whose viability has not been demonstrated yet.

In this direction, we proposed a novel operation mode for OBS networks, quasi-synchronous (QS) operation, with the aim of reaching performance benefits close to those obtained with perfect synchronization while keeping a moderate hardware complexity. In the QS-OBS scenario, we do not have the need for any intricate synchronization device. In contrast, we assume both that network links are designed such that the resulting propagation delays correspond to a multiple of a given fixed time slot duration and that bursts are released only at the beginning of such time slots. In addition, to take into account that perfect synchronization is practically impossible, we accept the presence of some time deviation. The existence of this time variation between the actual arrival of bursts and the beginning of time slots causes the bursts to be not perfectly aligned at the core switching nodes. For that, we refer to the QS operation.

Specifically, we model the time deviation of the QS transmission mode as a superposition of two different sources, denoted as drift and skew, respectively. We consider the drift as an irretrievable error due to structural inaccuracy of the different devices constituting an OBS node and to the physical impairments that may change the propagation time characteristics of the different channels of a fibre. In contrast, we consider the skew as the consequence of not having well-aligned hardware clock information amongst all network edge nodes. The skew is, hence, a retrievable time deviation error and, for that reason, we propose a method to bound its value to a range where there is no (or negligible) performance degradation.

The contribution of this paper is twofold. First, we present the details of the QS scheme, highlighting its architectural benefits compared to the synchronous and asynchronous cases. In particular, we provide a detailed performance study (through both analytic and simulation models) both to analyse the effects of the presence of the drifts and skews and to find a set of values that highlights the performance of the QS scheme. Second, we compare the performance of the asynchronous, synchronous and QS operation modes with the support of different deflection routing algorithms.

The rest of this paper is organized as follows. In Section 2, we provide complete information on the analysed network scenarios. In Section 3, we first present in

detail the analytic models for all the OBS transmission modes considered in this paper. Afterwards, we validate the models by simulation, with special emphasis on our novel analytic model for the QS operation in OBS networks. Then, in Section 4, we focus on the key parameters in the QS-OBS network and provide values for them that are in accordance with real OBS network scenarios. In Section 5, we present the re-synchronization mechanism that we propose with the purpose of maintaining such quasi-synchronization at edge nodes. Section 6 presents an overall comparison between the synchronous, asynchronous and our QS schemes under different deflection routing algorithms. Finally, the conclusions are presented in Section 7.

## 2. OBS network scenarios

In this study, we consider an E-OBS network scenario [2]: core switching nodes are enhanced with a pool of fibre delay coils that is inserted into the data path at the input port of the node. In conventional OBS (C-OBS) architectures, the processing offset time is provided at the edge node by delaying the transmission of a burst with respect to its control packet. In E-OBS, in contrast, the offset times are provided by means of these fibre delay coils at each core node. Thus, both the burst and control packet can be sent together from the edge node, avoiding several problems that result from the offset time variation inside the network, a feature that is inherent to C-OBS. For instance, concerning routing management, it is advantageous to provide offset times at each core node since routing decisions can be taken freely inside the network without any constraints on the length of the path. In contrast, in C-OBS, the maximal length of the routing path is related to the offset which, once introduced at the edge node, decreases at each hop.

In general, the transmission of optical data in the network can be either asynchronous or synchronous. Although both approaches have been studied extensively in the context of OPS networks, the research on OBS still concerns mostly the asynchronous approach.

### 2.1. Asynchronous OBS

In asynchronous OBS networks, optical bursts are released from edge nodes at arbitrary (random) instances of time, and they are not aligned when they arrive at core switching nodes (see Fig. 2.1A). Accordingly, the switching operation is performed asynchronously. The advantage of this approach is the simplification of the burstification process and the low complexity of the switching nodes.

### 2.2. Synchronous OBS

The idea of synchronous, or time-slotted, transmission in optical networks has been considered mainly in the context of OPS networks. In the case of a synchronous OBS scenario, the optical bursts are aligned and transmitted synchronously at the beginning of a fixed-duration time slot (see Fig. 2.1B). The main advantage of the synchronous approach, with respect to the asynchronous approach, is the improvement of the overall burst loss performance.

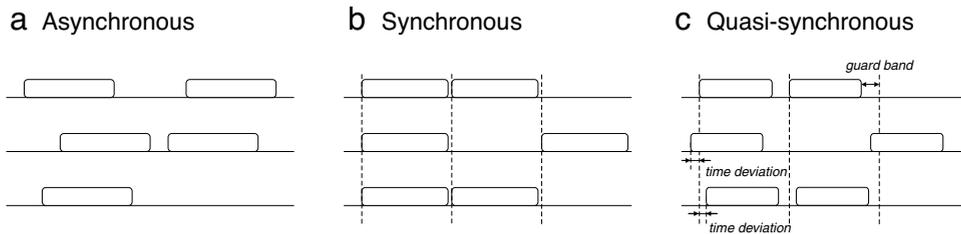


Fig. 2.1. OBS transmission modes.

In synchronous OBS networks, since bursts travelling over different length links may still arrive at a core node at different instances of time, their synchronization is achieved by means of a specialized node input interface. For example, in [7], a time-slotted OBS architecture called time-sliced OBS, which performs the switching in the time domain rather than in the wavelength one, is proposed. Such operation is achieved by means of a two-block stage: (a) a synchronizer consisting of variable delay lines which require feedback control information from the node controller; (b) an optical time slot interchanger (OTSI) consisting of optical crossbars and delay lines. However, as the authors conclude, there is still the need to undertake substantial additional work in order to prove the viability of this solution. Besides [7], no further works address such time-slotted OBS solution.

On the other hand, in [9] and [11], a variant of the time-slotted OBS operation is proposed under the name of time-synchronized OBS (SynOBS). The authors assume that the length of both bursts and slots is fixed, as we do in our QS-OBS. In SynOBS, synchronization is performed by means of time slot synchronizers and wavelength delay variation compensators. The synchronizers consist of cascaded variable-length delay lines, also known as tunable delay lines, and a number of  $2 \times 2$  switching devices. Their duty is to synchronize incoming data bursts from different input ports. Examples of such a mechanism can be found in [12]. The compensators, by contrast, are used to balance the different speeds at which wavelengths travel in a fibre. Assuming that the impact of delay variations among wavelengths can be neglected, these devices may be implemented with fixed-length delay lines. This is not, however, the case with tunable delay lines, whose design entails a rather complex control and structure. Note that the node controller is in charge of dynamically adjusting the delays applied to data streams from different input ports so that they all become synchronized. Furthermore, in the  $2 \times 2$  switches, high switching speeds (i.e., in the nanosecond range) are essential, and hence, semiconductor optical amplifier (SOA) switches are required. Both the cost that this structure entails (i.e., substantial extra hardware is necessary) and the impact that the added physical layer impairments have on the optical signal (i.e., SOA amplifiers also bring out some non-desirable effects such as power consumption, noise and nonlinearity) must be thoroughly evaluated before this solution can be considered viable.

In contrast, in our QS-OBS network operating under the E-OBS control architecture there is no need for such extra hardware. In E-OBS, pools of fixed-length delay coils

are used at each core node to provide bursts with enough time to compensate the processing and switching times (i.e., what electrical memories do at edge nodes in C-OBS). This provided time depends on the length of the delay coil exclusively and cannot be tuned; it is fixed and defined in the design phase of the network (we refer to [2] for more details). In addition, this pool of delay coils can act as a dispersion compensation unit to mitigate the chromatic dispersion.

### 2.3. Quasi-synchronous OBS

In the QS scenario, there are no synchronization devices; the network links are designed so that the resulting propagation delays correspond to a multiple of the slot size; the edge nodes are synchronized with each other and release data bursts at the beginning of time slots.

Since it is impossible to have all edge nodes perfectly synchronized, we assume that there exists some skew amongst their clocks. Moreover, we consider the presence of a local drift which accounts for the device inaccuracies and physical impairments. A possible representation of this scenario is shown in Fig. 2.2, where we model the drift as a normal random variable and the skew as a uniform random variable. Therefore, and due to the effect of both the drift and the skew, bursts arriving at core nodes are not perfectly aligned (see Fig. 2.1C), with the consequence of having performance degradation compared to the perfectly synchronized case.

In order to prevent the overlapping of de-synchronized bursts, and given that both sources of time deviation are independent, we propose two separate solutions for their adjustment. The problem can, therefore, be split into two different parts.

On the one hand, to correct the effect of the skew, we present in Section 5 a re-synchronization mechanism adapted to OBS networks. The basis of our mechanism was proposed in the past in [13] for distributing well-aligned hardware clock throughout the physical extent of a synchronous processor. Its scope is to bound the skew effect to a range of values that guarantees good performance results.

