12

Optical Burst-Switched Networks Exploiting Traffic Engineering in the Wavelength Domain

João Pedro[1,2] and João Pires[2]
[1]*Nokia Siemens Networks Portugal S.A.*
[2]*Instituto de Telecomunicações, Instituto Superior Técnico*
Portugal

1. Introduction

In order to simplify the design and operation of telecommunications networks, it is common to describe them in a layered structure constituted by a service network layer on top of a transport network layer. The service network layer provides services to its users, whereas the transport network layer comprises the infrastructure required to support the service networks. Hence, transport networks should be designed to be as independent as possible from the services supported, while providing functions such as transmission, multiplexing, routing, capacity provisioning, protection, and management. Typically, a transport network includes multiple network domains, such as access, aggregation, metropolitan and core, ordered by decreasing proximity to the end-users, increasing geographical coverage, and growing level of traffic aggregation.

Metropolitan and, particularly, core transport networks have to transfer large amounts of information over long distances, consequently demanding high capacity and reliable transport technologies. Multiplexing of lower data rate signals into higher data rate signals appropriate for transmission is one of the important tasks of transport networks. Time Division Multiplexing (TDM) is widely utilized in these networks and is the fundamental building block of the Synchronous Digital Hierarchy (SDH) / Synchronous Optical Network (SONET) technologies. The success of SDH/SONET is mostly due to the utilization of a common time reference, improving the cost-effectiveness of adding/extracting lower order signals from the multiplexed signal, the augmented reliability and interoperability, and the standardization of optical interfaces. SDH/SONET networks also generalized the use of optical fibre as the transmission medium of metropolitan and core networks. Essentially, when compared to twisted copper pair and coaxial cable, optical fibre benefits from a much larger bandwidth and lower attenuation, as well as being almost immune to electromagnetic interferences. These features are key to transmit information at larger bit rates over longer distances without signal regeneration.

Despite the proved merits of SDH/SONET systems, augmenting the capacity of transport networks via increasing their data rates is only cost-effective up to a certain extent, whereas

adding parallel systems by deploying additional fibres is very expensive. The prevailing solution to expand network capacity was to rely on Wavelength Division Multiplexing (WDM) to transmit parallel SDH/SONET signals in different wavelength channels of the same fibre. Nevertheless, since WDM was only used in point-to-point links, switching was performed in the electrical domain, demanding Optical-Electrical (OE) conversions at the input and Electrical-Optical (EO) conversions at the output of each intermediate node, as well as electrical switches. Both the OE and EO converters and the electrical switches are expensive and they represent a large share of the network cost.

Nowadays, transport networks already benefit from optical switching, thereby alleviating the use of expensive and power consuming OE and EO converters and electrical switching equipment operating at increasingly higher bit rates (Korotky, 2004). The main ingredients to support optical switching are the utilization of reconfigurable nodes, like Reconfigurable Optical Add/Drop Multiplexers (ROADMs) and Optical Cross-Connects (OXCs), along with a control plane, such as the Generalized Multi-Protocol Label Switching (GMPLS), (IETF, 2002), and the Automatically Switched Optical Network (ASON), (ITU-T, 2006). The control plane has the task of establishing/terminating optical paths (lightpaths) in response to connection requests from the service network. As a result, the current type of dynamic optical networks is designated as Optical Circuit Switching (OCS).

In an OCS network, bandwidth is allocated between two nodes by setting up one or more lightpaths (Zang et al., 2001). Consequently, the capacity made available for transmitting data from one node to the other can only be incremented or decremented in multiples of the wavelength capacity, which is typically large (e.g., 10 or 40 Gb/s). Moreover, the process of establishing a lightpath can be relatively slow, since it usually relies on two-way resource reservation mechanisms. Therefore, although the deployment of OCS networks only makes use of already mature optical technologies, these networks are inefficient in supporting bursty data traffic due to their coarse wavelength granularity and limited ability to adapt the allocated wavelength resources to the traffic demands in short time-scales, which can also increase the bandwidth waste due to capacity overprovisioning.

Diverse solutions have been proposed to overcome the limitations of OCS networks and improve the bandwidth utilization efficiency of future optical transport networks. The less disruptive approach consists of an optimized combination of optical and electrical switching at the network nodes. In this case, entire wavelength channels are switched optically at a node if the carried traffic flows, originated at upstream nodes, approximately occupy the entire wavelength capacity. Alternatively, traffic flows with small bandwidth requirements can be groomed (electrically) into one wavelength channel with enough spare capacity (Zhu et al., 2005). This hybrid switching solution demands costly OE/EO converters and electrical switches, albeit in/of smaller numbers/sizes than those needed in opaque implementations relying only on electrical switching. However, OCS networks with electrical grooming only become attractive when it is possible to estimate in advance the fractions of traffic to be groomed and switched transparently at each node, enabling to accurately dimension both the optical and electrical switches needed to accomplish an optimized trade-off between maximizing the bandwidth utilization and minimizing the electrical switching and OE/EO

conversion equipment. Otherwise, when the traffic pattern cannot be accurately predicted, this trade-off can become difficult to attain and both optical and electrical switches may have to be overdimensioned, hampering the cost-effectiveness of this hybrid approach.

The most advanced all-optical switching paradigm for supporting data traffic over optical transport networks is Optical Packet Switching (OPS). Ideally, OPS would replicate current store-and-forward packet-switched networks in the optical domain, thereby providing statistical multiplexing with packet granularity, rendering the highest bandwidth utilization when supporting bursty data traffic. In the full implementation of OPS, both data payload and their headers are processed and routed in the optical domain. However, the logical operations needed to perform address lookup are difficult to realize in the optical domain with state-of-the-art optics. Similarly to MPLS, Optical Label Switching (OLS) simplifies these logical operations through using label switching as the packet forwarding technique (Chang et al., 2006). In their simplest form, OPS networks can even rely on processing the header/label of each packet in the electrical domain, while the payload is kept in the optical domain. Nevertheless, despite the complexity differences of the implementations proposed in the literature, the deployment of any variant of OPS networks is always hampered by current limitations in optical processing technology, namely the absence of an optical equivalent of electronic Random-Access Memory (RAM), which is vital both for buffering packets while their header/label is being processed and for contention resolution (Tucker, 2006; Zhou & Yang, 2003), and the difficulty to fabricate large-sized fast optical switches, essential for per packet switching at high bit rates (Papadimitriou et al., 2003).

The above discussion highlighted that OCS networks are relatively simple to implement but inefficient for transporting bursty data traffic, whereas OPS networks are efficient for transporting this type of traffic but very difficult to implement with state-of-the-art optical technology. Next-generation optical networks would benefit from an optical switching approach whose bandwidth utilization and optical technology requirements lie between those of OCS and OPS. In order to address this challenge, an intermediate optical switching paradigm has been proposed and studied in the literature – Optical Burst Switching (OBS).

The basic premise of OBS is the development of a novel architecture for next-generation optical WDM networks characterized by enhanced flexibility to accommodate rapidly fluctuating traffic patterns without requiring major technological breakthroughs. A number of features have been identified as key to attain this objective (Chen et al., 2004). In order to overview some of them, consider an optical network comprising edge nodes, interfacing with the service network, and core nodes, as illustrated in Fig. 1. OBS networks grant intermediate switching granularity (between that of circuits and packets) via: assembling multiple packets into larger data containers, designated as data bursts, at the ingress edge nodes, enforcing per burst switching at the core nodes, and disassembling the packets at the egress edge nodes. Noteworthy, data bursts are only assembled and transmitted into the OBS network when data from the service network arrives at an edge node. This circumvents the stranded capacity problem of OCS networks, where the bandwidth requirements from the service network evolve throughout the lifetime of a lightpath and during periods of time can be considerably smaller than the provisioned capacity. Furthermore, the granularity at which the OBS network operates can be controlled through varying the number of packets contained in the data bursts, enabling to regulate the control and switching overhead.

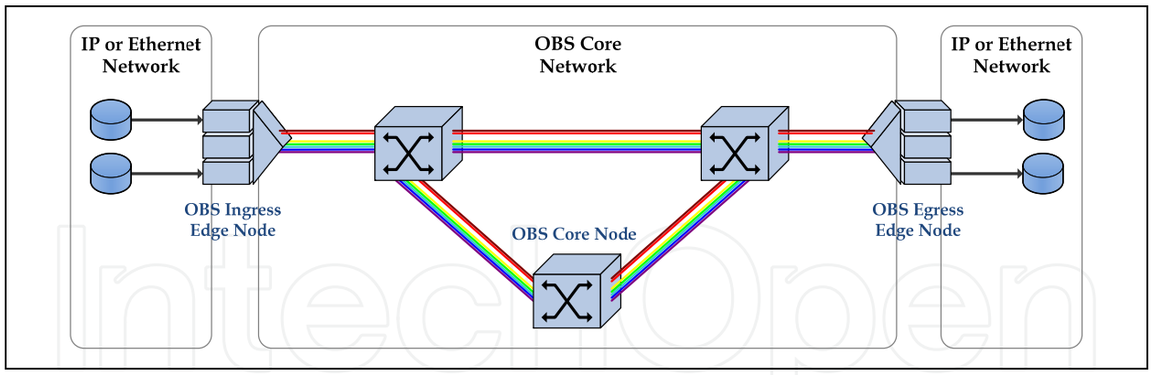


Fig. 1. Generic OBS network architecture.

In OBS networks, similarly to OCS networks, control information is transmitted in a separate wavelength channel and processed in the electronic domain at each node, avoiding complex optical processing functions inherent to OPS networks. More precisely, a data burst and its header packet are decoupled in both the wavelength and time domains, since they are transmitted in different wavelengths and the header precedes the data burst by an offset time. Channel separation of headers and data bursts, a distinctive feature of out-of-band signalling schemes, is suitable to efficiently support electronic processing of headers while preserving data in the optical domain, because OE/EO converters at the core nodes are only needed for the control channel. The offset time has a central role in OBS networks, since it is dimensioned to guarantee the burst header is processed and resources are reserved for the upcoming data burst before the latter arrives to the node. Accordingly, a data burst can cut through the core nodes all-optically, avoiding being buffered at their input during the time needed for header processing. Moreover, since the transmission of data bursts can be asynchronous, complex synchronization schemes are not mandatory. Combined, these features ensure OBS networks can be implemented without making use of optical buffering.

The prospects of deploying OBS in future transport networks can be improved provided that the bandwidth utilization achievable with OBS networks can be enhanced without significantly increasing their complexity or, alternatively, by easing their implementation without penalizing network performance. Noteworthy, OBS networks are technologically more demanding than OCS networks in several aspects. Firstly, although OBS protocols avoid optical buffering, OBS networks still demand some technology undergoing research, namely all-optical wavelength converters (Poustie, 2005) and fast optical switches scalable to large port counts (Papadimitriou et al., 2003). Secondly, the finer granularity of OBS is accomplished at the expense of a control plane more complex than the one needed for OCS networks (Barakat & Darcie, 2007). Nevertheless, the expected benefits of adopting a more bandwidth efficient optical switching paradigm fuelled significant research efforts in OBS, which even resulted in small network demonstrators (Sahara et al., 2003; Sun et al., 2005).

The performance of OBS networks is mainly limited by data loss due to contention for the same transmission resources between multiple data bursts (Chen et al., 2004). The lack of optical RAM limits the effectiveness of contention resolution in OBS networks. Wavelength conversion is usually assumed to be available to resolve contention for the same wavelength channel. In view of the complexity and immaturity of all-optical wavelength converters,

decreasing the number of converters utilized or using simpler ones without degrading performance would enhance the cost-effectiveness of OBS networks. Nevertheless, even if wavelength conversion is available, contention occurs when the number of bursts directed to the same link exceeds the number of wavelength channels. Moreover, the asynchronous transmission of data bursts creates voids between consecutive data bursts scheduled in the same wavelength channel, further contributing to contention. Consequently, minimizing these voids and smoothing burst traffic without resorting to complex contention resolution strategies would also improve the cost-effectiveness of OBS networks.