On the other hand, to overcome the problem posed by the drift present at the instant of burst departure, we introduce a guard band between the bursts so that they do not occupy the slots completely. This guard-time should be large enough to maintain (as much as possible) the burst alignment as in the case of synchronous operation but not too large in order to not decrease the bandwidth

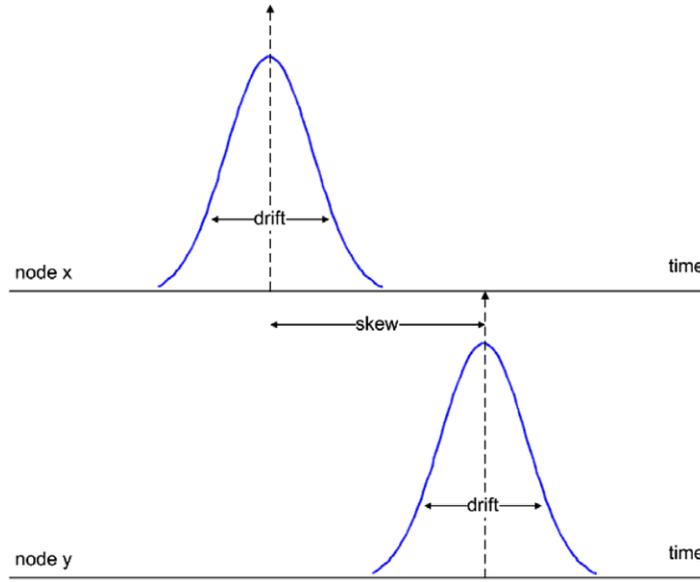


Fig. 2.2. Time deviation sources in the QS-OBS network.

utilization excessively. As a consequence, the duration of the guard-time results in a clear trade-off between utilization and performance. Notice that, in any mode of operation (i.e., asynchronous, synchronous and QS), a time margin must be included between every two consecutive bursts to allow for switching the bursts in intermediate nodes (both to process the information and to physically reconfigure the switches). Let us denote this margin as the basic guard-time. Since all modes of operation require such basic guard-time, it affects all of them equally. For this very reason, we neglect it in all scenarios considered in this paper for the sake of simplicity. Nonetheless, we do consider that this basic guard-time has a fixed value only in the synchronous approach because its input node interfaces can re-align the bursts perfectly. In the asynchronous and QS cases, by contrast, such basic guard-time is not necessarily maintained constant since burst arrivals may be affected by time variations and, as a result of this, the basic guard-time may be either increased or reduced. Such variations are what we define as drift. For this very reason, in the QS operation mode, we propose and evaluate the use of an additional guard-time whose purpose is to minimize the impact of the drifts on the network performance. Hereinafter in this paper, we only consider, if not explicitly stated, the additional guard-time, and thus, we will shorten it to guard-time.

Note that, in contrast to OPS networks, where optical packets have small size and even their small misalignment might result in a serious degradation of performance, in OBS there is possibly a higher margin for de-synchronization of burst transmission owing to much larger burst durations and, as a result, much larger guard bands. For these reasons, we expect that the results obtained with the QS operation should be somewhere between those of the asynchronous and synchronous cases.

In the next section, we present an analytic model of the drift deviation and study its impact on the node and

network behaviour. For the sake of simplicity, the analytic model presented relies on exponentially distributed drifts. Analytic and simulation results comparing the performance of the QS scheme with those of both the synchronous and asynchronous schemes are also provided. Afterwards, in Section 4 we make use of a more realistic model of the drifts (i.e., Gaussian-distributed drifts may fall on either side of the node clock pulse) and provide practical values for the time slot, guard-time and drift so that in subsequent sections the QS-OBS network performance can be effectively evaluated.

### 3. Evaluation of the drift impact

To evaluate the impact of the drift on the performance of the QS operation mode, we present a detailed performance study (through both analytic and simulation models) of the three different OBS transmission modes considered in this paper, namely asynchronous, synchronous and QS. First, we consider a single OBS node with  $W$  wavelengths in its output port  $p$ . Then, we extend the models to a network scenario using the reduced-load fixed point approximation proposed in [14]. It is worth pointing out that both the analytic and simulation studies presented in this section are independent of the time slot size. Hence, all drifts and guard-time values are given as a percentage of the time slot size. An analysis assuming absolute values for drifts (and consequently also absolute values for the time slot and guard-time) will be presented later in Section 4.

#### 3.1. Analytic models

We present the asynchronous, synchronous and QS analytic models, with special emphasis on our single-wavelength QS model. For the sake of mathematical tractability, let us assume in our analyses that burst arrivals can be modelled as a Poisson process with parameter  $\lambda$ .

We consider fixed burst size in all cases. For simplicity, we assume the service time  $\mu$  of bursts to be equal to 1, and, consequently, the traffic intensity  $\rho$  to be equal to  $\rho = \lambda/\mu = \lambda$ . Note that, whilst in the synchronous case the burst size is equal to the slot size, in the QS case it is decreased by the guard-time value used. In the next subsections we introduce the analytic models corresponding to each operation mode considered.

### 3.1.1. Asynchronous loss model

Since we are considering Poisson burst arrivals, it is well known (see, e.g., [15]) that an output port  $p$  can be modelled as a  $W$ -server loss system, and thus its blocking probability ( $Pb_{ASYN}$ ) under traffic intensity  $\rho$  is given by the following Erlang-B formula:

$$Pb_{ASYN}(\rho, W) = E(\rho, W) = \frac{\rho^W / W!}{\sum_{i=0}^W \rho^i / i!}. \quad (3.1)$$

### 3.1.2. Synchronous loss model

Under synchronous operation, bursts are only released at the exact instant of the slot start time. Due to the fact that the number of burst arrivals at each slot is Poisson distributed, we can model the probability of having exactly  $k$  burst arrivals during a slot  $N_j$  as follows:

$$P(k \text{ arrivals}) = P(N_j = k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (3.2)$$

Notice that, assuming a perfectly synchronized scenario, burst loss can only occur when within a slot time the number of burst arrivals is higher than the number of available wavelengths at  $p$  (i.e., higher than  $W$ ). Hence, we can analytically model the synchronous blocking probability ( $Pb_{SYNC}$ ) as follows [16]:

$$Pb_{SYNC}(\rho, W) = \frac{1}{\rho} \sum_{i=W+1}^{\infty} P(N_j = i)(i - W). \quad (3.3)$$

Each of the three terms in the formula models a different factor. Whilst  $P(N_j = i)$  represents the probability of an event to happen,  $(i - W)$  counts the number of bursts undergoing such an event. Eventually,  $\frac{1}{\rho}$  acts as a traffic normalization factor. Indeed, as long as  $\rho < 1$ , it corresponds to the server utilization or occupation rate. For example, this factor is responsible for capturing the impact that a reduction in the service time of bursts would have on the loss probability.

### 3.1.3. Quasi-synchronous loss model

In this section, we present our single-wavelength analytic model for the QS operation. In the QS scenario, each edge node generates bursts with a certain time deviation. This deviation can be modelled as a superposition of a skew and a drift. In this section, we only model the presence of the drift. Although modelling the drift as a normal random variable is a more realistic option, in our model, we assume, for the sake of simplicity, that the drifts follow an exponential distribution with parameter  $\alpha$ . We also consider the presence of a guard-time of value  $\sigma$  at the end of bursts.

Under QS network operation, burst loss can be caused by two different factors. The first loss factor is the consequence of a drift-based collision (i.e., a collision between two bursts that overlap in two consecutive slots due to incompatible drift values (see e.g., Fig. 3.1C). In the cases represented in Fig. 3.1A and B there is no possibility for a drift-based collision to occur. Notice that in the case 3.1B, the condition  $d' < \sigma$  is enough to guarantee that there is not a drift-based collision between slots  $N_1$  and  $N_2$ ; however, by adding the condition  $d'' < \sigma$ , we also prevent a possible drift-based collision with the next slot (i.e.,  $N_3$ ). The second loss factor results from an overflow-based collision (i.e., the arrival of more bursts than wavelengths available at  $p$  during one time slot). To model the drift-based collisions, we consider arrivals in two consecutive time slots, namely slot  $N_1$  and slot  $N_2$  in Fig. 3.1. Since arrivals are modelled using a Poisson process, we are able to estimate the probability of arrivals in two consecutive time slots as a simple product thanks to the independence between time intervals. To ease the mathematical development, we split the computation QS blocking probability ( $Pb_{QS}$ ) into two components depending on the number of arrivals in the first slot, giving the following structure:

$$Pb_{QS}(\rho, 1) = P_{QS0} + P_{QS1}, \quad (3.4)$$

where  $P_{QS0}$  refers to the case where there are no burst arrivals in the first slot  $N_1$ , and  $P_{QS1}$  refers to the case where there is a positive number of arrivals at  $N_1$ . Note that  $P_{QS0}$  corresponds to a synchronous loss model in a single-wavelength scenario.