In alternative or as a complement to contention resolution strategies, such as wavelength conversion, the probability of resource contention in an OBS network can be proactively reduced using contention minimization strategies. Essentially, these strategies optimize the resources allocated for transmitting data bursts in such way that the probability of multiple data bursts contending for the same network resources is reduced. Contention minimization strategies for OBS networks mainly consist of optimizing the wavelength assignment at the ingress edge nodes to decrease contention for the same wavelength channel (Wang et al., 2003), mitigating the performance degradation from unused voids between consecutive data bursts scheduled in the same wavelength channel (Xiong et al., 2000), and selectively smoothing the burst traffic entering the network (Li & Qiao, 2004). Albeit the utilization of these strategies can entail additional network requirements, namely augmenting the (electronic) processing capacity in order to support more advanced algorithms, it is expected that the benefits in terms of performance or complexity reduction will justify their support.

This chapter details two contention minimization strategies, which when combined provide traffic engineering in the wavelength domain for OBS networks. The utilization of this approach is shown to significantly improve network performance and reduce the number of wavelength converters deployed at the network nodes, enhancing their cost-effectiveness.

The remaining of the chapter is organized as follows. The second section introduces the problem of wavelength assignment in OBS networks whose nodes have no wavelength converters or have a limited number of wavelength converters. A heuristic algorithm for optimizing the wavelength assignment in these networks is described and exemplified. The third section addresses the utilization of electronic buffering at the ingress edge nodes of OBS networks, highlighting its potential for smoothing the input burst traffic and describing how it can be combined with the heuristic algorithm detailed in the previous section to attain traffic engineering in the wavelength domain. The performance improvements and node complexity reduction made possible by employing these strategies in an OBS network are evaluated via network simulation in the fourth section. Finally, the fifth and last section presents the final remarks of the work presented in this chapter.

2. Priority-based wavelength assignment

OBS networks utilize one-way resource reservation, such as the Just Enough Time (JET) protocol (Qiao & Yoo, 1999). The principles of burst transmission are as follows. Upon assembling a data burst from multiple packets, the ingress node generates a Burst Header Packet (BHP) containing the offset time between itself and the data burst, as well as the length of the data burst. This node also sets a local timer to the value of the offset time.

The BHP is transmitted via a control wavelength channel and processed at the control unit of each node along the routing path of the burst. The control unit uses the information in the BHP to determine the resources (e.g., wavelength channel in the designated output fibre link) to be allocated to the data burst during the time interval it is expected to be traversing the core node. This corresponds to a delayed resource reservation, since the resources are not immediately set up, but instead are only set up just before the arrival time of the data burst. Furthermore, the resources are allocated to the burst during the time strictly necessary for it to successfully pass through the node. This minimizes the bandwidth waste because these resources can be allocated to other bursts in non-overlapping time intervals. Before forwarding the BHP to the next node, the control unit updates the offset time, reducing it by the amount of time spent by the BHP at the node. Meanwhile, the data burst buffered at the ingress node is transmitted after the timer set to the offset time expires. In case of successful resource reservation by its BHP at all the nodes of the routing path, the burst cuts through the core nodes in the optical domain until it arrives to the egress node. Otherwise, when resource reservation is unsuccessful at a node, both BHP and data burst are dropped at that node and the failed burst transmission is signalled to the ingress node.

As a result of using one-way resource reservation, there is a large probability that data bursts arrive at a core node on the same wavelength channel from different input fibre links and being directed to the same output fibre link of that node. This leads to contention for the same wavelength channel at the output fibre link. These contention events can be efficiently resolved using wavelength converters and/or minimized in advance through an optimized assignment of wavelengths at the ingress nodes. In view of the immaturity of all-optical wavelength converters, strategies for minimizing the probability of wavelength contention become of paramount importance in order to design cost-effective OBS core nodes.

2.1 Problem statement

Consider an OBS network modelled as a directed graph $G = (V, E)$, where $V = \{v_1, v_2, ..., v_N\}$ is the set of nodes, $E = \{e_1, e_2, ..., e_L\}$ is the set of unidirectional fibre links and the network has a total of N nodes and L fibre links. Each fibre link supports a set of W data wavelength channels, $\{\lambda_1, \lambda_2, \ldots, \lambda_{W-1}, \lambda_W\}$. Let $\Pi = \{\pi_1, \pi_2, \ldots, \pi_{|\Pi|-1}, \pi_{|\Pi|}\}$ denote the set of routing paths used to transmit data bursts in the network, E_i denote the set of fibre links traversed by path $\pi_i \in \Pi$, and γ_i denote the average traffic load offered to path π_i. It is assumed that the average offered traffic load values are obtained empirically or based on long-term predictions of the network load. Ideally, this input information would be used to formulate a combinatorial optimization problem for determining a wavelength search ordering, that is, an ordered list of all W wavelength channels, for each routing path such that a relevant performance metric, like the average burst blocking probability, is minimized. However, blocking probability performance metrics can only be computed via network simulation or, in particular cases, estimated by solving a set of non-linear equations (Pedro et al., 2006a). As a result, the objective function cannot be expressed in terms of the problem variables in an analytical closed-form manner (Teng & Rouskas, 2005). Moreover, even if this was possible, the size of the solutions search space would grow steeply with the number of wavelength channels W and the number of routing paths

$|\Pi|$, since there are $(W!)^{|\Pi|}$ combinations of wavelength channel orderings. Consequently, for OBS networks of realistic size, this would prevent computing the optimum wavelength search orderings in a reasonable amount of time.

In view of the aforementioned limitations in both problem formulation and resolution, the wavelength search orderings must be computed without knowing the resulting average burst blocking probability and by relying on heuristic algorithms. Notably, when the core nodes have limited or no wavelength conversion capabilities, burst blocking probability is closely related with the expected amount of unresolved wavelength contention. Consider two routing paths, π_1 and π_2, that traverse a common fibre link. Clearly, the chances of data bursts going through these paths and contending for the same wavelength channel at the common fibre link are minimized if their ingress nodes search for an available wavelength using opposite orderings of the wavelengths, that is, the ingress node of π_1 uses, for instance, λ_1, λ_2, …, λ_{W-1}, λ_W, whereas the ingress node of π_2 uses λ_W, λ_{W-1}, …, λ_2, λ_1. This simple scenario is illustrated is Fig. 2 for $W = 4$, where most of the burst traffic on π_1 (π_2) will go through λ_1, λ_2 (λ_4, λ_3). However, in realistic network scenarios, each routing path shares fibre links with several other paths and, consequently, it is not feasible to have opposite wavelength search orderings for each pair of overlapping paths. Still, as long as it is possible for two overlapping paths to have two different wavelength channels ranked as the highest priority wavelengths, the probability of wavelength contention among data bursts going through these paths is expected to be reduced. This observation constitutes the foundation of the heuristic traffic engineering approaches described in the following.

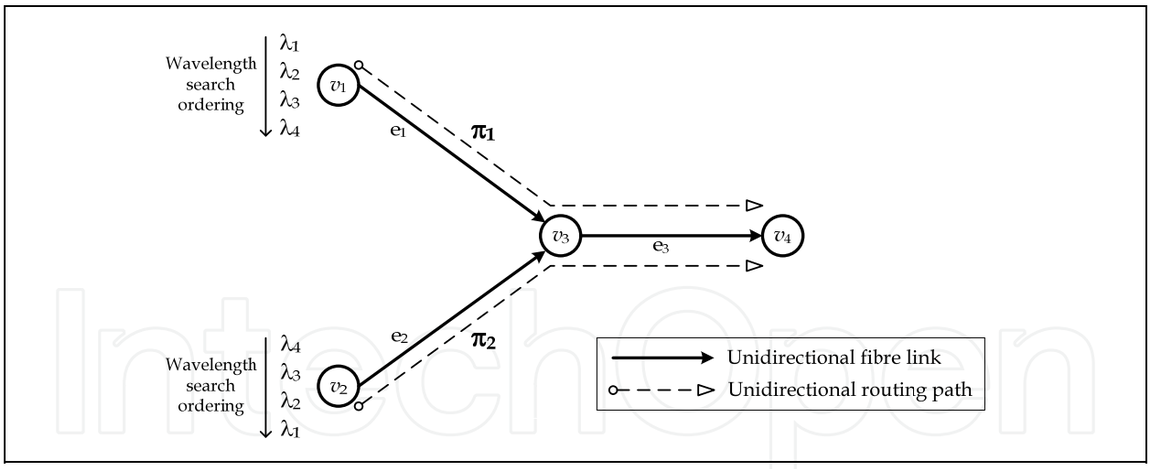


Fig. 2. Example OBS network with opposite wavelength search orderings.

2.2 Heuristic minimum priority interference

Intuitively, the chances of wavelength contention between data bursts going through different routing paths are expected to increase with both the average traffic load offered to the paths and with the number of common fibre links. Bearing this in mind, it is useful to define the concept of interference level of routing path π_i on routing path π_j with $i \neq j$ as,

$$I(\pi_i, \pi_j) = \gamma_i |E_i \cap E_j|, \quad (1)$$

where $|E_i \cap E_j|$ denotes the number of fibre links shared by both paths, and to define the combined interference level between routing paths π_i and π_j with $i \neq j$ as,

$$I^c(\pi_i, \pi_j) = I(\pi_i, \pi_j) + I(\pi_j, \pi_i) = (\gamma_i + \gamma_j)\,|E_i \cap E_j|\,. \tag{2}$$

The higher the combined interference level between two routing paths, the higher the likelihood that data bursts going through those paths will contend for the same fibre link resources. Consequently, routing paths with higher combined interference level should use wavelength search orderings as opposed as possible. This constitutes the basic principle exploited by First Fit-Traffic Engineering (FF-TE) (Teng & Rouskas, 2005), which was the first offline algorithm proposed to determine wavelength search orderings that are expected to reduce the probability of wavelength contention. However, this algorithm oversimplifies the problem resolution by computing a single wavelength search ordering for all the routing paths with the same ingress node. A detailed discussion of the limitations of the FF-TE algorithm is presented in (Pedro et al., 2006b). To overcome these shortcomings, the more advanced Heuristic Minimum Priority Interference (HMPI) algorithm, which computes an individual wavelength search ordering per routing path, is described below.

2.2.1 Algorithm description

The algorithm proposed in (Pedro et al., 2006b) for minimizing wavelength contention aims to determine an individual wavelength search ordering for each routing path with a reduced computational effort. The HMPI algorithm uses as input information the network topology, the routing paths and the average traffic load offered to the routing paths.

In order to determine the wavelength search ordering of a routing path, a unique priority must be assigned to each of the wavelengths. The wavelength ranked with the highest priority, called the primary wavelength, is expected to carry the largest amount of burst traffic going through the routing path. The other wavelengths, ordered by decreasing priority, expectedly carry diminishing amounts of burst traffic. In view of the importance of the primary wavelengths, the HMPI algorithm comprises a first stage dedicated to optimize them, consisting of the following three steps.

(**S1**) Reorder the routing paths of Π such that if $i < j$ one of the following conditions holds,

$$\sum_{\substack{\pi_k \in \Pi, \\ k \neq i}} I(\pi_i, \pi_k) > \sum_{\substack{\pi_k \in \Pi, \\ k \neq j}} I(\pi_j, \pi_k)\,; \tag{3}$$

$$\sum_{\substack{\pi_k \in \Pi, \\ k \neq i}} I(\pi_i, \pi_k) = \sum_{\substack{\pi_k \in \Pi, \\ k \neq j}} I(\pi_j, \pi_k) \text{ and } |E_i| > |E_j|\,. \tag{4}$$

(**S2**) Consider W sub-sets of the routing paths, one per wavelength, initially empty, that is, $|\Pi_j| = 0$ for $j = 1, \ldots, W$. Following the routing path ordering defined for Π, include path π_i in the sub-set Π_j such that for any $k \neq j$ one of the subsequent conditions holds,

$$\sum_{\substack{\pi_l \in \Pi_j, \\ l \neq i}} I^c(\pi_i, \pi_l) < \sum_{\substack{\pi_l \in \Pi_k, \\ l \neq i}} I^c(\pi_i, \pi_l)\,; \tag{5}$$

$$\sum_{\substack{\pi_l \in \Pi_j, \\ l \neq i}} I^c(\pi_i, \pi_l) = \sum_{\substack{\pi_l \in \Pi_k, \\ l \neq i}} I^c(\pi_i, \pi_l) \text{ and } |\Pi_j| > |\Pi_k|. \quad (6)$$

(**S3**) Select wavelength channel λ_j as the primary wavelength of all the paths in sub-set Π_j, that is,

$$P(\lambda_j, \pi_i) = \begin{cases} W, & \text{if } \pi_i \in \Pi_j \\ 0, & \text{otherwise} \end{cases}. \quad (7)$$

The first step of this stage of the HMPI algorithm is used to order the routing paths by decreasing interference level on the remaining paths. Ties are broken by giving preference to the longer routing paths. Considering W sub-sets of routing paths, the second step sequentially includes each routing path on the sub-set with minimum combined interference level between the routing path and the paths already included in the sub-set. Ties are broken by preferring the sub-set with larger number of paths. Finally, the third step assigns to all routing paths of a sub-set the primary wavelength associated with that sub-set. As a result of this stage, the routing paths with minimum combined interference level, carrying data bursts that are less prone to contend with each other for the same wavelength channel, will share the same primary wavelength.

In the second stage of the algorithm, the non-primary wavelengths for all routing paths are determined sequentially, starting with the second preferred wavelength channel and ending with the least preferred wavelength. When determining for each routing path the wavelength with priority $p < W$, it is intuitive to select one to which has been assigned, so far in the algorithm execution, the lowest priorities on routing paths that share fibre links with the routing path being considered. This constitutes the basic rule used in the second stage of the HMPI algorithm.

The following steps are executed for priorities $1 \leq p \leq W - 1$ in decreasing order and considering, for each priority p, all the routing paths according to the path ordering defined in the first stage of the algorithm.

(**S1**) Let $\Lambda = \{\lambda_j : P(\lambda_j, \pi_i) = 0, 1 \leq j \leq W\}$ denote the initial set of candidate wavelengths, containing all wavelengths that have been assigned a priority of zero on routing path π_i. If $|\Lambda| = 1$, go to (**S7**).

(**S2**) Let $\mathrm{P} = \{k : \exists \pi_l, l \neq i, P(\lambda_j, \pi_l) = k, |E_l \cap E_i| > 0, \lambda_j \in \Lambda\}$ be the set of priorities that have already been assigned to candidate wavelengths on paths that overlap with π_i.

(**S3**) Let $\psi = \min_{\lambda_j \in \Lambda} \max_{\pi_l \in \Pi} \{P(\lambda_j, \pi_l) : l \neq i, |E_l \cap E_i| > 0, P(\lambda_j, \pi_l) \in \mathrm{P}\}$ be the lowest priority among the set containing the highest priority assigned to each candidate wavelength on paths that share links with π_i. Update the set of candidate wavelengths as follows,

$$\Lambda \leftarrow \Lambda \setminus \{\lambda_j : \exists \pi_l, l \neq i, P(\lambda_j, \pi_l) > \psi, |E_l \cap E_i| > 0\}; \quad (8)$$

If $|\Lambda| = 1$, go to (**S7**).

(**S4**) Define $C(\lambda_j, e_m) = \sum\{\gamma_l : E_l \supset e_m, |E_l \cap E_i| > 0, P(\lambda_j, \pi_l) = \psi\}$ as the cost associated with wavelength channel $\lambda_j \in \Lambda$ on link $e_m \in E_i$ and $\alpha_e = \min_{\lambda_j \in \Lambda} \max_{e_m \in E_i} C(\lambda_j, e_m)$ as the minimum cost among the set containing the highest cost associated with each candidate wavelength on the fibre links of π_i. Update the set of candidate wavelengths as follows,

$$\Lambda \leftarrow \Lambda \setminus \{\lambda_j : \exists e_m, C(\lambda_j, e_m) > \alpha_e, e_m \in E_i\}; \quad (9)$$

If $|\Lambda| = 1$, go to (**S7**).

(**S5**) Define $C(\lambda_j, \pi_i) = \sum_{e_m \in E_i} C(\lambda_j, e_m)$ as the cost associated with wavelength λ_j on path π_i and $\alpha_\pi = \min_{\lambda_j \in \Lambda} C(\lambda_j, \pi_i)$ as the minimum cost among the costs associated with the candidate wavelengths on π_i. Update the set of candidate wavelengths as follows,

$$\Lambda \leftarrow \Lambda \setminus \{\lambda_j : C(\lambda_j, \pi_i) > \alpha_\pi\}; \quad (10)$$

If $|\Lambda| = 1$, go to (**S7**).

(**S6**) Update the set of priorities assigned to the candidate wavelengths as follows,

$$P \leftarrow P \setminus \{k : k \geq \psi\}; \quad (11)$$

If $|P| > 0$, go to (**S3**). Else, randomly select a candidate wavelength $\lambda \in \Lambda$.

(**S7**) Assign priority p to the candidate wavelength $\lambda \in \Lambda$ on path π_i, that is, $P(\lambda, \pi_i) = p$.

The first step of the second stage of the HMPI algorithm is used to define the candidate wavelength channels by excluding the ones that have already been assigned a priority larger than zero on the routing path, whereas the second step determines the priorities assigned to these wavelengths on paths that overlap with the routing path under consideration. The third, fourth and fifth step are used to reduce the number of candidate wavelengths. As soon as there is only one candidate wavelength, it is assigned to it the priority p on path π_i, concluding the iteration. In the third step, the highest priority already assigned to each of the candidate wavelength channels on paths that overlap with π_i is determined. Only the wavelengths with the lowest of these priorities are kept in the set of candidates. If needed, the fourth step tries to break ties by associating a cost with each candidate wavelength on each fibre link of π_i. This cost is given by the sum of the average traffic load offered to paths that traverse the fibre link and use the wavelength with priority ψ. The wavelengths whose largest link cost, among all links of π_i, is the smallest one (α_e) are kept as candidates. When there are still multiple candidate wavelengths, the fifth step associates a cost with each wavelength on path π_i, which is simply given by the sum of the cost associated to the wavelength on all links of the routing path. The candidate wavelengths with smallest path cost (α_π) are kept. If necessary, the sixth step removes the priorities equal or larger than ψ from the set of priorities assigned to candidate wavelengths on paths that overlap with the path being considered and repeats the iteration. Finally, if all priorities have been removed and there are still multiple candidate wavelengths, one of them is randomly selected.

As the outcome of executing the HMPI algorithm, each wavelength channel λ_j is assigned a unique priority on routing path π_i, $1 \leq P(\lambda_j, \pi_i) \leq W$. Equivalently, this solution for the priority assignment problem can be represented as an ordering of the W wavelengths, $\{\lambda^1(\pi_i), \lambda^2(\pi_i), \ldots, \lambda^j(\pi_i), \ldots, \lambda^W(\pi_i)\}$, where $\lambda^j(\pi_i)$ denotes the j^{th} wavelength channel to be searched when assigning a wavelength to data bursts directed to routing path π_i. In order to enforce these search orderings, each of these lists must be uploaded from the point where they are computed to the ingress nodes of the routing paths. Hence, assuming single-path routing, each ingress node will have to maintain at most $N - 1$ lists of ordered wavelengths.

The computational complexity of the HMPI algorithm, as derived in (Pedro et al., 2009c), is given by $O(W^2 \cdot |\Pi|^2)$, that is, in the worst case it scales with the square of the number of wavelength channels times the square of the number of routing paths.

2.2.2 Illustrative example

In order to give a better insight into the HMPI algorithm, consider the example OBS network of Fig. 3, which has 6 nodes and 8 fibre links (Pedro et al., 2009c). The number of routing paths used to transmit bursts in the network is $|\Pi| = 6$ and each fibre link supports a number of wavelength channels $W = 4$. Moreover, the average traffic load offered to each routing path is 1, except for routing path π_4, which has an average offered traffic load of 1.2, that is, $\gamma_i = 1$ for $i = 1, 2, 3, 5, 6$ and $\gamma_4 = 1.2$.

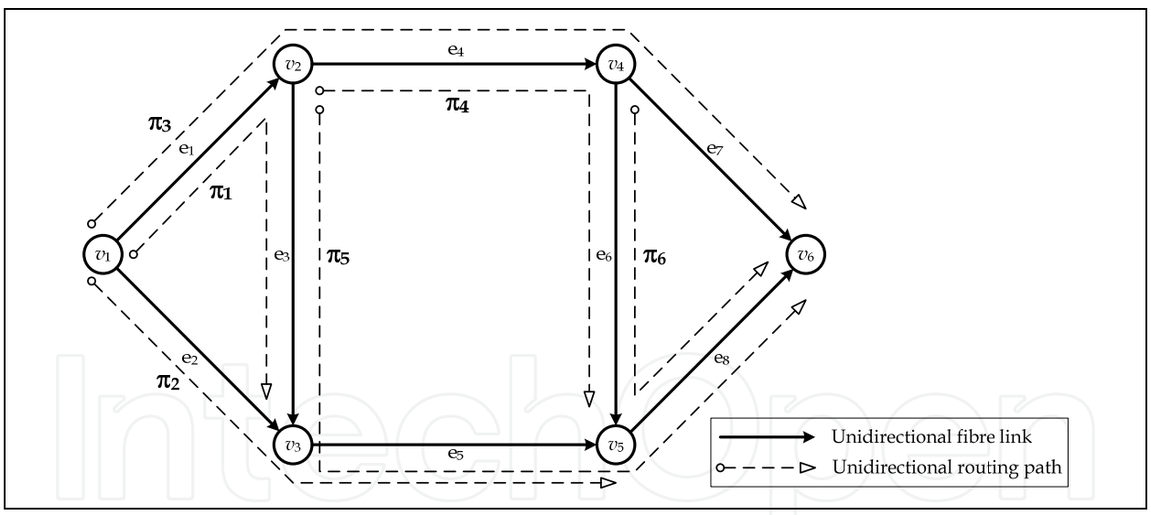


Fig. 3. OBS network used to exemplify the HMPI algorithm (Pedro et al., 2009c).

The HMPI algorithm starts by computing the interference level of all pairs of routing paths, as shown in Table 1. Step (**S1**) of the first stage of the algorithm orders the routing paths by decreasing order of their interference level over other paths, which results in the path order $\{\pi_5, \pi_4, \pi_3, \pi_1, \pi_6, \pi_2\}$. The path with the highest interference level over other paths is π_5, which overlaps with three paths, and the path with the second highest interference level over other paths is π_4, which overlaps with two paths. Although π_3, π_1 and π_6 also overlap with two paths, π_4 is offered more traffic load and consequently can cause more contention. In addition, π_3 precedes π_1 and π_6 because it is longer than the later paths. Since paths π_1 and

π_6 are tied, the path with the smallest index was given preference. Finally, the path with the lowest interference level over other paths is π_2.