As aforementioned, in the QS operation mode, there are two different classes of burst loss, namely drift-based collisions and overflow collisions. It is important to notice that, in a perfectly synchronized scenario, where only overflow collisions exist, burst  $W + 1$  and any subsequent arrivals are lost with a probability equal to 1 because no more than  $W$  wavelengths are available. However, in the QS mode, this is not true at all since earlier arrivals may be lost due to drift-based collisions. In our analytic model, we take this fact into account, and the term  $B_r$  accounts for such reduced probabilities. Eventually, the formulation for the two components in Eq. (3.4) corresponds to

$$\begin{aligned} P_{QS0} &= P(N_1 = 0)Pb_{SYNC}(\rho, 1) \\ &= \frac{P(N_1 = 0)}{\rho} \sum_{l=2}^{\infty} P(N_2 = l)(l - 1) \\ P_{QS1} &= \frac{1}{\rho} \sum_{k=1}^{\infty} P(N_1 = k) \sum_{l=1}^{\infty} P(N_2 = l) \\ &\quad \times \left( \sum_{n=1}^l A_n + \sum_{r=1}^{l-1} B_r \right), \end{aligned} \quad (3.5)$$

where  $k$  and  $l$  refer to the number of burst arrivals at slot  $N_1$  and slot  $N_2$ , respectively. Besides,  $A_n$  represents the burst loss probability caused by drift-based collisions for the burst arrival number  $n$  at slot  $N_2$ , and  $B_r$  denotes the burst loss probability caused by overflow-based collisions for the burst arrival number  $r + 1$  at slot  $N_2$  (i.e., notice that the first burst arrival cannot be lost due to overflow). Since

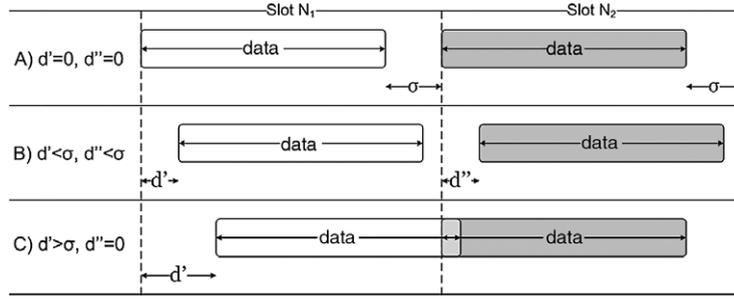


Fig. 3.1. Exponential drift collision scenario.

the  $B_r$  values are dependent on those of  $A_n$ , we are able to derive them as follows:

$$B_r = 1 - \overline{B}_r \quad (3.7)$$

$$\overline{B}_r = (1 - A_{r+1}) \prod_{i=1}^r A_i. \quad (3.8)$$

In Eq. (3.8), we can observe that the burst arrival  $r + 1$  at slot  $N_2$  will not be lost due to overflow if and only if all the preceding arrivals have been lost due to drift-based collisions ( $\prod_{i=1}^r A_i$ ) and it is not lost owing to its own drift ( $1 - A_{r+1}$ ). Hence, our goal in this study is to analytically model  $A_n$ .

If we take a look at Fig. 3.1, we observe that a collision will occur if and only if the drifts corresponding to two consecutive reservations fulfil the following condition:

$$d' > d'' + \sigma. \quad (3.9)$$

Our objective is to find an analytic model for the calculation of the burst loss probability, taking into account losses that are the direct consequence of the cases where condition (3.9) is fulfilled.

Let us start by considering the random variables  $Z$  and  $U$  that correspond to  $d'$  and  $d''$ , respectively. Since we consider a single-wavelength scenario, our interest lies in the minimum drift value (i.e., the first burst arrivals at each slot). Then,  $Z$  is equal to  $\min\{d'_1, d'_2, \dots, d'_k\}$ , where  $d'_1, d'_2, \dots, d'_k$  form an increasing sequence of exponentially distributed drifts at slot  $N_1$ . Likewise,  $U$  is equal to  $\min\{d''_1, d''_2, \dots, d''_l\}$ , where  $d''_1, d''_2, \dots, d''_l$  form an increasing sequence of exponentially distributed drifts at slot  $N_2$ . Note that the smaller the drift, the earlier the arrival.

First of all, we find the cumulative distribution function of the  $Z$  random variable.

$$\begin{aligned} F_Z(z) &= P(Z < z) = P(\min\{d'_1, d'_2, \dots, d'_k\} \leq z) \\ &= 1 - P(\min\{d'_1, d'_2, \dots, d'_k\} > z) \\ &= 1 - P(\{d'_1 > z\} \cap \{d'_2 > z\} \cap \dots \cap \{d'_k > z\}) \\ &= 1 - P(d'_1 > z)^k = 1 - [1 - F_D(z)]^k, \end{aligned} \quad (3.10)$$

where  $F_D(x)$  is the cumulative distribution function of an exponential random variable  $D$  with parameter  $\alpha$ , and which corresponds to

$$F_D(x) = 1 - e^{-\alpha x} \forall x \in [0, \infty); \quad (3.11)$$

then,

$$\begin{aligned} F_Z(z) &= 1 - [1 - (1 - e^{-\alpha z})]^k \\ &= 1 - e^{-\alpha k z} \quad \forall z \in [0, \infty), k \in \mathbb{N}. \end{aligned} \quad (3.12)$$

Note that  $Z$  is an exponential distribution with parameter  $\alpha k$ .

Hereinafter, we assume that  $d'_k$  and  $d''_l$  are equally distributed with parameter  $\alpha$ . Therefore, if  $Z$  is exponential with parameter  $\alpha k$ ,  $U$  is also exponential with parameter  $\alpha l$ .

The term  $P(Z > U + \sigma)$  refers to the probability that the first burst to arrive at slot  $N_2$  is lost as a result of a drift-based collision. Then,

$$P(Z > U + \sigma) = P(Z > U + \sigma | Z > \sigma)P(Z > \sigma), \quad (3.13)$$

and if we apply the memoryless property of the exponential distribution, we have that

$$P(Z > U + \sigma | Z > \sigma) = P(Z > U); \quad (3.14)$$

hence,

$$P(Z > U + \sigma) = P(Z > U)P(Z > \sigma). \quad (3.15)$$

The first term is derived as follows:

$$P(Z > U) = \int_0^\infty \int_0^x (\alpha k e^{-\alpha k x} \alpha l e^{-\alpha l y}) dy dx = \frac{l}{l+k}, \quad (3.16)$$

and the second term corresponds to

$$P(Z > \sigma) = 1 - P(Z < \sigma) = 1 - F_Z(\sigma) = e^{-\alpha k \sigma}; \quad (3.17)$$

thus, we finally have

$$P(Z > U + \sigma) = e^{-\alpha k \sigma} \frac{l}{l+k}. \quad (3.18)$$

Note that the above expression can only be applied when we compute the drift-loss probability for the first burst arrival in the second slot. For all successive new arrivals  $n \in [2, \infty)$ , we must take into account their arrival times which consist of the arrival time of the first burst (i.e.,  $U + \sigma$ ) plus the distance between two consecutive drifts (i.e.,  $d''_n - d''_{n-1}$ ) times the number of arrivals after the first one (i.e.,  $n - 1$ ). Thus, we propose the following formulation:

$$\begin{aligned} P(Z > (n-1)(d''_n - d''_{n-1}) + U + \sigma) \\ &= P(Z > (n-1)(d''_n - d''_{n-1}) \\ &\quad + U + \sigma | Z > \sigma)P(Z > \sigma). \end{aligned} \quad (3.19)$$

Again, if we recursively apply the memoryless property of the exponential distribution, we obtain the following:

$$\begin{aligned}
P(Z > (n-1)(d''_n - d''_{n-1}) + U + \sigma | Z > \sigma) \\
&= P(Z > (n-1)(d''_n - d''_{n-1}) + U) \\
&= P(Z > (n-1)(d''_n - d''_{n-1}) \\
&\quad + U | Z > U)P(Z > U)
\end{aligned} \quad (3.20)$$

$$\begin{aligned}
P(Z > (n-1)(d''_n - d''_{n-1}) + U | Z > U) \\
&= P(Z > (n-1)(d''_n - d''_{n-1})) \\
&= P\left(\frac{Z}{(n-1)} > (d''_n - d''_{n-1})\right).
\end{aligned} \quad (3.21)$$

Hence, we eventually have that

$$\begin{aligned}
P(Z > (n-1)(d''_n - d''_{n-1}) + U + \sigma) \\
&= P\left(\frac{Z}{(n-1)} > (d''_n - d''_{n-1})\right) \\
&\quad \times P(Z > U)P(Z > \sigma).
\end{aligned} \quad (3.22)$$

Now the objective is to obtain a valid expression for  $P\left(\frac{Z}{(n-1)} > (d''_n - d''_{n-1})\right)$ . We define two new random variables, namely  $Q$  and  $T$ , that correspond to

$$Q = \frac{Z}{(n-1)} \quad (3.23)$$

$$T = d''_n - d''_{n-1} = X - Y. \quad (3.24)$$

It is easy to note that  $Q$  is an exponential random variable with parameter  $\alpha k(n-1)$ . However, in order to compute the probability  $P(Q > T)$ , we have to derive the density function of  $T$ , which corresponds to the subtraction of two exponential and independent random variables with the same parameter. Since the sequence is in increasing order, we must take into account the constraint  $X > Y$ .