$I(\pi_i, \pi_j)$	π_1	π_2	π_3	π_4	π_5	π_6
π_1	--	0	1	0	1	0
π_2	0	--	0	0	1	0
π_3	1	0	--	1	0	0
π_4	0	0	1.2	--	0	1.2
π_5	1	1	0	0	--	1
π_6	0	0	0	1	1	--

Table 1. Interference level of the routing paths.

Step (**S2**) starts by creating one sub-set of routing paths per wavelength, that is, Π_1, Π_2, Π_3, Π_4. Following the determined path order, π_5 is included in the first empty sub-set, Π_1. Path π_4 is also included in Π_1, because $I^C(\pi_4, \pi_5) = 0$ and Π_1 has more paths than the remaining sub-sets. Since path π_3 overlaps with π_4, $I^C(\pi_3, \pi_4) = 2.2$, and π_4 is already included in Π_1, π_3 is included in the empty sub-set Π_2. Moreover, path π_1 overlaps with both π_5 and π_3 and thus it is included in empty sub-set Π_3. Path π_6 can be included in sub-sets Π_2 and Π_3 because it only overlaps with the paths of Π_1. The tie is broken by selecting the sub-set with smallest index, that is, Π_2. Similarly, path π_2 is also included in this sub-set as it does not overlap with the paths in Π_2 and Π_3 and $|\Pi_2| > |\Pi_3|$. Since every path has been included in one sub-set, $\Pi_1 = \{\pi_4, \pi_5\}$, $\Pi_2 = \{\pi_2, \pi_3, \pi_6\}$ and $\Pi_3 = \{\pi_1\}$, step (**S3**) concludes the first stage of the algorithm by making λ_1 the primary wavelength of paths π_4 and π_5, λ_2 the primary wavelength of paths π_2, π_3 and π_6, and λ_3 the primary wavelength of path π_1. The other wavelengths are temporarily assigned priority 0 on the routing paths. Table 2 shows the priorities assigned to the wavelengths on the routing paths after the entire HMPI algorithm has been executed.

$P(\lambda_j, \pi_i)$	π_1	π_2	π_3	π_4	π_5	π_6
λ_1	1	1	1	4	4	1
λ_2	2	4	4	1	1	4
λ_3	4	3	2	2	2	3
λ_4	3	2	3	3	3	2

Table 2. Wavelengths priority on the routing paths.

The second stage of the algorithm is initiated with $p = 3$ and proceeds path by path according to the order already defined. For path π_5, the algorithm starts by creating the initial set of candidate wavelengths, $\Lambda = \{\lambda_2, \lambda_3, \lambda_4\}$, in (**S1**). Since this path overlaps with π_1, π_2 and π_6, the set of priorities assigned to wavelengths of Λ on these paths, determined in (**S2**), is $P = \{0, 4\}$. Wavelength λ_4 is assigned priority 0 on all paths that overlap with π_5 and thus $\rho = 0$. Accordingly, in (**S3**) the set of candidate wavelengths is updated, $\Lambda = \{\lambda_4\}$, and λ_4 is assigned priority 3 on path π_5. For path π_4, $\Lambda = \{\lambda_2, \lambda_3, \lambda_4\}$, $P = \{0, 4\}$, and $\rho = 0$. The set of

candidate wavelengths is updated to $\Lambda = \{\lambda_3, \lambda_4\}$, because both λ_3 and λ_4 are assigned priority 0 on paths that overlap with π_4. In this particular case, the algorithm cannot break the tie and in (**S7**) randomly selects wavelength λ_4 to be assigned priority 3 on path π_4. For the remaining paths, there is only one candidate wavelength whose priority on other paths equals ρ. Wavelength λ_4 is assigned priority 3 on paths π_3 and π_1 and wavelength λ_3 is assigned this priority on paths π_6 and π_2.

The second stage of the algorithm is executed again, but with $p = 2$. For path π_5, the initial set of candidate wavelengths is $\Lambda = \{\lambda_2, \lambda_3\}$. Both wavelengths are assigned priority 4 on at least one of the paths that overlaps with π_5 ($\rho = 4$), λ_2 on π_2 and π_6 and λ_3 on π_1. Paths π_1, π_2, and π_6 share with π_5 links e_3, e_5 and e_8, respectively, and the average traffic load offered to these paths is 1. Thus, according to (**S4**), the cost associated with λ_2 and λ_3 on each link is at most 1 ($\alpha_e = 1$). However, λ_2 has this link cost on two links, which in (**S5**) results in a cost $C(\lambda_2, \pi_5) = 2$, whereas λ_3 has this link cost on a single link, $C(\lambda_3, \pi_5) = 1$. Consequently, $\alpha_\pi = 1$ and the set of candidate wavelengths is updated to $\Lambda = \{\lambda_3\}$. For path π_4, $\Lambda = \{\lambda_2, \lambda_3\}$, $P = \{0, 3, 4\}$, and $\rho = 3$. Only wavelength λ_3 is used with a priority smaller or equal than 3 in all links, which reduces the set of candidates to λ_3. In the case of path π_3, $\Lambda = \{\lambda_1, \lambda_3\}$ and λ_1 is assigned priority 4 on π_4, whereas λ_3 is assigned this priority on π_1. Since $\gamma_4 > \gamma_1$, the highest link cost associated to λ_1 is larger than that for λ_3, and the candidate wavelengths are reduced to λ_3. For path π_1, $\Lambda = \{\lambda_1, \lambda_2\}$ and both these wavelengths observe $\rho = 4$, $\alpha_e = 1$ and $\alpha_\pi = 1$. The algorithm has to randomly select one of the wavelengths (λ_2). For both π_6 and π_2, $\Lambda = \{\lambda_1, \lambda_4\}$, $\rho = 3$, but only λ_4 is assigned a priority smaller or equal to 3 in all of the links. The set of candidate wavelengths is reduced to $\Lambda = \{\lambda_4\}$.

Finally, for $p = 1$ the wavelength assignment is trivial, because there is only one wavelength still assigned priority 0 on each path. The complete wavelength search ordering of each path can be obtained from Table 2. The following observations show that these orderings should effectively reduce contention. Firstly, overlapping paths do not share the same primary wavelength. Instead, primary wavelengths are reused by link-disjoint routing paths (e.g., λ_2 is the primary wavelength of π_2, π_3 and π_6). Secondly, paths use with smallest possible priority the primary wavelengths of overlapping paths (e.g., π_1, π_2 and π_6 overlap with π_5 and use the primary wavelength of this path with priority 1).

3. Traffic engineering in the wavelength domain

Noteworthy, at the ingress edge nodes of an OBS network, data bursts are kept in electronic buffers before a wavelength channel is assigned to them and they are transmitted optically towards the egress edge nodes. Clearly, the flexibility of scheduling data bursts in the wavelength channels is considerably higher when the bursts are still buffered at the ingress nodes than when they have already been converted to the optical domain. For instance, a data burst can be delayed at one of the ingress buffers by the exact amount of time required for a wavelength channel to become available in the designated output fibre link. This procedure is not possible at the core nodes due to the lack of optical RAM. The capability of delaying data bursts at an ingress node by a random amount of time, not only increases the chances of successfully scheduling bursts at the output fibre link of their ingress nodes, but also enables implementing strategies that reduce in advance the probability of contention at the core nodes.

The Burst Overlap Reduction Algorithm proposed in (Li & Qiao, 2004) exploits the additional degree of freedom provided by delaying data bursts at the electronic buffers of the ingress nodes to shape the burst traffic departing from these nodes in such way that the probability of contention at the core nodes can be reduced. The principle underlying BORA is that a decrease on the number of different wavelength channels allocated to the data bursts assembled at an ingress node can smooth the burst traffic at the input fibre links of the core nodes and, as a result, reduce the probability that the number of overlapping data bursts directed to the same output fibre link exceeds the number of wavelength channels. In its simpler implementation, BORA relies on using the same wavelength search ordering at all the ingress nodes of the network and utilizing the buffers in these nodes to transmit the maximum number of bursts in the first wavelength channels according to such ordering. In order to limit the extra transfer delay incurred by data bursts, as well as the added buffering and processing requirements, the ingress node can impose a maximum ingress burst delay, $\Delta t_{\max}^{\mathrm{RAM}}$, defined as the maximum amount of time a data burst can be kept at an electronic buffer of its ingress node excluding the time required to assemble the burst and the offset time between the data burst and its correspondent BHP.

The concept of BORA is appealing in OBS networks with wavelength conversion, since these algorithms have not been designed to mitigate wavelength contention. Moreover, BORA algorithms do not account for the capacity fragmentation of the wavelength channels, which is also a performance limiting factor in OBS networks. These limitations have motivated the development of a novel strategy in (Pedro et al., 2009b) that also exploits the electronic buffers of the ingress edge nodes to selectively delay data bursts, while providing a twofold advantage over BORA: enhanced contention minimization at the core nodes and support of core node architectures with relaxed wavelength conversion capabilities.

The first principle of the proposed strategy is related with the availability of RAM at the ingress nodes. In the process of judiciously delaying bursts to schedule them using the smallest number of different wavelength channels, the delayed bursts can be scheduled with minimum voids between them and the preceding bursts already scheduled on the same wavelength channel. This is only possible because the bursts assembled at the node can be delayed by a random amount of time. The serialization of data bursts not only smoothes the burst traffic, with the consequent decrease of the chances of contention at the core nodes, but also reduces the fragmentation of the wavelengths capacity at the output fibre links of the ingress nodes. These serialized data bursts traverse the core nodes, where some of them must be converted to other wavelength channels to resolve contention. The wavelength conversions break the series of data bursts and, as a result, create voids between a burst converted to another wavelength channel and the bursts already scheduled on this wavelength. A large number of these voids lead to wasting bandwidth, as the core nodes will not be able to use them to carry data.

In essence, the first key principle consists of serializing data bursts at the ingress nodes to mitigate the voids between them. Noticeably, if these bursts traverse a set of common fibre links without experiencing wavelength conversion, the formation of unusable voids is reduced at those links. Hence, the second key principle of the proposed strategy consists of improving the probability that serialized bursts routed via the same path are kept in the same wavelength channel for as long as possible. This can reduce the number of unusable

voids created in the fibre links traversed before wavelength conversion is used, improving network performance.

The task of keeping the data bursts, which are directed to the same routing path and have been serialized at the ingress node, in the same wavelength channel requires minimizing the chances that bursts on overlapping routing paths contend for the same wavelength channel and, as a result, demand wavelength conversion. This objective is the same as that of the HMPI algorithm presented in Section 2. For that reason, the strategy proposed in (Pedro et al., 2009b), which is designated as Traffic Engineering in the wavelength domain with Delayed Burst Scheduling (TE-DBS), combines the wavelength contention minimization capability of HMPI with selectively delaying data bursts at the electronic buffers of their ingress nodes not only to smooth burst traffic, but also to maximize the amount of data bursts carried in the wavelength channels ranked with the highest priorities by HMPI.

The key principles of the TE-DBS strategy can be illustrated with the example of Fig. 4. The OBS network depicted comprises six nodes and five fibre links. Three paths, π_1, π_2, and π_3,

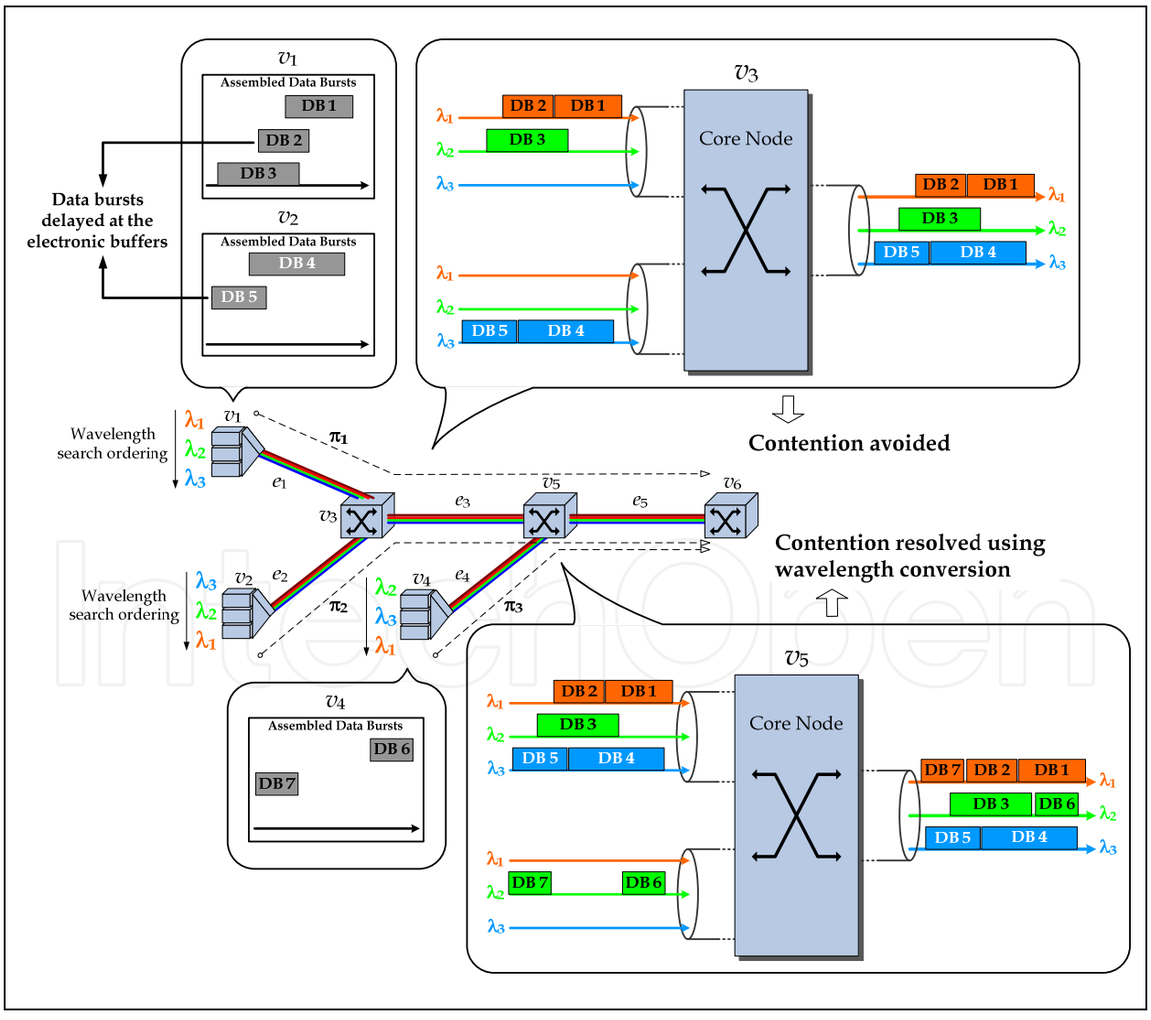


Fig. 4. Example of using TE-DBS to minimize contention at the core nodes.

are used to transmit bursts between one of the three ingress nodes, v_1, v_2, and v_4, and node v_6. Contention between bursts from different input fibre links and directed to the same output fibre link can occur at core nodes v_3 and v_5. Each ingress node uses its own wavelength search ordering and selectively delays bursts with the purpose of transmitting them on the wavelength channels which have been ranked with the highest priorities by an algorithm for minimizing contention in the wavelength domain. Similarly to what occurs with BORA, a maximum ingress burst delay, $\Delta t_{\max}^{\mathrm{RAM}}$, is imposed at each ingress node.

As can be seen, v_1 has assembled three data bursts (DB 1, DB 2, and DB 3), which overlap in time, and v_2 has assembled two data bursts (DB 4 and DB 5), which also overlap in time. The first two bursts assembled by v_1 are transmitted in wavelength channel λ_1, whereas the third cannot be transmitted in this wavelength without infringing the maximum ingress burst delay and, therefore, has to be transmitted in λ_2. The two bursts assembled by v_2 are transmitted in the wavelength ranked with highest priority, λ_3. These bursts traverse v_3, where contention is avoided since the bursts arrive in different wavelengths. Meanwhile, the ingress node v_4 has assembled two data bursts (DB 6 and DB 7) and transmits them in the wavelength ranked with highest priority, λ_2. All seven data bursts traverse core node v_5, where DB 7 must be converted to another wavelength in order to resolve contention.

The major observations provided by this example are as follows. Similarly to using BORA, the burst traffic is smoothed at the ingress nodes, reducing contention at the core nodes from an excessive number of data bursts directed to the same output fibre link. Moreover, since the burst traffic of routing paths π_1, π_2, and π_3 is mostly carried in different wavelengths, contention for the same wavelength channel is also reduced. As a result, the pairs of bursts serialized at the ingress nodes, DB 1 and DB 2 in routing path π_1 and DB 4 and DB5 in routing path π_2, can be kept in the same wavelength channel until they reach node v_6, mitigating the fragmentation of the capacity of wavelengths λ_1 and λ_3 in the fibre links traversed by routing paths π_1 and π_2. Since this is accomplished through minimizing the probability of wavelength contention, it can also relax the wavelength conversion capabilities of the core nodes without significantly degrading network performance.

The TE-DBS strategy requires the computation of one wavelength search ordering, $\{\lambda^1(\pi_i), \lambda^2(\pi_i), \ldots, \lambda^W(\pi_i)\}$, for each routing path π_i. The HMPI algorithm is used to optimize offline the wavelength search orderings. These orderings are stored at the ingress nodes and the control unit of these nodes uses them for serializing data bursts on the available wavelength channel ranked with the highest priority on the routing path the bursts will follow.

4. Results and discussion

This section presents a performance analysis of the framework for traffic engineering in the wavelength domain TE-DBS, described in the Section 3, and assuming the HMPI algorithm, detailed in Section 2, is employed offline to optimize the wavelength search ordering for each routing path in the network.

The results are obtained via network simulation using the event-driven network simulator described in (Pedro et al., 2006a). The network topology used in the performance study is a 10-node ring network. All of the network nodes have the functionalities of both edge and

core nodes and the resource reservation is made using the JET protocol. It is also assumed that all the wavelength channels in a fibre link have a capacity μ = 10 Gb/s, the time required to configure an optical space switch matrix is t_g = 1.6 μs, each node can process the BHP of a data burst in t_p = 1 μs and the offset time between BHP and data burst is given by $t_g + h_i \cdot t_p$, where h_i is the number of hops of burst path $\pi_i \in \Pi$. The switch matrix of each node is assumed to be strictly non-blocking. Unless stated otherwise, the simulation results were obtained assuming W = 32 wavelength channels per fibre link.

The traffic pattern used in the simulations is uniform, in the sense that a burst generated at an ingress node is randomly destined to one of the remaining nodes. Bursts are always routed via the shortest path. Both the data burst size and the burst interarrival time are negative-exponentially distributed. An average burst size of 100 kB is used, which results in an average burst duration of 80 μs. In the network simulations, increasing the average offered traffic load is obtained through reducing the average burst interarrival time. The average offered traffic load normalized to the network capacity is given by,

$$\Gamma = \frac{\sum_{\pi_i \in \Pi} \gamma_i \cdot h_i^{SP}}{L \cdot W \cdot \mu}, \tag{12}$$

where h_i^{SP} is the number of links traversed between the edge nodes of $\pi_i \in \Pi$.

In OBS networks, the most relevant performance metric is the average burst blocking probability, which measures the average fraction of burst traffic that is discarded by the network. The network performance can also be evaluated via the average offered traffic load that results in an objective average burst blocking probability B_{obj}. This metric is estimated by performing simulations with values of Γ spaced by 0.05, determining the load values between which the value with blocking probability B_{obj} is located and then using linear interpolation (with logarithmic scale for the average burst blocking probability). All of the results presented in this section were obtained through running 10 independent simulations for calculating the average value of the performance metric of interest, as well as a 95% confidence interval on this value. However, these confidence intervals were found to be so narrow that have been omitted from the plots for improving readability.

The majority of OBS proposals assumes the utilization of full-range wavelength converters deployed in a dedicated configuration, that is, one full-range wavelength converter is used at each output port of the switch matrix, as illustrated in Fig. 5. Each full-range wavelength converter must be capable of converting any wavelength at its input to a fixed wavelength at its output and if a node has M output fibres, its total number of converters is $M \cdot W$.

Fig. 6 plots the average burst blocking probability as a function of the maximum ingress burst delay for different values of the offered traffic load and considering both TE-DBS and the previously described BORA strategy. It also displays the blocking performance that corresponds to delaying bursts at the ingress nodes whenever a free wavelength channel is not immediately found. More precisely, the DBS strategy consists of delaying a data burst at its ingress node by the minimum amount of time, upper-bounded to the maximum ingress burst delay, such that one wavelength becomes available in the output fibre link.

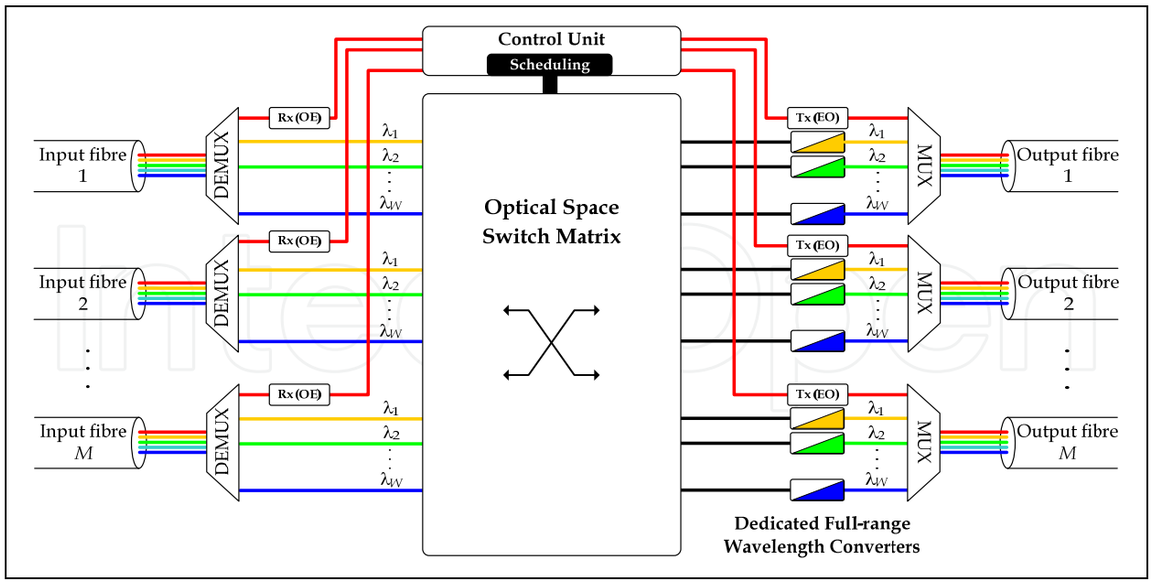


Fig. 5. OBS core node architecture with dedicated full-range wavelength converters.