$F_T(t)$  is defined as

$$F_T(t) = P(T \leq t) = P(X - Y \leq t). \quad (3.25)$$

Now, let us define the limits of the integration taking into account the constraint  $X > Y$ . The region is, therefore, defined by these two inequalities:

$$(a) x \leq y + t; \quad (b) x > y. \quad (3.26)$$

Hence,

$$\begin{aligned}
F_T(t) &= \int_0^\infty \int_y^{y+t} f_{XY}(x, y) dx dy \\
&= \int_0^\infty \int_y^{y+t} \alpha^2 e^{-(x+y)} dx dy \\
&= \frac{1}{2}(1 - e^{-\alpha t}).
\end{aligned} \quad (3.27)$$

The density function  $f_T(t)$  is obtained through the derivation of  $F_T(t)$ .

$$f_T(t) = \frac{dF_T(t)}{dt} = \frac{\alpha}{2} e^{-\alpha t} \quad \forall t \in [0, \infty). \quad (3.28)$$

Therefore, now it is possible to compute the probability  $P(Q > T)$  for the arrival number  $n$ , as follows:

$$\begin{aligned}
P(Q > T)_n &= \int_0^\infty \int_0^x (\alpha k(n-1) e^{-\alpha k(n-1)x} \frac{\alpha}{2} e^{-\alpha y}) dy dx \\
&= \frac{1}{2(k(n-1) + 1)} \quad \forall k \in \mathbb{N}.
\end{aligned} \quad (3.29)$$

Summarizing, to compute the drift-loss probability for the arrival number  $n \in [2, \infty)$  in the second slot, we apply the following formula:

$$\begin{aligned}
P(Z > (n-1)(d''_n - d''_{n-1}) + U + \sigma) \\
&= P(Z > \sigma)P(Z > U)P(Q > T)_n.
\end{aligned} \quad (3.30)$$

Finally, we can define the  $A_n$  values as follows:

$$A_n = \begin{cases} e^{-\alpha k \sigma} \frac{l}{l+k}, & n = 1 \\ e^{-\alpha k \sigma} \frac{l}{l+k} \frac{1}{2(k(n-1)+1)}, & \text{otherwise} \end{cases}. \quad (3.31)$$

A boundary condition can be inferred from this model. If  $\alpha = 0$  (i.e., there is no drift at edge nodes), then  $A_n = 0$  and  $B_r = 1$ . Therefore,  $Pb_{QS} = Pb_{SYNC}$  (i.e., the QS loss model becomes a synchronous loss model). Besides, it is easy to observe that, for a large number  $k$  of arrivals at slot  $N_1$ ,  $A_n$  will tend to a low probability value. In contrast, for a small number  $k$  of arrivals,  $A_n$  will tend to a high probability value. The same can be inferred from the number  $l$  of arrivals at slot  $N_2$ . Notice that the higher the number of arrivals, the smaller the drift of the first burst arrival.

### 3.2. Analytic models validation

In order to validate the analytic models, we compare their numerical results with those obtained through simulations.

#### 3.2.1. Single OBS node

We first consider a scenario composed by an isolated single OBS node with only one wavelength in its output port  $p$ . In all of our studies presented hereinafter, we consider, if not specifically given differently, the following simulation scenario: (1) simulations are conducted on the JAVOBS [17] network simulator; (2) the NSFNET (US network) [18] network topology, which has 14 nodes and 21 bidirectional links, is the selected network topology in all the experiments; (3) all network links have equal length and the resulting propagation delays correspond to a multiple of the slot size; (4) each channel has a capacity of 10 Gbps; (5) the load is normalized to the link capacity; (6) the traffic is uniformly distributed; (7) a single-path shortest-path routing algorithm is used; (8) both the processing and switching times are neglected; and (9) full wavelength conversion is available at all nodes.

In Figs. 3.2 and 3.3, we present the simulation and numerical results obtained using the QS model with guard-time values ranging from approximately  $\sigma = 0.033\%$  to  $33\%$  of the slot size. The drifts follow an exponential distribution with a parameter  $\alpha$  such that  $E(D) = \frac{1}{\alpha} \approx 0.77\%$  and  $E(D) \approx 5\%$  of the slot size, respectively, in each figure. Besides, the asynchronous and synchronous cases are used as benchmark references. As can be observed in both figures, all simulation results perfectly match those obtained through the analytic expressions presented before. It is interesting to notice that whilst a negligible value of the guard-time results in a performance close to that of the asynchronous operation, gradual increments of the guard-time display performances that tend towards

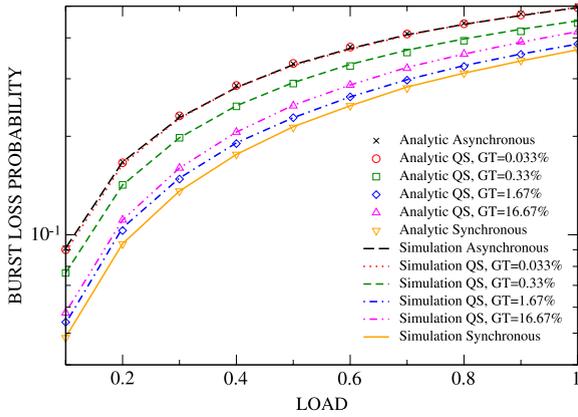


Fig. 3.2. Single-node validation with 0.77% drift.

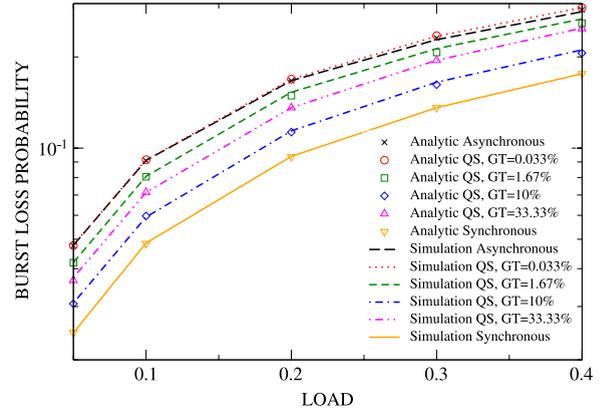


Fig. 3.3. Single-node validation with 5% drift.

that of the perfectly synchronized case. However, as stated earlier, we cannot increase as much as we want the value of the guard-time, since this will result in a performance degradation due to the overload cost that we should pay; indeed, in order to maintain a constant load when decreasing the size of the burst (i.e., increasing the guard-time), the number of bursts launched into the network is increased. As a direct consequence of these results, it is clear that finding an optimal value for the guard-time is of major importance in this problem. Although the derived analytic model only concerns a single-wavelength scenario, it allows one to gain a valuable insight into the evaluation of the QS operation in OBS networks.

### 3.2.2. OBS network

Taking advantage of the analytic model presented, we can gain more insight in the QS operation by analysing its behaviour in a network scale. In order to obtain the network-wide burst loss probability we apply the reduced-load fixed point approximation [14] model. It assumes that blocking events occur independently from link to link along any route  $r$ , where  $r = (l_1, \dots, l_{k-n}, \dots, l_k)$  is an ordered set of links which connect a source node to a destination node. It also assumes that the routes are predefined by means of using, for example, a shortest-path algorithm, and thus, the load assigned to each network link can be obtained. Then, considering a vector of stationary link blocking probabilities, the reduced offered load resulting from blocking can be approximated, i.e., for any route  $r$  that crosses link  $l_{k-n}$ , the load it offers to link  $l_{k-n}$  is reduced by blocking events occurred in its preceding links, that is, links  $l_1, \dots, l_{k-n-1}$ . In order to obtain a solution for the vector of link blocking probabilities, the authors of [14] make use of a successive substitution procedure for which convergence is guaranteed.

In our models, we compute the loss probability in each network link using either the Erlang-B loss formula (3.1) (asynchronous case), the synchronous loss formula (3.3) (synchronous case) or the QS loss formula (3.4) (quasi-synchronous case). Note that, in the QS case, the arrivals are still Poisson distributed since the drifts also follow an exponential distribution. Thus, the reduced-load fixed point approximation can also be applied in this case.

Due to the fact that our analytic model is only valid for a single-wavelength scenario, we consider links equipped with only one single channel. We present the results in Fig. 3.4 for a drift equal to a 0.77% of the slot size and guard-time values ranging from approximately  $\sigma = 0.033\%$  to 3.33% of the slot size, and in Fig. 3.5 for a drift equal to a 5% of the slot size and guard-time values ranging from  $\sigma = 0.033\%$  to 13.33% of the slot size. Note that, as a consequence of having only one channel in each link, we have to consider a very light loaded scenario in order to be able to plot significant values of the loss probability. We observe, in both figures, a similar behaviour to that obtained with the isolated OBS node. Consequently, with the use of an optimal value of the guard-time, we expect that the performance displayed by the QS operation will achieve results close to those obtained with perfect synchronization.

## 4. Time slot and guard-time dimensioning

The aim of this section is the dimensioning of the two main defining parameters of the QS-OBS network, namely the time slot size and the guard-time. Both parameters are of crucial importance in the design of the QS-OBS network, and hence, they impact seriously on the network performance. They not only must be carefully engineered in order to obtain good network performance results, but also must be dimensioned so that they are in accordance with the presence of the drift phenomenon in the considered OBS network scenarios. For this purpose, we first provide a reasonable model for the drifts, and second, we perform a simulation study to evaluate the impact that both the time slot and guard-time size have on the QS-OBS network performance.