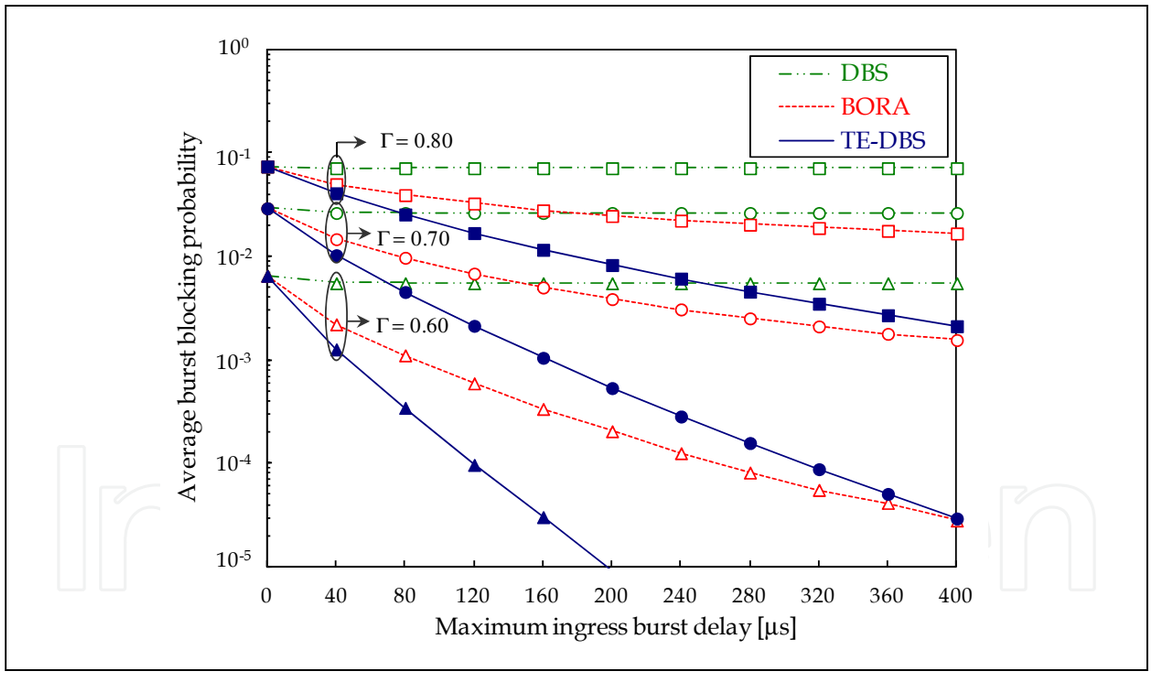


Fig. 6. Network performance with dedicated full-range wavelength converters for different values of the average offered traffic (Pedro et al., 2009a).

The curves for DBS show that exploiting the electronic buffers at the ingress nodes only for contention resolution does not improve blocking performance. On the contrary, with both BORA and TE-DBS the average burst blocking probability is decreased as the maximum ingress burst delay is increased, confirming that these strategies proactively reduce the probability of contention by selectively delaying bursts at their ingress nodes.

The results also indicate TE-DBS is substantially more efficient than BORA in exploiting larger maximum ingress burst delays to reduce the burst blocking probability. The proposed strategy outperforms BORA for the same maximum ingress burst delay or, alternatively, requires a smaller maximum ingress burst delay to attain the same blocking performance of BORA. Particularly, the decrease rate of the burst losses with increasing the maximum ingress burst delay is considerably larger with TE-DBS than that with BORA. In addition, with TE-DBS the slope of the curves of the burst blocking probability is much steeper for smaller values of the average offered traffic load, a trend less pronounced with BORA.

Table 3 presents the average traffic load that can be offered to the network as to support an objective average burst blocking probability, B_{obj}, of 10^{-3} and 10^{-4}. The results include two values of the maximum ingress burst delay for BORA and TE-DBS, Δt_{max}^{RAM} = 200 μs and Δt_{max}^{RAM} = 400 μs, and the case of immediate burst scheduling at the ingress nodes, $\Delta t_{max}^{RAM} = 0$.

B_{obj}	$\Delta t_{max}^{RAM} = 0$	Δt_{max}^{RAM} = 200 μs		Δt_{max}^{RAM} = 400 μs	
		BORA	TE-DBS	BORA	TE-DBS
10^{-3}	0.522	0.654	0.723	0.689	0.782
10^{-4}	0.453	0.584	0.659	0.632	0.729

Table 3. Average offered traffic load for an objective average burst blocking probability of 10^{-3} and 10^{-4} (Pedro et al., 2009a).

The OBS network supports more offered traffic load for the same average burst blocking probability when using the TE-DBS and BORA strategies instead of employing immediate burst scheduling. In addition, the former strategy provides the largest improvements in supported offered traffic load. For instance, with $B_{obj} = 10^{-3}$, the network supports 32% more offered traffic load when using BORA with a maximum ingress burst delay of 400 μs instead of immediate burst scheduling, whereas when using the TE-DBS strategy the performance improvement is more expressive, enabling an increase of 50% in offered traffic load.

In order to provide evidence of the principles underlying contention minimization with BORA and TE-DBS, the first set of results differentiates the burst blocking probability at the ingress nodes (ingress bursts) and at the core nodes (transit bursts). Fig. 7 plots the average burst blocking probability, discriminated in terms of ingress bursts and transit bursts, as a function of the maximum ingress burst delay for $\Gamma = 0.70$.

The plot shows that without additional delays at the ingress nodes, the blocking probability of ingress bursts and of transit bursts are of the same order of magnitude. However, as the maximum ingress burst delay is increased, the blocking probability of ingress bursts is rapidly reduced, as a result of the enhanced ability of ingress nodes to buffer bursts during longer periods of time. This holds for the three channel scheduling algorithms. Therefore, the average burst blocking probability of transit bursts becomes the dominant source of blocking. Notably, using DBS does not reduce burst losses at the core nodes, rendering this strategy useless, whereas BORA and TE-DBS strategies exploit the selective ingress delay to reduce blocking of transit bursts. Moreover, TE-DBS is increasingly more effective than BORA in reducing these losses, which supports its superior performance displayed in Fig. 6.

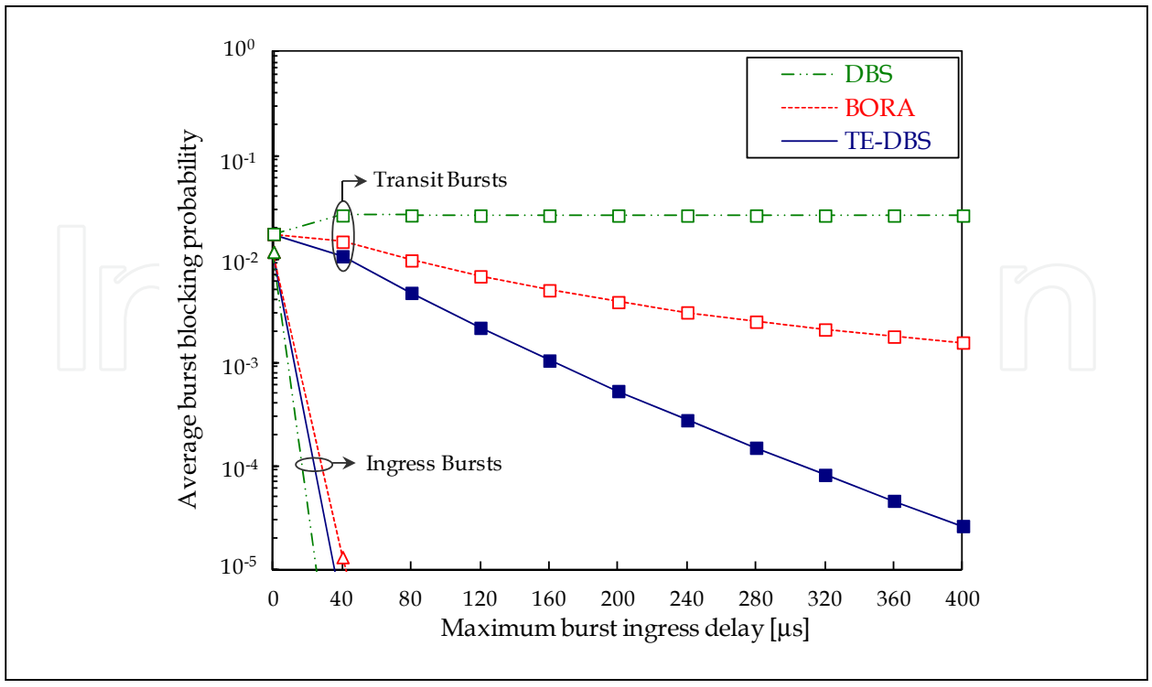


Fig. 7. Average burst blocking probability of ingress and transit bursts (Pedro et al., 2009a).

The major dissimilarity between the TE-DBS and BORA strategies is the order by which free wavelength channels are searched to schedule the data bursts assembled at the ingress nodes. Particularly, the TE-DBS strategy exploits the selective delaying of data bursts at the electronic buffers of these nodes not only to smooth the burst traffic entering the core network, similarly to BORA, but also to proactively reduce the unusable voids formed between consecutive data bursts scheduled in the same wavelength channel. As described in Section 3, complying with the latter objective demands enforcing that the serialized data bursts are kept in the same wavelength for as long as possible along their routing path, which means that contention for the same wavelength among bursts on overlapping paths must be minimized. Intuitively, the success of keeping the serialized data bursts in the same wavelength channel for as long as possible should be visible in the form of a reduced number of bursts experiencing wavelength conversion at the core nodes. In order to observe this effect, Fig. 8 presents the average wavelength conversion probability, defined as the fraction of transit data bursts that undergo wavelength conversion, as a function of the maximum ingress burst delay for different values of the average offered traffic load.

The curves for TE-DBS exhibit a declining trend as the maximum ingress burst delay increases, with this behaviour being more pronounced for smaller average offered traffic load values. These observations confirm that the probability of the data bursts serialized at the ingress nodes being kept in the same wavelength channel, as they go through the core nodes, is higher for larger values of the maximum ingress burst delay and smaller values of offered traffic load. Conversely, with BORA the wavelength conversion probability remains insensitive to variations in both the maximum ingress burst delay and offered traffic load, corroborating the fact that it cannot reduce the utilization of wavelength conversion at the core nodes. The reduced wavelength contention characteristic of the TE-DBS strategy, which

is absent in BORA, is critical to mitigate the fragmentation of the wavelengths capacity, resulting in the smaller transit burst losses reported with TE-DBS in Fig. 7 and ultimately explaining the enhanced blocking performance provided by this strategy.

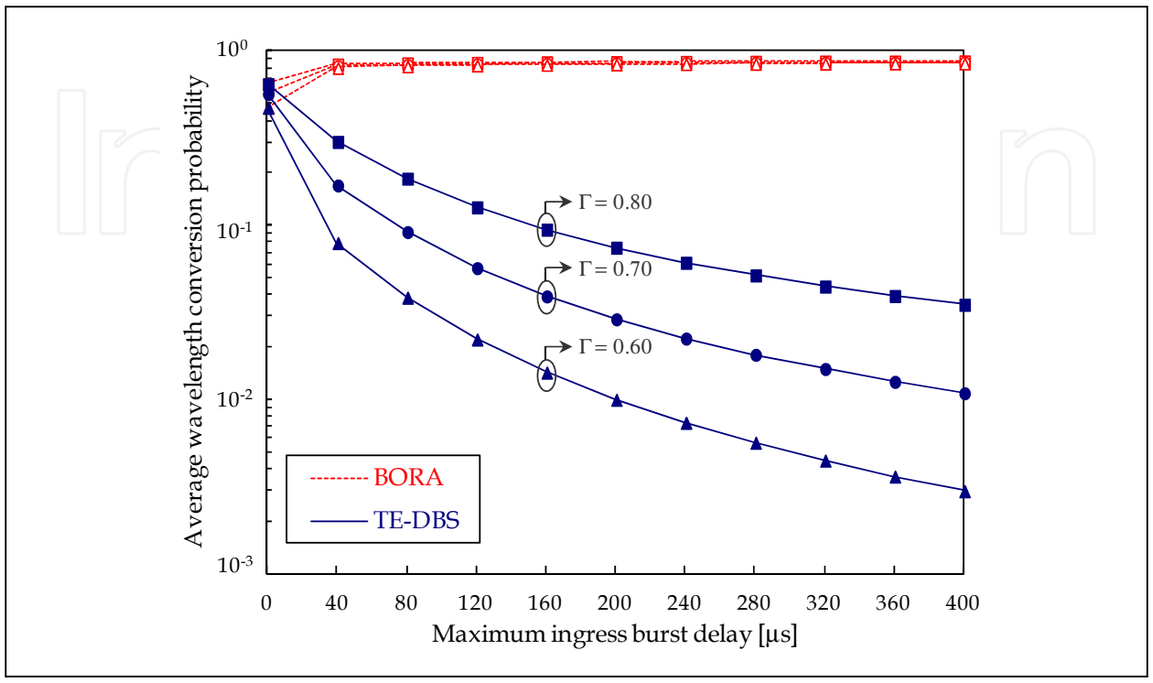


Fig. 8. Average wavelength conversion probability (Pedro et al., 2009b).

Fig. 9 shows the blocking performance as a function of the maximum ingress burst delay for different numbers of wavelength channels and Γ = 0.80. The results indicate that the slope of the average burst blocking probability curves for TE-DBS increases with the number of wavelength channels, augmenting the performance gain of using this strategy instead of BORA. This behaviour is due to the fact that when the number of wavelength channels per fibre link increases the effectiveness of the HMPI algorithm in determining appropriate wavelength search orderings improves, enhancing the isolation degree of serialized burst traffic from overlapping routing paths on different wavelength channels.

In principle, only a fraction of transit bursts experience wavelength contention, demanding the use of a wavelength converter. Consequently, the deployment of a smaller number of converters, in a shared configuration, has been proposed in the literature. Converter sharing at the core nodes can be implemented on a per-link or per-node basis, depending on whether each converter can only be used by bursts directed to a specific output link or can be used by bursts directed to any output link of the node (Chai et al., 2002). The latter sharing strategy enables to deploy a smaller number of converters. Fig. 10 exemplifies the architecture of a core node with C full-range wavelength converters shared per-node, where $C \leq M \cdot W$. In this core node architecture, each wavelength converter must be capable of converting any wavelength channel at its input to any wavelength channel at its output and the switch matrix has to be augmented with C input ports and C output ports.

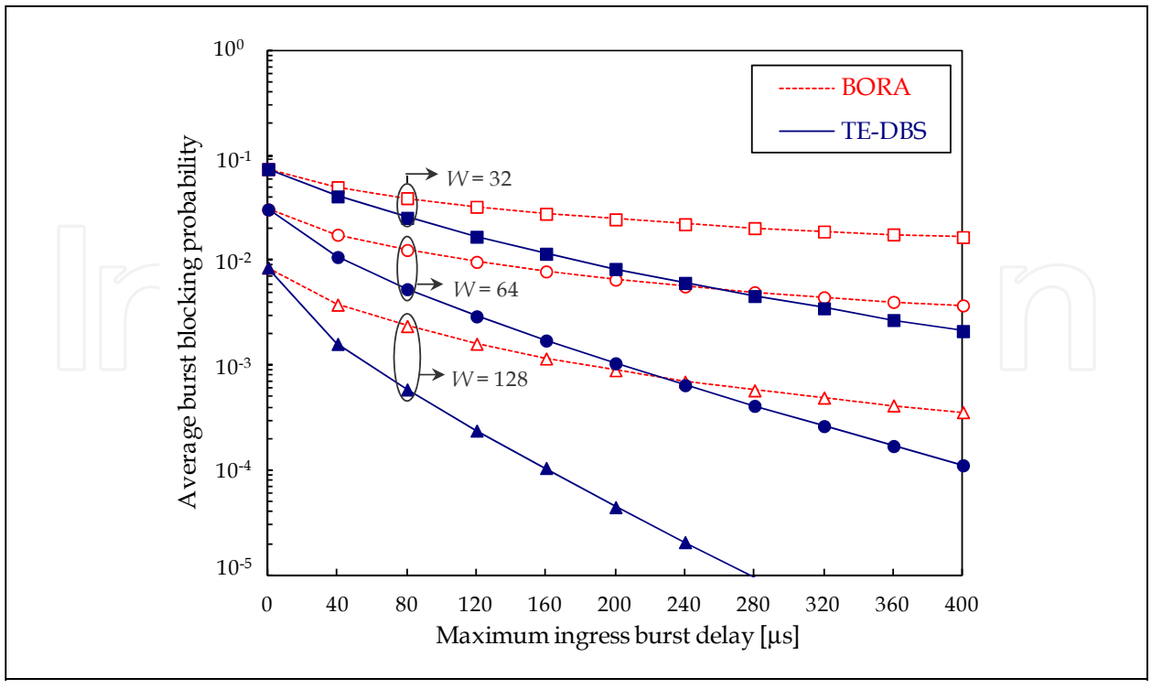


Fig. 9. Network performance for different numbers of wavelength channels (Pedro et al., 2009b).

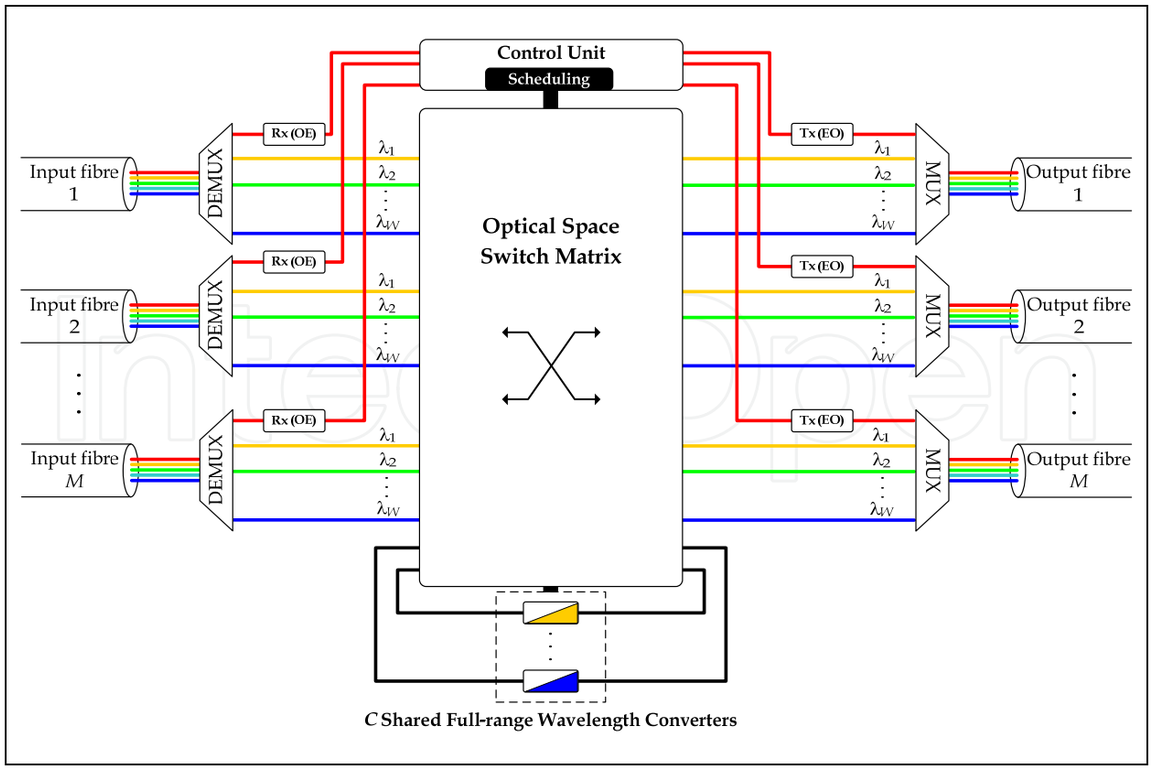


Fig. 10. OBS core node architecture with shared full-range wavelength converters.

The minimization of wavelength contention experienced by transit bursts is a key enabler for TE-DBS to improve the loss performance of OBS networks. Particularly, the simulation results presented in Fig. 8 confirm that the utilization of this strategy reduces the probability of wavelength conversion, and consequently the utilization of the wavelength converters, as the maximum ingress burst delay is increased. This attribute can extend the usefulness of TE-DBS to OBS networks with shared full-range wavelength converters because in this network scenario the lack of available converters at the core nodes can become the major cause of unresolved contention, specially for small values of C.

In order to illustrate the added-value of the TE-DBS strategy in OBS networks whose core nodes have shared full-range wavelength converters, consider the 10-node ring network with $W = 32$. When using wavelength converters in a dedicated configuration, each node of this network needs $M \cdot W = 64$ converters. Fig. 11 plots the average burst blocking probability as a function of the number of shared full-range wavelength converters at the nodes, C, for different values of the average offered traffic load and using BORA and TE-DBS strategies with $\Delta t_{\max}^{\mathrm{RAM}} = 160$ μs.

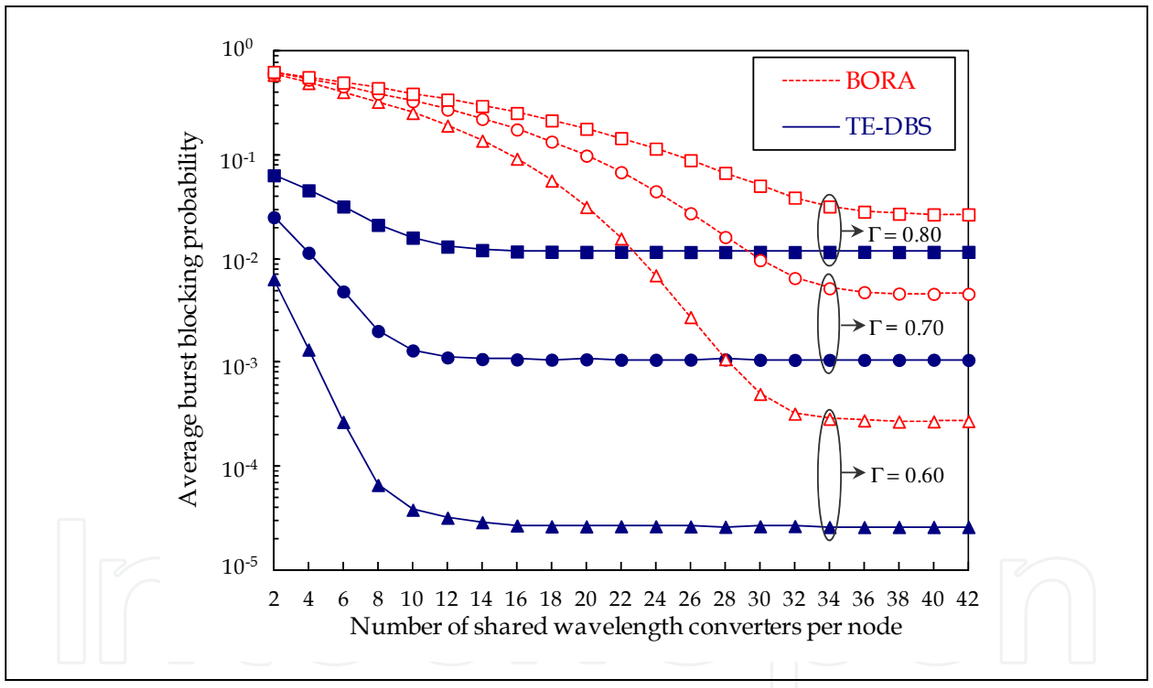


Fig. 11. Network performance with shared full-range wavelength converters for different values of the average offered traffic load (Pedro et al., 2009a).