### 4.1. Drift model

As mentioned in Section 2, we consider that the best way to model the drifts that may arise in a real OBS network is by making use of a Gaussian-distributed random variable (see Fig. 2.2). Specifically, we consider that the drifts follow a normal distribution denoted by  $N(0, \xi^2)$ , where  $\xi$  corresponds to the standard deviation of the random variable expressed in  $\mu\text{s}$ . Hence, we are assuming that 68% of the drift values fall within the region

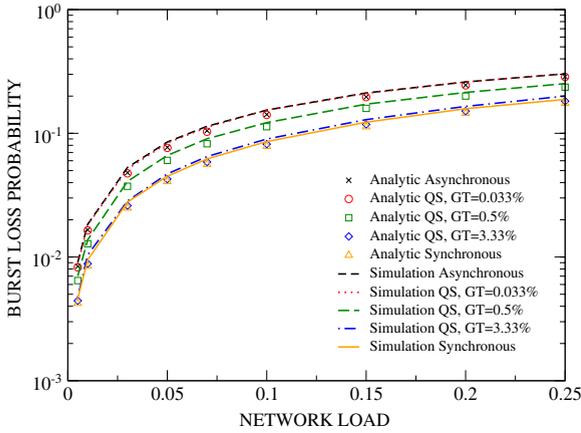


Fig. 3.4. NSFNET analytic results with a drift equal to 0.77%.

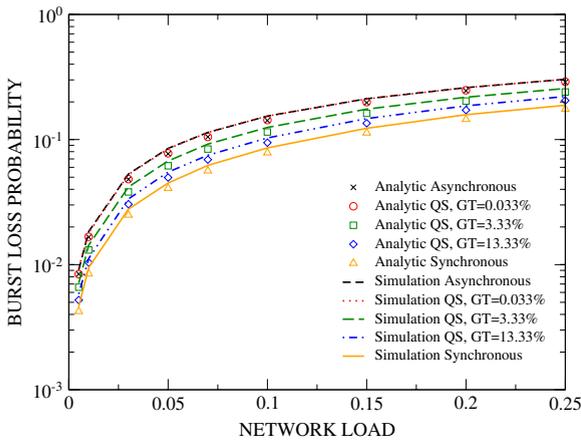


Fig. 3.5. NSFNET analytic results with a drift equal to 5%.

$[-\xi \mu\text{s}, \xi \mu\text{s}]$ , 95% of them  $\in [-2\xi \mu\text{s}, 2\xi \mu\text{s}]$  and 99.7%  $\in [-3\xi \mu\text{s}, 3\xi \mu\text{s}]$ .

We consider that, in real OBS network scenarios, drifts are the result of the contribution of several inaccuracies. On the one hand, there are the contributions of the inevitable structural inaccuracy of the devices constituting an OBS node which may introduce random time differences between two identical bursts (e.g., control pulse activating the lasers, inaccurate wavelength conversion, delayed processing time, etc.). On the other hand, the physical impairments of the fibres may change the propagation time characteristics of the different channels (recall that a burst can change the wavelength randomly along its path to solve a contention); nonetheless, considering that a signal cannot usually remain in the optical domain more than 1500 km before requiring a full regeneration, the effect of the wavelength walk-off can be considered to be limited to approximately  $1\mu\text{s}$  [20]. In summary, it is reasonable to model the drift phenomenon as a Gaussian distribution, where the majority of the bursts experience a drift within a given region while scattered events can fall in any zone.

## 4.2. Evaluation of the time slot size

In the QS-OBS network, both the time slot and guard-time size have a strong influence on the network performance results. Indeed, intuitively, one can remark that the larger the time slot, the smaller the influence of the (absolute) drift. Thus, whilst for the guard-time a trade-off between burst overlapping reduction and bandwidth utilization must be evaluated, for the time slot only a very large size would provide the optimality. However, the time slot size cannot be freely selected since other factors inherent to OBS networks must be taken into account. Indeed, the burst assembly algorithm performance as well as the basic guard-time already mentioned in Section 2 impact on the time slot dimensioning directly. Notice that the burst assembly time is limited and large time slots may lead to inefficient resource utilization due to the incapacity of the burst assembler to produce large bursts in time. On the other hand, since the basic guard-time is required between every two consecutive time slots, a small time slot size would also compromise resource utilization. An evaluation of the optimal time slot size for the SynOBS network which takes into consideration the issue of both the burst assembly and the basic guard-time can be found in [19]. Since our main objective is to evaluate the impact that both the drifts and skews have on the QS-OBS network, the evaluation of both the burst assembly process and the basic guard-time is left out of the scope of this paper. However, their consideration when determining the time slot size is certainly a subject for further research on QS-OBS networks.

In particular, here we assess the impact that the selection of the time slot size has on the overall QS-OBS network performance, in terms of the overall burst loss probability, considering several guard-time values. In order to do so, we conduct a series of simulations on the NSFNET network topology considering links are equipped with 16 wavelengths each, and the load is set to  $\rho = 0.5$ . We assume two different drift scenarios, namely  $\xi = 0.3 \mu\text{s}$  (see Fig. 4.1) and  $\xi = 0.9 \mu\text{s}$  (see Fig. 4.2). The time slot sizes considered range from  $5 \mu\text{s}$  to  $100 \mu\text{s}$ , and the guard-time values from 0 to 5% of the time slot size. The asynchronous and synchronous cases are used as performance references again. From the results obtained in both figures, it is easy to note that both factors together define a clear trade-off between drift correction (i.e., higher slot sizes benefit from larger guard-times) and resource utilization (i.e., larger guard-times increase the number of packets injected into the network). Whilst larger time slots require smaller guard-times to minimize the drift effect, smaller time slots struggle to reduce the drift impact unsuccessfully. Indeed, in such cases, no clear minimum is observed regardless of the guard-time value used, particularly in Fig. 4.2.

Hereinafter in this paper, we consider a time slot value of  $30 \mu\text{s}$  (a maximum burst size value of 0.3 Mb considering 10 Gbps links), since we assume that this provides a fair trade-off between the performance achieved and the requirements it imposes on both the burst assembler and the system (i.e., basic guard-time issue).

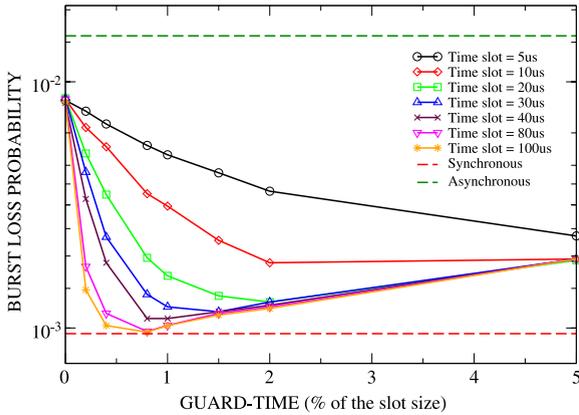


Fig. 4.1. Time slot and guard-time evaluation for drifts generated by  $\xi = 0.3 \mu\text{s}$ .

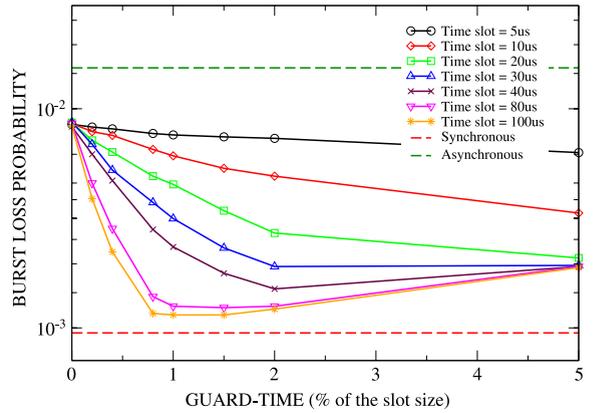


Fig. 4.2. Time slot and guard-time evaluation for drifts generated by  $\xi = 0.9 \mu\text{s}$ .

### 4.3. Optimal guard-time

In Section 3.2, we realized that a key factor for achieving the desired performance in the QS operation is the finding of an optimal value for the guard-time. This guard-time results in a clear trade-off between the reduction of the number of drift-based collisions (i.e., the increase of the guard-time reduces the probability of fulfilling condition (3.9)) and the increase of the number of overflow-based collisions (i.e., to maintain a constant load, a higher number of packets has to be injected into the network). To analyse the optimal value for the guard-time, we conduct several simulations in both the SIMPLE [21] (a network that consists of 6 nodes and 8 bidirectional links) and the NSFNET network topologies. In particular, we run a set of simulations in both networks considering a 30  $\mu\text{s}$  time slot and links equipped with 8, 16 and 32 wavelengths. The results for the NSFNET network topology are presented in Fig. 4.3 under two different load scenarios, namely  $\rho = 0.5$  and  $\rho = 0.6$ .

Fig. 4.3 clarifies the concept of the aforementioned guard-time trade-off. It is easy to observe that the value of the guard-time cannot be freely increased since it implies strong performance degradation. It is also relevant that, by considering a higher drift, the minimum point moves to higher guard-time values (right-hand side of the figure). From all the results obtained, we can conclude that under scenarios falling inside the typical OBS operating range (i.e., scenarios such that the network-wide burst loss probability  $\in [10^{-3}, 10^{-6}]$ ), there exists an optimal range for the selection of the guard-time. Despite the fact that we present the results for  $\xi$  values up to 3  $\mu\text{s}$  (recall that  $\xi$  represents the standard deviation of the Gaussian-distributed drift model), we assume that this value, as mentioned in Section 4.1, should not exceed 0.9  $\mu\text{s}$ . Taking this fact into account, we plot in the figure the optimal range for the guard-time. We observed in further analysis, not included here, that in such a range the optimal value for the guard-time does not depend either on the topology or on the number of wavelengths or the load. It is interesting to see that in such an optimal range there are only insignificant differences in the performance obtained. For these reasons, we select a guard-time value of 0.5  $\mu\text{s}$  (i.e.,

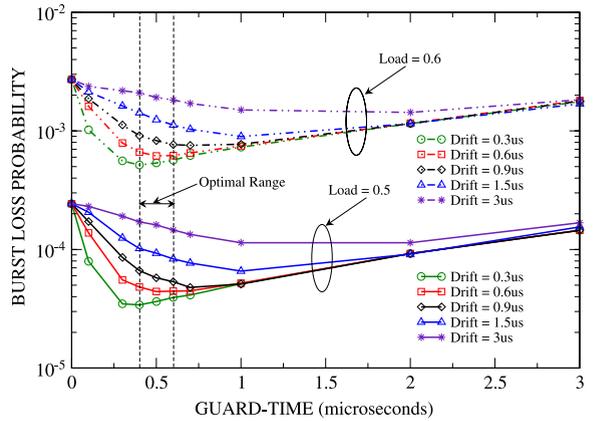


Fig. 4.3. Optimal guard-time study.

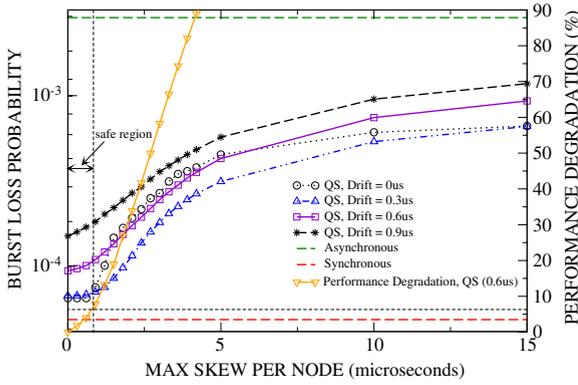
1.67% of the slot size for a 30  $\mu\text{s}$  time slot) which implies a reduction of the fixed burst size to 0.295 Mb. Hereinafter in this paper, we consider the 0.5  $\mu\text{s}$  guard-time as the optimal value for all  $\xi \in [0.3 \mu\text{s}, 1.5 \mu\text{s}]$ . Notice that the optimal guard-time range can also be spotted in Figs. 4.1 and 4.2.

## 5. Evaluation of the skew impact

Reaching a perfect synchronization at the edge nodes, so that all have the same clock information all the time, is practically impossible. Therefore, we assume that there exists some skew amongst their clocks. In the next subsections, we first evaluate the impact of the skew in the performance of the QS operation, and later, we propose a novel scheme to reduce the negative impact of such skew.

### 5.1. Performance degradation due to skew

In order to evaluate the skew impact, we consider that each edge node has clock information with a certain time deviation from their network counterparts. In particular, we assume that this time deviation can be considered constant in time within the  $\mu\text{s}$ – $\text{ms}$  scales (i.e., it changes at a larger time scale). In our model, we consider that



**Fig. 5.1.** Skew impact on the QS-OBS network performance under a load of 0.4.

(at a given time) each node has its own skew value; thus, we assign to each node a value generated according to a uniform random distribution that can be defined as  $U(0, \varphi)$ , where  $\varphi$  represents the maximum skew value expressed in  $\mu\text{s}$ .

To evaluate the degradation introduced by the skew, we present in Figs. 5.1 and 5.2 the results obtained under QS operation making use of the guard-time value obtained in Section 4.3, that is, a guard-time equal to  $0.5 \mu\text{s}$  (i.e., a burst size of  $0.295 \text{ Mb}$  in a time slot of  $30 \mu\text{s}$ ). We consider drifts with  $\xi$  values ranging from  $0 \mu\text{s}$  to  $0.9 \mu\text{s}$ . In the cases at hand, we assume links equipped with 16 wavelengths each and a network load equal to  $\rho = 0.4$  and  $\rho = 0.7$ , respectively, in each figure. In the x-axis we plot values of  $\varphi$  ranging from 0 to  $15 \mu\text{s}$  (i.e., 50% of the slot size). Whilst the left y-axis displays the performance degradation in terms of the network-wide burst loss probability, the right y-axis gives a percentage figure that is computed according to the following formula:

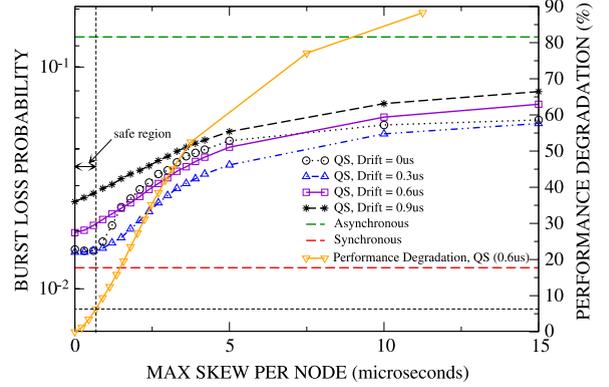
$$\text{Degradation}[\%] = \frac{BLP_{\varphi} - BLP_{\varphi=0}}{BLP_{\varphi=0}} * 100, \quad (5.1)$$

where  $BLP_{\varphi}$  corresponds to the network-wide burst loss probability for a particular  $\varphi$ . For the sake of readability, in the figures we only plot the performance degradation of the  $0.6 \mu\text{s}$  drift case.

As was to be expected, the performance of the QS operation is strongly worsened as a consequence of the presence of the skew. However, we noticed that there exists a region of interest for values of  $\varphi$  up to approximately  $0.9 \mu\text{s}$  of the slot size (referred to as the safe region in both figures) where no degradation is observed (i.e., the performance obtained remains nearly flat) and up to  $1.8 \mu\text{s}$  for a 25% degradation. Therefore, our objective is to devise a solution that is able to permanently guarantee that the skew present in the network falls within that safe region or, at least, within acceptable performance levels (e.g.,  $1.8 \mu\text{s}$  region).

## 5.2. Skew re-synchronization mechanism

In our study, we assume that the skew values may vary slowly and not faster than in a minutes scale. In



**Fig. 5.2.** Skew impact on the QS-OBS network performance under a load of 0.7.

fact, external factors such as changes in temperature or a voltage drift could cause transitions in the skew values. Therefore, we must also guarantee that the devised mechanism achieves its objective within a time scale that is not affected by possible skew transitions within such high-order scales. In consequence, aside from reaching the re-synchronization of the network, this mechanism also has to achieve it within an acceptable convergence time (e.g., within the milliseconds scale).

We propose a novel re-synchronization mechanism adapted to OBS networks. The basis of this technique is an averaging scheme proposed in [13] for distributing a well-aligned hardware clock throughout the physical extent of a synchronous processor. In fact, the averaging scheme, with minor variation, is used by many computer networks to maintain coherent notions of absolute time. The goal of this scheme is that all skew values converge to a common value. For our purposes, this final value is not significant, and it is only the final uniformity that matters. In this simple scheme, the skew of each local clock is driven towards the average skew of its topological neighbours (i.e., nodes that are directly connected by a physical link to a particular node). In this way, when all nodes have the same skew, all driving forces are zero and a stationary point is reached. It is clear that, in order to execute this mechanism inside a network, an exchange of information is needed. To convey such information inside an OBS network, we take advantage of burst control packets (BCPs). It is worth pointing out that contention between BCPs is generally neglected in OBS networks, and thus, the performance of the mechanism is not dependent on the load.

To explain our OBS-averaging scheme, let us first define  $L_x$  as the local clock reference at node  $x$  and  $RTT_{y,x}$  as the round-trip time between neighbouring nodes  $y$  and  $x$  (we assume each node computes the  $RTT$  to its neighbours by sending periodic pings through the control channel). Node  $x$  includes the time information  $L_x$  in each of its BCPs. A node  $y$  is able to compute its current clock skew relative to node  $x$  as

$$\delta(y, x) = L_y - \frac{RTT_{y,x}}{2} - L_x. \quad (5.2)$$

The exact skew can be determined if all values in (5.2) are exact. However, some uncertainties could be present, for

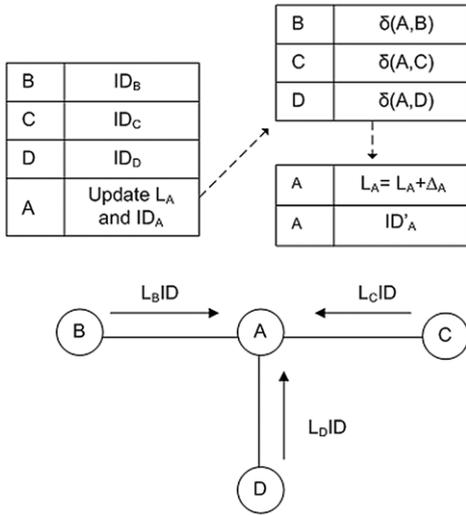


Fig. 5.3. Unique identifier usage in the re-synchronization mechanism.

instance, due to possible variations of the node processing time. This could introduce evaluation errors, and therefore, affect the synchronization process. Nevertheless, in Section 5.1, we show that a region of 0.9 μs or even of 1.8 μs is acceptable in the skew synchronization; thus, uncertainties causing errors that fall within such region do not impact negatively upon the entire process.

Once the network operation is started, each node  $x$  includes in every BCP it generates the following tuple:

$$\langle L_x, ID_x \rangle, \tag{5.3}$$

where  $ID_x$  is a unique identifier that is assigned to each different value of  $L_x$ ; that is, every time  $L_x$  is updated, a new and different unique  $ID_x$  is assigned to  $L_x$ . This stamp is used to prevent a node from updating its local clock before it has updated the clock information from all of its neighbours.

Fig. 5.3 clarifies how the identifier is used. At any time node  $A$  receives a BCP coming from any of its neighbours, it extracts the carried tuple and updates the information concerning this neighbour. To store this information, node  $A$  maintains two databases with information related to each of the nodes constituting its set of topological neighbours  $K = \{n_1, n_2, \dots, n_{|K|}\}$ , or, in this particular example,  $K = \{B, C, D\}$ . The first one (upper left-hand side) stores the unique identifiers that are being received. The value of a unique stamp is updated if and only if the newly received value differs from the stored one. Accordingly, once an update of any identifier is performed, its corresponding time information is also updated in the second database (upper right-hand side). For the sake of clarity, we have not depicted how node  $A$  sends its  $ID_A$  to nodes  $B, C$  and  $D$ .

Hence, node  $A$  cannot update either its  $L_A$  or its  $ID_A$  until it has updated information from all of its neighbours. As shown in Fig. 5.3, once  $A$  has a new identifier for all of its neighbours, it proceeds to update its local clock information according to the information stored in the time database (i.e., clock skews). Note that we illustrate an update on the identifier by replacing  $ID_A$  with  $ID'_A$ . In order to update  $L_A$ , node  $A$  performs the following two steps:

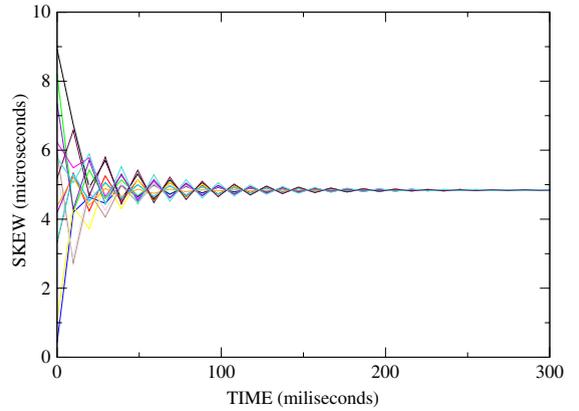


Fig. 5.4. Convergence for skew values generated by  $U(0, 10)$ .

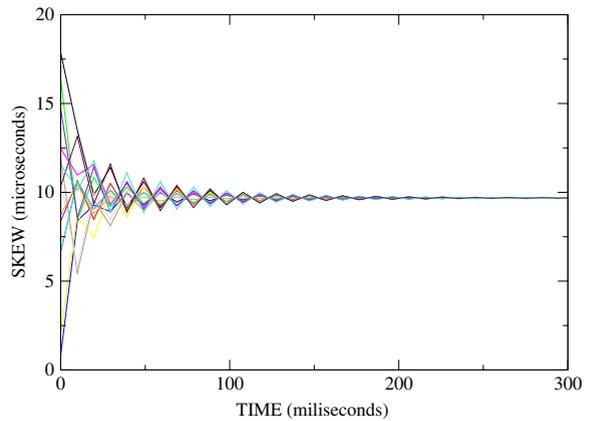


Fig. 5.5. Convergence for skew values generated by  $U(0, 20)$ .

- (1) Skew average computation:  $\Delta_A = \frac{1}{|K|} \sum_{i=1}^{|K|} \delta(A, n_i)$
- (2) Local clock update:  $L_A = L_A + \Delta_A$ .

In the event of a node not sending any control packet, which will prevent the mechanism to continue working properly (i.e., notice that nodes do not update their local clocks until they receive information from all neighbouring nodes), our mechanism makes use of an additional time-out reference to avoid such a deadlock state and trigger an automatical re-start of the re-synchronization mechanism without taking that particular node into consideration.

In order to analyse the convergence properties of the mechanism, in this particular experiment, we consider the real length of the NSFNET network links. Hence, we have both correct information on propagation delays and a valid estimation of the convergence time. To point out that it is the uniformity that matters and not the final stationary point, we present in Figs. 5.4 and 5.5 the results obtained using two different uniform distributions for the skew, namely  $U(0, 10)$  and  $U(0, 20)$ . In the figures, each line represents the value of the skew in one of the nodes. In these two different cases, the stationary points are found at approximately 4.8 μs and 9.7 μs, respectively. In spite of the difference found between them, in both cases a nearly perfect re-synchronization is reached at about 200 ms after the mechanism is triggered. Note that this implies that the

time-out, utilized to prevent that a node not sending any control packet alters the re-synchronization procedure, must be enough larger than this value, and can be set up to 10 s, for example.

As a conclusion, these results show that the skew effect can be completely erased from the network when our re-synchronization mechanism is effectively applied.

## 6. Deflection routing support in a QS-OBS network

In this final section, we present the results of the performance of the QS operation for OBS networks making use of the optimal guard-time value of  $0.5 \mu\text{s}$  (i.e., a burst size of 0.295 Mb) found in Section 4.3 and also considering that the skew impact in the network can be bounded to negligible values thanks to the re-synchronization mechanism presented in Section 5. However, this time, and in order to improve the performance of the synchronous strategies with respect to the asynchronous one, we consider routing algorithms that make use of effective deflection routing policies. To be precise, we evaluate the performance of the following deflection techniques: (1) load-based reflection routing algorithm (LBRR) [22] (without resource pre-allocation); (2) reflection-based deflection routing (RDR) [23]; (3) multi-topology routing (MTR) [24]; and (4) conventional deflection routing (DR) [25]. For comparison purposes, we use a time-to-live (TTL) field in the control packets that limits the number of hops a burst can undertake.

First, we analyse the behaviour of the QS operation regarding the number of wavelengths in each link. For this purpose, we consider the LBRR algorithm under both 32 wavelengths (see Fig. 6.1) and 64 wavelengths (see Fig. 6.2). A reflection routing algorithm allows sending a contending burst towards a neighbouring node (reflection neighbour) on the condition that this reflection neighbour, after receiving the burst, will intend to return the burst back or, in other words, reflect it. In this study, we set the maximum number of reflections a burst can undergo to three. We provide the results for the QS operation under drifts corresponding to three different  $\xi$  values, namely  $0.3 \mu\text{s}$ ,  $0.9 \mu\text{s}$  and  $1.5 \mu\text{s}$ . In the plot, the asynchronous and synchronous cases are included as benchmark indicators. As can be seen, even for high values of  $\xi$ , such as  $1.5 \mu\text{s}$ , the improvement of the QS operation with respect to the asynchronous operation is noticeable in both cases. In fact, as the number of wavelengths increases (see Fig. 6.2), the impact of the drift is substantially lessened (i.e., the more channels that are available, the more chances a burst has to find a compatible drift). Therefore, the performance improvement achieved by the QS operation becomes, inside the typical OBS operating range, approximately an order of magnitude, which is a very significant figure. It is also worth mentioning that, with the increase of the number of wavelengths, the performance gain of the synchronous case with respect to the asynchronous one is also significantly improved.

Second, we evaluate the performance of the QS operation with respect to both the asynchronous and synchronous cases considering all the aforementioned deflection routing algorithms. In order to fairly compare the

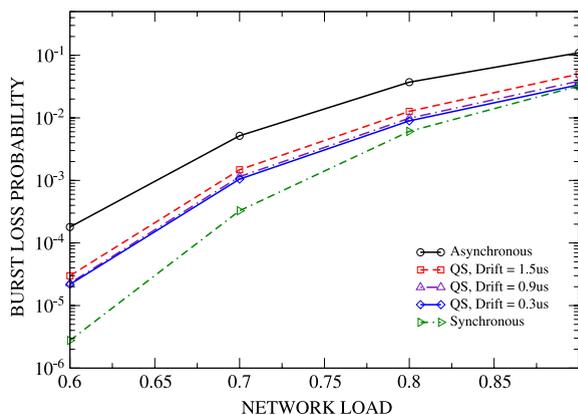


Fig. 6.1. QS performance using the LBRR algorithm with 32 wavelengths.

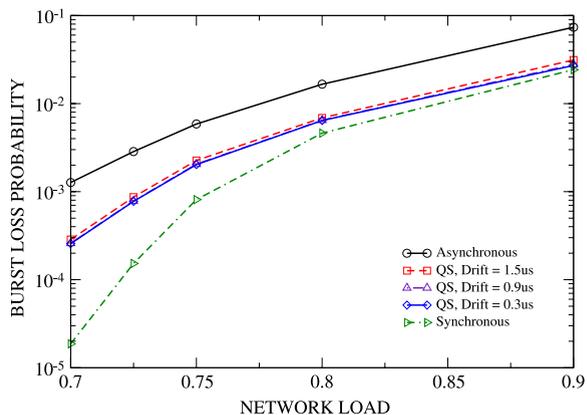


Fig. 6.2. QS performance using the LBRR algorithm with 64 wavelengths.

different strategies, the TTL is set to six hops more than the shortest-path route. To be precise, since the MTR algorithm applied to the NSFNET network only allows undergoing a maximum of three deflections per burst (i.e., MTR divides the NSFNET topology into four different layers), we also set the maximum number of reflections a burst can undergo to three in the LBRR algorithm (i.e., the aforementioned six hops). In this case, the number of wavelengths per link is set to 32 and the load injected into the network to  $\rho = 0.65$ . This load value has been selected in order to obtain burst loss probability values which fall inside the typical OBS operating range. The results obtained are presented in Table 6.1. We also include the results obtained under shortest-path routing (SPR). In all the five different cases, the performance improvement obtained with the QS operation (considering a drift with  $\xi = 0.3 \mu\text{s}$ ) is higher than 77%, and in some cases, approximately an order of magnitude (e.g., LBRR and RDR algorithms). Notice that the benefit is slightly less under the SPR algorithm. It is, hence, possible to achieve significant improvements when the QS transmission mode is used together with effective deflection routing techniques.

## 7. Conclusions

In this paper, we have proposed a novel QS operation mode for OBS networks aiming to achieve performance

**Table 6.1**

Performance comparison between the asynchronous, synchronous and QS transmission modes.

Routing algorithm	Asynchronous	QS(0.3 $\mu$ s)	Synchronous
SPR	$6.33 * 10^{-3}$	$1.60 * 10^{-3}$	$1.30 * 10^{-3}$
DR	$7.74 * 10^{-4}$	$1.76 * 10^{-4}$	$2.19 * 10^{-5}$
MTR	$4.08 * 10^{-4}$	$7.94 * 10^{-5}$	$2.82 * 10^{-5}$
LBRR	$1.33 * 10^{-3}$	$1.94 * 10^{-4}$	$3.18 * 10^{-5}$
RDR	$9.37 * 10^{-4}$	$1.45 * 10^{-4}$	$5.57 * 10^{-6}$

results close to those obtained with perfect synchronization. We contrasted the architectural benefits and requirements of the QS scheme with those of the synchronous one. We also showed that with the use of effective deflection routing techniques the performance improvement of the synchronous mode with respect to the asynchronous one is brought to a very interesting range. Therefore, the idea of devising a novel architecture, without the technical requirements imposed by the perfectly synchronized case, such as our QS operation mode, gains momentum.

For this purpose, we considered two different sources of time deviation in our QS-OBS scenario, namely skew and drift. We developed an analytic model to test the performance of the QS operation in both an isolated node and the NSFNET network topology. The model presented is exact for the case when the drifts are exponentially distributed. Through numerical examples, and with the use of a more realistic distribution of the drifts, an optimal range for the selection of the optimal guard-time value was found. We also observed that within the region of interest such a value is not dependent on any network parameters. In order to correct the skew effect, we proposed an averaging scheme that effectively limits the impact of this time phenomenon. Thanks to this scheme, the performance of the QS operation remains unaffected. Eventually, we performed a set of simulations to study the behaviour of the QS approach when deflection routing policies are applied.

The main benefit of the proposed QS transmission mode is the significant improvement attained with respect to the asynchronous mode in terms of the overall burst loss probability. In the specific examples considered in this paper, improvements as high as an order of magnitude have been observed when the QS transmission mode is used together with effective deflection routing algorithms.

## Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 247674 (STRONGEST-project). The authors would also like to thank the support received from the Spanish Ministry of Science and Innovation under the "COPERNic" project (Ref. TEC2009-13252), the Catalan Government under the contract SGR-1140, and the Polish Ministry of Science and Higher Education under the contract 643/N-COST/2010/0.

## References

- [1] C. Qiao, M. Yoo, Optical burst switching (OBS)—a new paradigm for an optical Internet, *J. High Speed Netw.* 8 (1) (1999) 69–84.
- [2] M. Klinkowski, D. Careglio, J. Solé-Pareta, M. Marciniak, Performance overview of the offset time emulated OBS network architecture, *IEEE/OSA J. Lightwave Technol.* 27 (14) (2009) 2751–2764.
- [3] F. Kuo, The ALOHA system, *Computer Networks*, Prentice-Hall, 1973, p. 501–518.
- [4] C. Guillemot, et al., Transparent optical packet switching: the European ACTS KEOPS project approach, *IEEE/OSA J. Lightwave Technol.* 16 (12) (1998) 2117–2134.
- [5] O. Ozturk, E. Karasan, N. Akar, Performance evaluation of slotted optical burst switching systems with quality of service differentiation, *IEEE/OSA J. Lightwave Technol.* 27 (14) (2009) 2621–2633.
- [6] Z. Zhang, L. Liu, Y. Yang, Slotted optical burst switching (SOBS) networks, *Comput. Commun.* 30 (8) (2007) 3471–3479.
- [7] J. Ramamirtham, J. Turner, Time sliced optical burst switching, in: *Proceedings of IEEE INFOCOM '03*, San Francisco, CA, USA, Apr., 2003, pp. 2030–2038.
- [8] F. Farahmad, V.M. Vokkarane, J.P. Jue, Practical priority contention resolution for slotted optical burst switching networks, in: *Proceedings of IEEE/SPIE 1st Int. Workshop Optical Burst Switching*, WOBS 2003, Dallas, USA, Sep., 2003.
- [9] A. Rugsachart, R.A. Thompson, An analysis of time-synchronized optical burst switching, in: *Proceedings of Workshop on High Performance Switching and Routing*, HPSR 2006, Poznan, Poland, June, 2006.
- [10] O. Pedrola, S. Rumley, D. Careglio, M. Klinkowski, P. Pedroso, C. Gaumier, J. Solé-Pareta, A performance survey on deflection routing techniques for OBS networks, in: *Proceedings of International Conference on Transparent Optical Networks*, ICTON 2009, São Miguel, Azores, Portugal, June, 2009.
- [11] A. Rugsachart, Time-synchronized optical burst switching, Ph.D. dissertation, University of Pittsburgh, 2007.
- [12] A. Pattavina, Architectures and performance of optical packet switching nodes for IP networks, *IEEE/OSA J. Lightwave Technol.* 23 (3) (2005) 1023–1032.
- [13] Gill A. Pratt, John Nguyen, Distributed Synchronous Clocking, *IEEE Trans. Parallel Distrib. Syst.* 6 (3) (1995) 314–328.
- [14] Z. Rosberg, Hai Le Vu, M. Zukerman, J. White, Performance analyses of optical burst-switching networks, *IEEE J. Sel. Areas Commun.* 21 (7) (2003) 1187–1197.
- [15] J. Teng, G. Rouskas, A detailed analysis and performance comparison of wavelength reservation schemes for optical burst switched networks, *Photonic Netw. Commun.* 9 (3) (2005) 311–335.
- [16] S.L. Danielsen, B. Mikkelsen, C. Joergensen, T. Durhuus, K.E. Stubkjaer, WDM packet switch architectures and analysis of the influence of tuneable wavelength converters on the performance, *IEEE/OSA J. Lightwave Technol.* 15 (2) (1997) 219–226.
- [17] O. Pedrola, S. Rumley, M. Klinkowski, D. Careglio, C. Gaumier, J. Solé-Pareta, JAVOBS: a flexible simulator for OBS network architectures, *J. Netw.* 5 (2) (2010) 256–264.
- [18] K.C. Claffy, G.C. Polyzos, H.W. Braun, Traffic characteristics of the T1 NSFNET backbone, in: *Proceedings of the IEEE INFOCOM '93*, San Francisco, CA, USA, Mar., 1993, pp. 885–892.
- [19] A. Rugsachart, R. Thompson, Optimal timeslot size for synchronous optical burst switching, in: *Proceedings of BROADNETS'07*, Raleigh, North Carolina, Sep., 2007.
- [20] G.P. Agrawal, *Fiber-Optic Communication Systems*, Wiley, Hoboken, NJ, 2002.
- [21] M. Klinkowski, M. Pioro, D. Careglio, M. Marciniak, J. Solé-Pareta, Routing optimization in optical burst switching networks, in: *Proceedings of 11th IFIP Working Conference on Optical Network Design and Modelling*, ONDM2007, Athens, Greece, May, 2007.
- [22] J. Perelló, S. Spadaro, J. Comellas, G. Junyent, A load-based reflection routing protocol with resource pre-allocation for contention resolution in OBS networks, *Eur. Trans. Telecommun.* 20 (2009) 1–7.
- [23] M. Morita, H. Tode, K. Murakami, Reflection-based deflection routing in OPS networks, *IEICE Trans. Commun.* E91-B (2) (2008) 409–417.
- [24] S. Gjessing, A Novel Method for Re-Routing in OBS Networks, *IEEE ISCT 2007*, New York, USA, Oct., 2007, pp. 22–27.
- [25] C. Hsu, T. Liu, N. Huang, Performance Analysis of Deflection Routing in Optical Burst-Switched Networks, in: *Proceedings of IEEE INFOCOM 2002*, New York, USA, June, 2002.