The blocking performance curves clearly show that the OBS network using TE-DBS can benefit not only in terms of enhanced blocking performance, but also from enabling using simplified core node architectures. More precisely, the burst loss curves indicate that for very small numbers of shared wavelength converters, the utilization of TE-DBS results in a burst blocking probability that can be multiple orders of magnitude lower than that obtained using BORA. Furthermore, using TE-DBS demands a much smaller number of shared wavelength converters to match the blocking performance of a network using core

nodes with dedicated wavelength converters. Particularly, with TE-DBS around 16 shared converters per node are enough to match the loss performance obtained with 64 dedicated converters, whereas with BORA this number more than doubles, since around 36 shared converters are required. The larger savings in the number of wavelength converters enabled by TE-DBS also mean that the expansion of the switch matrix to accommodate the shared converters is smaller, leading to an even more cost-effective network solution.

5. Conclusions

Optical burst switching is seen as a candidate technology for next-generation transport networks. This chapter has described and analyzed the performance benefits of a strategy to enforce traffic engineering in the wavelength domain in OBS networks. The TE-DBS strategy is based on using the HMPI algorithm to optimize offline the order by which wavelength channels are searched for each routing path and employing at the ingress nodes a selective delaying of data bursts as a way to maximize the amount of burst traffic sent via the wavelength channels ranked with highest priority. Both the HMPI offline algorithm and the online selective delaying of bursts were revisited and exemplified.

A network simulation study has highlighted the performance improvements attained by using TE-DBS in an OBS network with dedicated full-range wavelength converters and with shared full-range wavelength converters. It was shown that the utilization of the TE-DBS strategy enables to reduce the average burst blocking probability for a given average offered traffic load, or augment the average offered traffic load for an objective burst blocking probability, when compared to utilizing a known contention minimization strategy. The simulation results shown that increasing the maximum delay a burst can experience at the ingress node and augmenting the number of wavelength channels per link can improve the effectiveness of the TE-DBS strategy and also provided evidence of the burst serialization and traffic isolation in different wavelengths inherent to this strategy. Finally, the analysis confirms that the utilization of TE-DBS in OBS networks with shared full-range wavelength converters can provide noticeable savings in the number of expensive all-optical wavelength converters and a smaller increase in the size of the switch matrix of the core nodes.

6. References

Barakat, N. & Darcie, T. (2007). The Control-Plane Stability Constraint in Optical Burst Switching Networks. *IEEE Communications Letters*, Vol. 11, No. 3, (March 2007), pp. 267-269, ISSN 1089-7798

Chai, T.; Cheng, T. ; Shen, G.; Bose, S. & Lu, C. (2002). Design and Performance of Optical Cross-Connect Architectures with Converter Sharing. *Optical Networks Magazine*, Vol. 3, No. 4, (July/August 2002), pp. 73-84, ISSN 1572-8161

Chang, G.; Yu, J. ; Yeo, Y.; Chowdhury, A. & Jia, Z. (2006). Enabling Technologies for Next-Generation Packet-Switching Networks. *Proceedings of the IEEE*, Vol. 94, No. 5, (May 2006), pp. 892-910, ISSN 0018-9219

Chen, Y.; Qiao, C. & Yu, X. (2004). Optical Burst Switching: A New Area in Optical Networking Research. *IEEE Network*, Vol. 18, No. 3, (May/June 2004), pp. 16-23, ISSN 0890-8044

IETF (2002). *RFC 3945: Generalized Multi-Protocol Label Switching (GMPLS) Architecture*, Internet Engineering Task Force, September 2002

ITU-T (2006). *Recommendation G.8080: Architecture for the Automatically Switched Optical Network (ASON)*, International Telecommunication Union – Telecommunication Standardization Sector, June 2006

Korotky, S. (2004). Network Global Expectation Model: A Statistical Formalism for Quickly Quantifying Network Needs and Costs. *IEEE/OSA Journal of Lightwave Technology*, Vol. 22, No. 3, (March 2004), pp. 703-722, ISSN 0733-8724

Li, J. & Qiao, C. (2004). Schedule Burst Proactively for Optical Burst Switched Networks. *Computer Networks*, Vol. 44, (2004), pp. 617-629, ISSN 1389-1286

Papadimitriou, G.; Papazoglou, C. & Pomportsis, A. (2003). Optical Switching: Switch Fabrics, Techniques, and Architectures. *IEEE/OSA Journal of Lightwave Technology*, Vol. 21, No. 2, (February 2003), pp. 384-405, ISSN 0733-8724

Pedro, J.; Castro, J.; Monteiro, P. & Pires, J. (2006a). On the Modelling and Performance Evaluation of Optical Burst-Switched Networks, *Proceedings of IEEE CAMAD 2006 11th International Workshop on Computer-Aided Modeling, Analysis and Design of Communication Links and Networks*, pp. 30-37, ISBN 0-7803-9536-0, Trento, Italy, June 8-9, 2006

Pedro, J.; Monteiro, P. & Pires, J. (2006b). Wavelength Contention Minimization Strategies for Optical-Burst Switched Networks, *Proceedings of IEEE GLOBECOM 2006 49th Global Telecommunications Conference*, paper OPNp1-5, ISBN 1-4244-0356-1, San Francisco, USA, November 27-December 1, 2006

Pedro, J.; Monteiro, P. & Pires, J. (2009a). On the Benefits of Selectively Delaying Bursts at the Ingress Edge Nodes of an OBS Network, *Proceedings of IFIP ONDM 2009 13th Conference on Optical Network Design and Modelling*, ISBN 978-1-4244-4187-7, Braunschweig, Germany, February 18-20, 2009

Pedro, J.; Monteiro, P. & Pires, J. (2009b). Contention Minimization in Optical Burst-Switched Networks Combining Traffic Engineering in the Wavelength Domain and Delayed Ingress Burst Scheduling. *IET Communications*, Vol. 3, No. 3, (March 2009), pp. 372-380, ISSN 1751-8628

Pedro, J.; Monteiro, P. & Pires, J. (2009c). Traffic Engineering in the Wavelength Domain for Optical Burst-Switched Networks. *IEEE/OSA Journal of Lightwave Technology*, Vol. 27, No. 15, (August 2009), pp. 3075-3091, ISSN 0733-8724

Poustie, A. (2005). Semiconductor Devices for All-Optical Signal Processing, *Proceedings of ECOC 2005 31st European Conference on Optical Communication*, Vol. 3, pp. 475-478, ISBN 0-86341-543-1, Glasgow, Scotland, September 25-29, 2005

Qiao, C. & Yoo, M. (1999). Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet. *Journal of High Speed Networks*, Vol. 8, No. 1, (January 1999), pp. 69-84, ISSN 0926-6801

Sahara, A.; Shimano, K.; Noguchi, K.; Koga, M. & Takigawa, Y. (2003). Demonstration of Optical Burst Data Switching using Photonic MPLS Routers operated by GMPLS Signalling, *Proceedings of OFC 2003 Optical Fiber Communications Conference*, Vol. 1, pp. 220-222, ISBN 1-55752-746-6, Atlanta, USA, March 23-28, 2003

Sun, Y.; Hashiguchi, T. ; Minh, V.; Wang, X.; Morikawa, H. & Aoyama, T. (2005). Design and Implementation of an Optical Burst-Switched Network Testbed. *IEEE*

Communications Magazine, Vol. 43, No. 11, (November 2005), pp. s48-s55, ISSN 0163-6804

Teng, J. & Rouskas, G. (2005). Wavelength Selection in OBS Networks using Traffic Engineering and Priority-Based Concepts. *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 8, (August 2005), pp. 1658-1669, ISSN 0733-8716

Tucker, R. (2006). The Role of Optics and Electronics in High-Capacity Routers. *IEEE/OSA Journal of Lightwave Technology*, Vol. 24, No. 12, (December 2006), pp. 4655-4673, ISSN 0733-8724

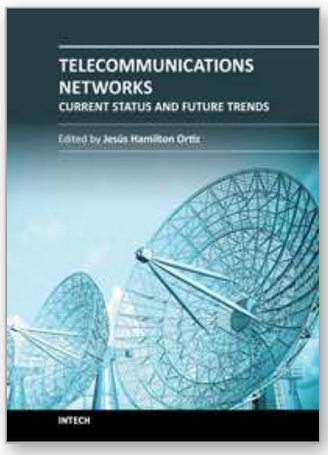
Wang, X.; Morikawa, H. & Aoyama, T. (2003). Priority-Based Wavelength Assignment Algorithm for Burst Switched WDM Optical Networks. *IEICE Transactions on Communications*, Vol. E86-B, No. 5, (2003), pp. 1508-1514, ISSN 1745-1345

Xiong, Y.; Vandenhoute, M. & Cankaya, H. (2000). Control Architecture in Optical Burst-Switched WDM Networks. *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 10, (October 2000), pp. 1838-1851, ISSN 0733-8716

Zang, H.; Jue, J. ; Sahasrabuddhe, L.; Ramamurthy, R. & Mukherjee, B. (2001). Dynamic Lightpath Establishment in Wavelength-Routed WDM Networks. *IEEE Communications Magazine*, Vol. 39, No. 9, (September 2001), pp. 100-108, ISSN 0163-6804

Zhou, P. & Yang, O. (2003). How Practical is Optical Packet Switching in Core Networks?, *Proceedings of IEEE GLOBECOM 2003 49th Global Telecommunications Conference*, pp. 2709-2713, ISBN 0-7803-7974-8, San Francisco, USA, December 1-5, 2003

Zhu, K.; Zhu, H. & Mukherjee, B. (2005). *Traffic Grooming in Optical WDM Mesh Networks,* Springer, ISBN 978-0-387-25432-6, New York, USA



Telecommunications Networks - Current Status and Future Trends

Edited by Dr. Jesús Ortiz

ISBN 978-953-51-0341-7
Hard cover, 446 pages
Publisher InTech
Published online 30, March, 2012
Published in print edition March, 2012

This book guides readers through the basics of rapidly emerging networks to more advanced concepts and future expectations of Telecommunications Networks. It identifies and examines the most pressing research issues in Telecommunications and it contains chapters written by leading researchers, academics and industry professionals. Telecommunications Networks - Current Status and Future Trends covers surveys of recent publications that investigate key areas of interest such as: IMS, eTOM, 3G/4G, optimization problems, modeling, simulation, quality of service, etc. This book, that is suitable for both PhD and master students, is organized into six sections: New Generation Networks, Quality of Services, Sensor Networks, Telecommunications, Traffic Engineering and Routing.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen