Virtual Optical Bus: An Efficient Architecture for Packet-Based Optical Transport Networks

Ahmad Rostami and Adam Wolisz

***Abstract*—The virtual optical bus (VOB) is presented as a novel architecture for packet-based optical transport networks. The VOB is an evolutionary networking architecture based on the optical burst/packet switching (OBS/OPS) paradigm with a higher performance—in terms of packet loss rate and network throughput. The achieved gain comes at a cost of a marginal increase in the delay that packets experience at the ingress edge of the network, where we can still use inexpensive electrical buffers. In the VOB architecture, flows of traffic between nodes in the network are grouped into clusters and within each of the clusters a special form of coordination on packet transmission is introduced. This coordination ensures collision-free packet transmission within each cluster. Additionally, clustering of flows and selection of paths for clusters are done in a way that the interaction among routes of clusters in the network is minimized. This leads to a reduction of packet collisions in the network and also an increase in the network throughput. Design issues related to the VOB architecture are discussed and two design examples are presented that illustrate the high potential of this approach.**

***Index Terms*—Optical burst/packet switching; Contention avoidance; Traffic engineering; Traffic shaping**.

I. Introduction

The exponential growth of traffic in the Internet has turned optics into the technology of choice for transmission in the metro/core networks. In addition, it has created some concerns over the scalability of electronic switching, thereby pushing towards all-optical networks. Nonetheless, optics does not provide good support for packet switching. Specifically, there are three major issues hindering realization of an all-optical packet switching approach in the near future: low speed of all-optical switches, immature all-optical processors, and lack of true optical buffers. One promising approach to addressing these issues is optical burst switching (OBS); see [1]. OBS consists of grouping packets into bursts, out-of-band signaling, and cut-through switching that collectively eliminate the need for fast optical switches and all-optical processors. It, however, fails to fully eliminate the problem of buffering in the network. In fact, buffers are needed in packet switched networks for two purposes. First, they are used to keep the packets in nodes while the node controllers read and process the packets' headers. Additionally, buffers are the main tools to mitigate the contention problem over the output ports of packet switches. Although OBS eliminates the need for buffers with regard to their first role, it does not provide an efficient solution for contention resolution.

In our earlier work in [2], we demonstrated that in a transport network based on the OBS, traffic shaping at the entrance edge of the network offers high potential for contention mitigation. More specifically, the losses due to lack of buffers in transit switches are reduced by buffering—shaping—of data packets at the ingress edge of the network, where we can still use inexpensive electrical buffers. In this work we aim at even more efficient contention reduction by suggesting a novel architecture and protocol for the optical transport network. In our approach, several ingress nodes inside the network are grouped together into a cluster. The set of links connecting the nodes in a cluster form an optical path that is termed a *virtual optical bus* (VOB). A path formed in this way is called a virtual bus since the physical links to which a path is associated can be used by several paths simultaneously. A form of coordination among nodes belonging to the same VOB on injecting their packets into the network is applied to eliminate the potential of collisions among packets within the same VOB. In a VOB network—a network that applies the VOB approach—all flows are grouped into VOBs in such a way that each flow belongs to only one VOB and the flows belonging to different VOBs have the minimum potential for contention with each other inside the network.

The rest of this paper is structured as follows. In Section II, we present the architecture of the system under consideration and discuss the outline of our approach. In Section III we specify the VOB and present the design of the VOB MAC protocol. In Section IV we discuss the design of VOB networks and present an ILP formulation for that. Section V investigates the performance of the VOB MAC protocol and its comparison with OBS. In Section VI two VOB network design examples are presented that demonstrate the operation and performance of the VOB network. Related work is presented in Section VII, and finally conclusions are given in Section VIII.

Manuscript received May 21, 2010; revised August 24, 2010; accepted August 24, 2010; published October 19, 2010 (Doc. ID 128821).

The authors are with the Telecommunication Networks Group (TKN), Technical University of Berlin, 10587 Berlin, Germany (e-mail: rostami@tkn.tu-berlin.de).

Digital Object Identifier 10.1364/JOCN.2.000901

1943-0620/10/110901-14/$15.00

II. System Architecture and Outline of Our Solution

Consider an OBS transport network, e.g., [3], having N nodes and L unidirectional WDM links, where links connect nodes following an arbitrary topology. Any given link l can support W_l wavelength channels, each operating at a rate of C Gb/s.

The network operates in the asynchronous burst switching mode, where bursts (i.e., jumbo data packets destined to the same target node) are assembled at the edge node and injected into the network. The information required by intermediate nodes to forward a data packet towards its destination is carried in a separate header packet, which is released to the network simultaneously with the data packet but on a separate signaling channel. At each intermediate node, data bursts go through an optical delay line while the header packet is converted to the electrical domain, processed by the node controller, and converted again back to optics (O/E/O conversion); see Fig. 1.

Each node n in the network is equipped with T_n tunable optical transmitters and R_n tunable optical receivers. Each tunable transmitter and receiver could be tuned to, respectively, send or receive data on any of the wavelength channels of the links attached to the node. Furthermore, a reactive contention resolution mechanism is implemented in each node n, which incorporates a set of G_n internal wavelength converters as well as F_n WDM fiber delay lines (FDLs) per output link of that node. Note that, if a node operates only as a transit node, it does not need to contain any add/drop facilities. Similarly, if a node is only an edge node it does not need any wavelength converter or FDL buffer for contention resolution. Figure 1 shows an abstract model of the node that can operate alternatively as an ingress/egress edge node and as a transit node.

Let us assume temporarily that the ingress nodes work independently—with no coordination among each other—for injecting traffic into the network, which is natural for the packet switching paradigm. This carries a natural potential of collision among packets in the network. A collision occurs when two or more data packets are going to use the same channel on a given link for an overlapping period of time.

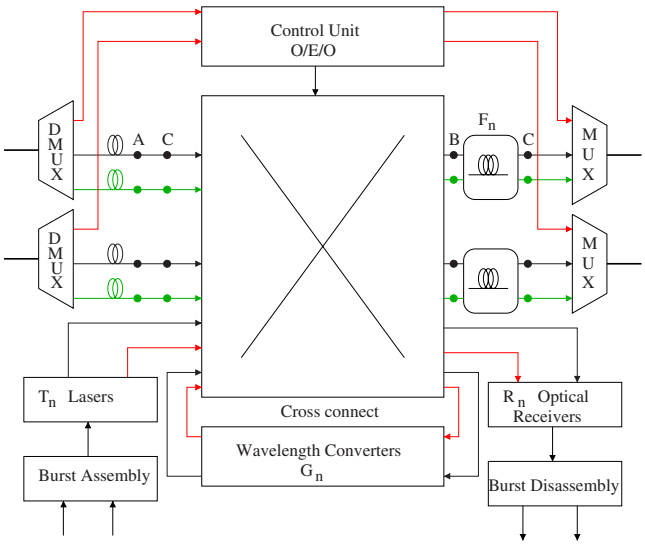

Fig. 1. (Color online) A generic architecture of the switch assumed in the VOB.

In the electronic packet-switched networks, efficient avoidance of this adverse phenomenon is usually done by introducing some buffers that hold colliding packets during times of contentions. Unfortunately, the all-optical nodes have limited buffering potential—through FDL buffers—which due to lack of the random access property can offer only predefined fixed delay values. Accordingly, the only way to limit the contention is an appropriate traffic shaping at the input edge of the network. An optimal idealistic approach for this purpose would be global scheduling of all sources so that every single packet transmission from any source in the network is coordinated with all other sources so as to avoid collision in the transit nodes. This could—theoretically—be realized by introducing a central controller to the network that schedules all packet transmissions in the network in a collision-free manner. In this way, up to some maximum utilization the controller can ensure that the packet will not collide with any other packet along its path to destination. Although this approach would potentially lead to the highest throughput in the network and the minimum amount of collisions, this might be too costly and not practical due to the necessary ideal time synchronization of all components, complexity, as well as signaling and scheduling overhead. A natural alternative is decomposition of the global scheduling of all sources into smaller sets with local scheduling within each set and hoping that the results of the local scheduling will not collide too much with each other. On account of this, we postulate grouping of traffic sources into disjoint clusters and apply a form of coordination—traffic shaping—among associated sources in each cluster. In this way only the groups of sources remain independent and subject to collisions, thereby reducing the collision rate in the network.

A crucial issue for clustering is setting the requirements for assigning sources to clusters. Taking into account that the final goal is to decrease the packet collisions in the network, a natural requirement is to form a cluster out of the sources that potentially share the same set of links in their path and thus their packets are likely to collide with each other. Furthermore, we have to take care that the sets of routes of flows from different clusters should intersect minimally—ideally be disjoint. In the latter case there would be—by definition—no collisions between the sources belonging to different clusters either! This might, however, be impossible in reality; thus we suggest keeping the inter-cluster collisions on a low level by finding for them paths with minimal interaction.

Concerning the scheduling within the clusters, we need a local scheduling mechanism that is feasible and at the same time efficient. In fact, as opposed to the global scheduling, the local scheduling pertains only to a limited number of flows in the network and has therefore much lower complexity. Nevertheless, we will go a step further and target a distributed local scheduling mechanism rather than a centralized approach. As for the efficiency, the same criteria are applied to the local scheduling as the global scheduling. That is, a local scheduling algorithm operating on a local set

of sources should schedule the corresponding packets with minimum—ideally no—collisions within that local set.

To evaluate the effectiveness of our approach in improving the performance of the network we consider three metrics: packet loss rate, access delay of packets, and network throughput. Packet loss rate will be calculated as the fraction of generated packets by the ingress nodes that are lost in the network due to collisions. The access delay is calculated as the difference between the time a packet is generated by an ingress node and the time it is released to the network, and it is calculated for successfully delivered packets. The maximum achievable throughput of the network will also be evaluated and compared with that of a classical OBS network.

III. Virtual Optical Bus

Our approach for coordination of flows within each cluster will be based on the abstraction of the VOB—a virtual unidirectional bus accommodating a set of origin–destination (O-D) flows, defined as a flow of packets originating at an ingress edge node destined to a specific egress node in the considered network, on their entire route.

The following features collectively give the formal definition of the VOB.

- A VOB is defined over a specified sequence of connected links, i.e., it is a directed simple path with a single origin and destination.
- Each VOB is piecewise associated with a specific wavelength channel. On a multichannel link with wavelength converters available, a VOB can use any channel on that link; however, flows associated with that VOB may not use more than one channel at any given time.
- Each O-D flow must be associated with one and only one VOB.

Note that the third feature implies that both the origin and destination of a given flow should be associated with the same VOB, meaning that the traffic of an O-D flow will not be split/switched over multiple VOBs.

Following the description given above, establishing the coordination among flows associated with a VOB is now translated into designing a medium access control (MAC) protocol that ensures collision-free, efficient, and fair access to the VOB for the associated traffic flows. For this purpose we use a MAC protocol that is based on the buffer-insertion protocol, which was first introduced in the mid-1980s and was afterwards improved and used in many other works; see [4]. In the buffer insertion, each intermediate active node on a bus—an intermediate node that injects traffic into the bus—is equipped with an extra buffer called an insertion buffer. The buffer is used to actively delay transit packets flowing on the bus in order to prevent them from colliding with local packets in transmission. Also, at every intermediate node, priority is given to the transit traffic over the local traffic. In this way, the required capacity of the insertion buffer would be small. Specifically, it must hold a transit packet until the transmission of an ongoing local packet is completed; therefore, it only needs capacity equal to the transmission time of a maximum-length packet.

A known phenomenon associated with the buffer insertion protocol is the fairness issue. In fact, in a physical bus using buffer-insertion MAC protocol, those nodes closer to the beginning point of the bus have better and faster access to the bus. This might result in unfair access to the bus for the nodes closer to the end of the bus. Nevertheless, this is not the case for the VOB. In fact, as we discuss in Section IV, in designing a VOB network, O-D flows are associated with a VOB based on their average data rate; that is, it is assumed that a single upstream node does not have the potential to monopolize the VOB in the long term. Nevertheless, in order to guarantee that downstream O-D flows do not have to wait excessively long times before they are granted access to the VOB, in our design each source node in the network is further equipped with a token bucket shaper.

In order to adopt the buffer-insertion protocol in the VOB, we need an insertion buffer per VOB per any active intermediate node being part of that VOB. That is, in a network with N nodes each node would require at most $N-1$ insertion buffers since it generates traffic destined to at most $N-1$ other nodes in the network. Since the insertion buffer is used to buffer transit optical packets already traveling over the bus, it must be realized in the optical domain to avoid the need for the electro-optical conversion. Taking into account that the required capacity of an insertion buffer is small, it can be easily realized by a single FDL as shown in Fig. 2.

The operation of the VOB MAC protocol is as follows:

- At any node i being part of the VOB there is a token bucket shaper that is filled with new tokens at rate t_i (bytes/s). The bucket has the maximum size of $T_{max,i}$ (bytes) and T_i (bytes) denotes the available tokens in the bucket at any given time.
- Any active node i on the VOB can inject a packet, which is at the head of its local transmit queue, into the VOB as soon as the following three conditions are met: i) the corresponding transmission channel is idle, ii) the packet size is smaller than the available token T_i in the corresponding bucket, and iii) the insertion FDL is idle or the gap between the time that the packet inside the FDL starts leaving the buffer and the current time is larger than the time required to transmit the local packet. If the active node is the most upstream node of the VOB, it does not have an insertion buffer

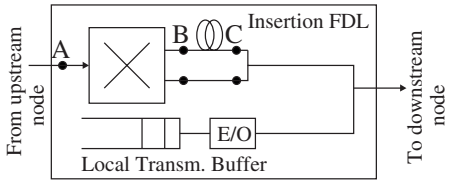

Fig. 2. Abstract model of a single ingress node equipped with a FDL insertion buffer on the transit path. There are two possibilities for incorporating the insertion buffers into the architecture of the switch depicted in Fig. 1: at the input or output of the switch (see points marked as A, B, and C). In the former case, the insertion buffers are placed between the delay lines and the switch fabric, which necessitates additional 1×2 switching elements before the insertion buffers to allow for bypassing the insertion buffers if required. In the latter case, the insertion buffers could be incorporated into the FDL buffering unit. In either case, additional ports on the main switch would be needed.

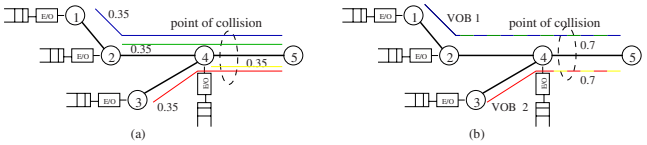

Fig. 3. (Color online) Example of applying the VOB: (a) a simple network with four O-D flows, (b) the same network after applying the VOB.

and therefore it does not need to consider condition iii. In any case, immediately after the packet transmission is scheduled, T_i is reduced by the packet size.

- If a transit packet arrives at an intermediate node of the VOB while a local packet is being transmitted, it will be delayed by the insertion FDL.

This operation guarantees a collision-free packet transmission within a VOB. In Section V, we investigate the operation and performance of the proposed MAC protocol under different settings in detail.

To elaborate more about the VOB operation let us now present an example. Figure 3 depicts a network with 5 nodes, in which nodes 1–4 generate traffic destined to node 5, each with an intensity equal to 35% of the capacity of a single WDM channel. Packets generated at each node go through an electro-optical (E/O) conversion before transmission. The link $(4\rightarrow5)$ is offered the aggregate load of 1.4 through multiplexing four independent O-D flows. If we suppose that this link has two data channels and two tunable wavelength converters with no FDL buffer and that the traffic of each source is injected according to the Poisson process, by applying the known Erlang-B formulae [5] we will observe some 29% packet loss rate on this link under OBS operation. Now, let us assume we establish two VOBs over the physical routes $(1\rightarrow2\rightarrow4\rightarrow5)$ and $(3\rightarrow4\rightarrow5)$, respectively [Fig. 3(b)]. The former VOB accommodates O-D flows $(1\rightarrow5$ and $2\rightarrow5)$ and the latter carries the traffic of O-D flows $(3\rightarrow5$ and $4\rightarrow5)$. In this case, by using the MAC protocol described above, node 2 (4) will schedule its local packets in a way that they do not overlap with packets of node 1 (3), i.e., no intra-VOB collision occurs. Consequently, we have only two independent traffic flows being multiplexed into link $(4\rightarrow5)$, and therefore no packet collision will occur on this link anymore since it already has two channels.

IV. VOB Network Design

Design of a VOB network consists of grouping all O-D flows into VOBs and that has to be done with the objective of minimizing the packet collision rate in the network. We refer to this step as VOB layout design, which can formally be stated as follows. Having a network with given resources, as explained in Section II, and an O-D traffic matrix, how to cover the traffic matrix by a set of VOBs, with respect to the limitations imposed by the topology and the resources available within each node and link, in such a way that the packet collision in the network is minimized. The VOB layout design should give us complete information about the beginning, the end, and the route of all VOBs as well as the IDs of flows to be covered by each VOB. It should be noted that VOB layout design involves assigning O-D flows to VOBs and that, in turn, implies routing of O-D flows in the network. In general, there are two possibilities for routing bidirectional flows between two specific nodes in the network: symmetric or asymmetric routing. In our design, we do not enforce symmetric routing and this can cause some bidirectional flows to take different routes on opposite directions. Nevertheless, for some scenarios such as the ring network topology the symmetry between the routes are to a high extent achieved.

Taking into account that the intracluster collisions—collisions among the flows belonging to the same VOB—are completely suppressed by the proposed MAC protocol, the objective function of the VOB layout design translates into minimizing the intercluster collisions across the network. Nevertheless, minimization of the intercluster collisions cannot be solved directly and in a straightforward manner, if at all possible, since it depends on many factors such as traffic characteristics and routing. Consequently, we resort to an alternative approach, where the objective function is set to minimization of the maximum number of VOBs that are multiplexed to any link in the VOB network. This objective function will minimize the interaction among routes of VOBs in the network, thereby reducing the intercluster collision across the network.

In the following, we apply the path-based method [6] to formulate the VOB layout design as an integer linear programming (ILP) problem. In order to facilitate finding a solution to the problem, we try to limit the optimization search space. This is achieved through generating a list of many potential VOBs in the network and also for each potential VOB determining all O-D flows that it could support. This process is carried out in a preprocessing phase as explained below.

A. Preprocessing

Let us assume that all links and O-D flows in the network are, respectively, populated in sets $\mathbb{L}$ with L elements and $\mathbb{F}$ with F elements. We first apply the k-shortest path (KSP) routing algorithm [7] to the given topology and populate the set of all k-shortest paths between all pairs of nodes in the network in $\mathbb{P}$ with size P. Note that $\mathbb{P}$ is the set of VOB candidates in the network.

Now we form the matrix Θ $(P\times L)$ that relates links in the network to the paths populated in $\mathbb{P}$. That is,

$$\theta(p,l) := \begin{cases} 1 & \text{if link } l \in \mathbb{P} \text{ is used in path } p \in \mathbb{P} \\ 0 & \text{otherwise} \end{cases}.$$

In the next step, we form Γ, which is a $P\times L\times F$ matrix and whose elements are calculated as

$$\gamma(p,l,f) := \begin{cases} 1 & \text{if O-D flow } f \in \mathbb{F} \text{ could be supported by path } p \in \mathbb{P} \text{ on link } l \in \mathbb{L} \\ 0 & \text{otherwise} \end{cases}.$$

In case there are any constraints regarding the routing of some O-D flows in the network they have to be applied during the formation of Γ. One such constraint can be the maximum allowable hop count for routing of O-D flows. We finally note that for a given network topology the preprocessing phase needs to be done only once.

B. ILP Formulation

a) Assumptions: For our formulation we make the following assumptions:

- Between each two nodes in the network there will be at most one VOB in each direction.
- All nodes support full wavelength conversion. Although this is not mandatory for the VOB design framework, this assumption will simplify the layout design process.

b) Parameters and Variables: The parameters used in the formulation are the following:

- $\Lambda=[\lambda^f]$ is the given demand matrix, where λ^f denotes the intensity of the O-D flow $f \in \mathbb{F}$ (normalized to a single wavelength capacity).
- A_{max} is the maximum allowable load on a single WDM channel in the network normalized to the capacity of a single wavelength.

Also, the variables are

- x_p

$$x_p := \begin{cases} 1 & \text{if } p \in \mathbb{P} \text{ is selected as a VOB} \\ 0 & \text{otherwise} \end{cases}.$$

- y_p^f

$$y_p^f := \begin{cases} 1 & \text{if flow } f \in \mathbb{F} \text{ is supported by path p} \in \mathbb{P} \\ 0 & \text{otherwise} \end{cases}.$$

c) Objective Function: We set the ILP objective function to

$$\text{Min} \max_{l \in \mathbb{L}} \sum_{p \in \mathbb{P}} \theta(p,l) x_p. \tag{1}$$

The objective function minimizes the maximum number of multiplexing VOBs over any link in the network.

d) Optimization Constraints:

$$\sum_{p \in \mathbb{P}} y_p^f = 1 \quad \forall f \in \mathbb{F}. \tag{2}$$

$$\sum_{f \in \mathbb{F}} y_p^f \lambda^f \gamma(p,l,f) \leq x_p A_{max} \quad \forall l \in \mathbb{L}, \quad p \in \mathbb{P}. \tag{3}$$

In constraint 2 each traffic flow is associated with one and only one VOB. Also, constraint 3 ensures that traffic offered to any part of a VOB on any link in the network does not exceed a fraction of the WDM channel capacity equal to A_{max}, which is a design parameter used to control the access delay of traffic sources.

V. Evaluation of VOB MAC Protocol: An Example

In this section, we evaluate the VOB MAC protocol and compare it with that of classical OBS. The VOB MAC protocol presented in Section III is a new variation of the well-studied buffer-insertion protocol. In the buffer insertion, where a physical bus is shared among a set of nodes, 100% throughput can be achieved as long as the total load offered to each segment of the shared bus is smaller than the transmission capacity of that segment. The only concern would be delay of accessing the bus, which rapidly increases with the distance of a node from the head end of the shared bus. This is the so-called fairness issue associated with the buffer-insertion protocol. The issue stems basically from the inherent priority of upstream sources in accessing the bus that allows them to grab as much capacity as they want without considering the needs of downstream nodes. This can effectively be controlled by limiting the injection rate of each source to the bus, e.g., by means of a token bucket controller. Nevertheless, in a VOB with a load-controlled insertion-buffer protocol, 100% throughput cannot be achieved, since a FDL-based insertion buffer cannot utilize the whole capacity of an optical link. That the type of buffer used in realizing the FDL-based insertion buffer does not support the random access property leads to the fragmentation of the data channel, and this consequently hinders the optimum utilization of channel capacity. Therefore, it is of great importance to investigate the delay-throughput characteristics of the MAC protocol proposed for the VOB.

Let us consider the delay throughput of a single VOB on a link in the network. The performance of the VOB MAC protocol on this link depends on several factors, including the number of VOBs multiplexed onto the link and their associated traffic intensity, the number of wavelength channels on the link, and the number of upstream nodes. Therefore, we need a sophisticated scenario that allows us to consider all these parameters. For this purpose we have designed a test scenario as depicted in Fig. 4.

The network in Fig. 4 is a single-bottleneck topology that consists of B parallel branches of nodes, $(n+1)\times B+1$ nodes, and $(n+1)\times B$ unidirectional links, where a link connecting any two nodes is realized by a WDM link with W_l data channels and one control channel. We further make the following assumptions:

- Nodes $S_{i,1},\ldots,S_{i,n+1}$ generate traffic destined for node D_i $(1 \leq i \leq B)$.
- Aggregate load generated for node D_i $(1 \leq i \leq B)$ is less than the transmission capacity of a single wavelength channel.
- The network operates under either of the two modes: classical OBS and VOB. In the latter case, one VOB is established on any branch i $(1 \leq i \leq B)$ and all nodes on the branch are associated with that VOB; see Fig. 4. Accordingly, in this case, each active node on the VOB is assumed to be equipped with the VOB MAC protocol

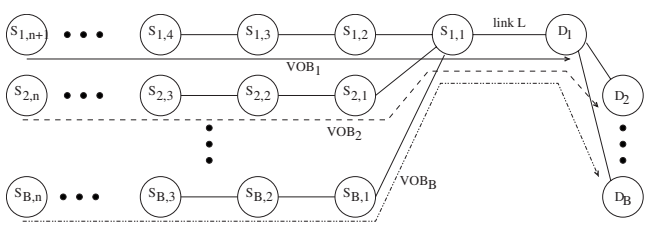

Fig. 4. Single-bottleneck scenario used for MAC protocol evaluation.

and corresponding FDL-based insertion buffer. In addition, it is assumed that the local transmission buffer at each node has an unlimited buffering capacity.

- The number of wavelength channels on any link at the left side of the bottleneck link L is equal to n, and that of the bottleneck link L is equal to W_L. Each link on the right side of link L has one wavelength channel.
- Each source node is equipped with one tunable optical transmitter and n tunable wavelength converters, and it does not have any FDL buffer for contention resolution.
- At each source node, bursts of fixed length are generated according to the Poisson process with intensity ρ (the same for all sources), as normalized to the wavelength channel capacity. The traffic model is in line with the analysis presented in [8].

To quantify the performance metrics, we developed a simulation platform of the network using the discrete event simulator OMNeT++ [9]. Simulation experiments are carried out for both OBS and VOB network operation at a variety of combinations of $B=\{2,3,4\}$, $n=\{2,3,4,6\}$, $W=\{1,2,3,4\}$, and at different load values. In the experiments, wavelength channel capacity and burst length are set to 10 Gb/s and 10 kB,[1] respectively. Also, the token arrival rate parameter of the token bucket shaper in each source node is set to be equal to 1.1 times the average load value of that source, and the bucket size is set to allow transmission of up to 20 consecutive packets for each source.

We evaluate the packet loss rate, throughput, and access delay—as defined at the end of Section II—of the designed scenario. The loss rate is estimated at the entrance of the bottleneck link L as this is the only point in the network where packets might get dropped due to contentions. The throughput is considered at the network level and the delay is estimated as experienced by node $S_{1,1}$ in accessing VOB_1.

[1]That is, the burst transmission time is equal to 8 μs. A network with this setting would need fast optical cross-connects as well as tunable lasers, receivers, and converters with the switching/tuning times in the range of μs or less. There are hints that such equipment is either partly available or would be realizable in the near future (see [10]).

In the following we discuss the achieved results in detail; however, for the sake of brevity the detailed graphs are shown only for the case of $B=3$.[2] In the analysis given below, we distinguish between two modes of operation for VOB: mode I corresponds to the case where the number of VOBs B associated with the bottleneck link L is smaller than or equal to W_L (i.e., $W_L \geq B$), and mode II refers to the other cases, i.e., $W_L < B$.

A. Loss Rate and Throughput

Figures 5 and 6 depict the loss rate on the bottleneck link and the network throughput, respectively, as a function of the load offered to the bottleneck link. First and foremost it is observed that applying the VOB greatly reduces the loss rates and as a result, the achieved throughput of the VOB network is higher than that of the OBS network. In fact, in VOB mode I the loss rate is always zero, and therefore the throughput is equal to the offered load. This holds true for any number of nodes n and any number of VOBs B. Therefore, in this mode, we observe the largest difference between the VOB and OBS networks in terms of the loss rate and the throughput. In mode II burst losses occur under both VOB and OBS networks, and the loss rate increases with the offered load. Nevertheless, the loss rate of the VOB network is still much smaller than that of the OBS network. In addition, in this mode the difference between the two approaches in terms of the throughput increases with load. Table I summarizes the throughput gain of the VOB network over the OBS network under different settings of the parameters and at 70% bottleneck-link load. It is observed that the throughput gain of the VOB for the considered scenario can be as high as 22.7%.

Now let us consider the impact of the number of nodes n. Although in VOB mode I both the throughput and loss rates are independent of the number of nodes n, this is not the case for OBS, in which both the loss rate and the throughput deteriorate with n. In other words, in this mode the gain of the VOB network over the OBS network—in terms of throughput and loss rate—increases with n. In VOB mode

[2]More detailed results of the experiments are presented in [11].

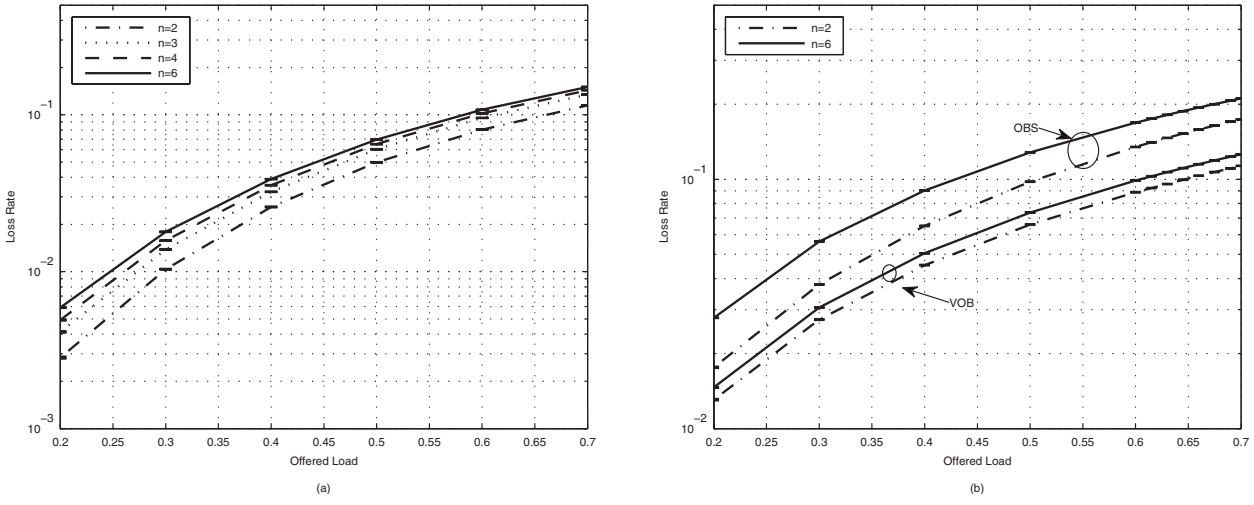

Fig. 5. Loss rate on the bottleneck link against the offered load at $B=3$, (a) OBS, $W=3$. The loss rate for the VOB network is zero, (b) $W=2$. The error bars represent 90% confidence intervals.

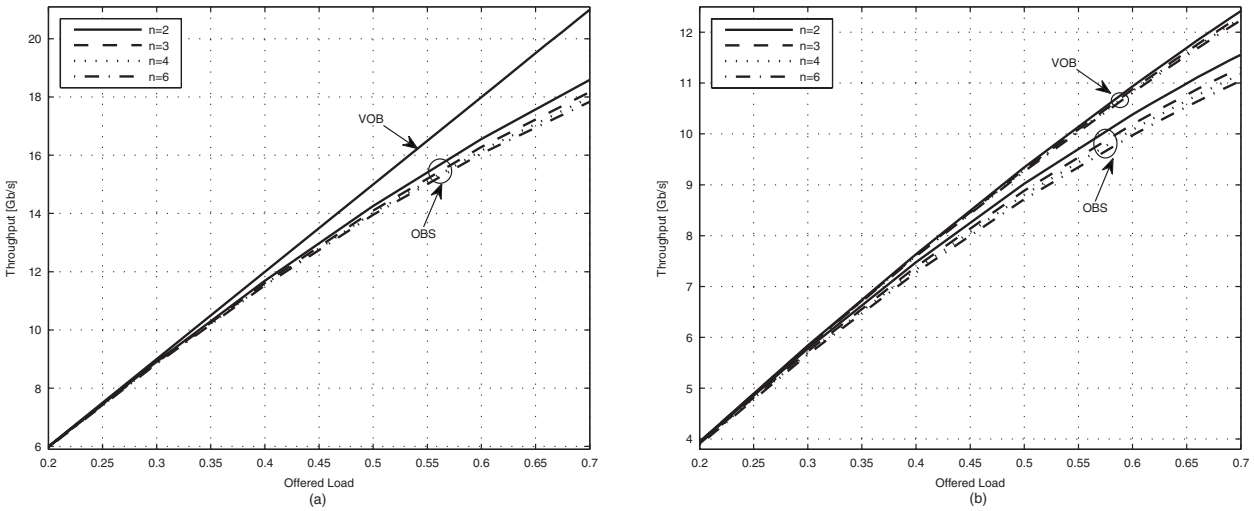

Fig. 6. Throughput of the network against the offered load to the bottleneck link at $B=3$, (a) $W=3$, (b) $W=2$. 90% confidence intervals are not shown since they are too small.

II, similar to OBS, the throughput of the VOB network decreases with n for a given setting. Nevertheless, the reduction of throughput for the VOB network is much smaller than that of OBS, which helps the VOB network to outperform OBS more largely with n.

Finally, we analyze the impact of the number of VOBs. From Table I it is seen that the throughput gain of the VOB network over the OBS network in mode I decreases with B. For instance, at $n=6$ the throughput gain reduces from 22.7% at $B=2$ to 10.9% at $B=4$. To clarify the reason, we note that in this mode the number of VOBs B is equal to the number of wavelength channels on the bottleneck link. That is, at $B=4$ the number of wavelength channels on the bottleneck link is twice larger than that at $B=2$, whereas the normalized offered load is fixed. On the other hand, at a fixed load the drop rate should decrease with the number of available wavelength channels. That is, the drop rate of the OBS network decreases with B, while the drop rate of the VOB network is fixed and equal to zero in this mode. As a result, the difference between the VOB and OBS networks reduces with B. In VOB mode II the throughput gain depends very much on the ratio B/W_L as depicted in Table I. For instance, the throughput gains are similar for cases $B=4/W_L=2$ and $B=2/W_L=1$. Additionally, the throughput gain of the VOB network improves as B/W_L approaches 1.

B. Access Delay

In the previous subsection it was demonstrated that the VOB network achieves a better performance—in terms of loss rate and throughput—in comparison with the classical OBS. This gain, however, comes at the cost of an increase in the access delay, i.e., the time it takes for a source node to inject its traffic into the network. In this subsection, we investigate the access delay characteristics of the VOB.

To characterize the access delay of the VOB at a given node it is important to take into account the impact of the number of upstream nodes, which actively inject traffic into the VOB. In our analysis the access delay is estimated at node $S_{1,1}$ while accessing VOB_1 as a function of total load offered to VOB_1 on the bottleneck link. To account for different numbers of upstream nodes, we vary n such that at a given value of $n=N$ there would be N nodes upstream to node $S_{1,1}$.

Figure 7 depicts the average access delay of node $S_{1,1}$ as a function of total load offered to VOB_1 on the bottleneck link. The delay values are shown at $B=3$ and $W_L=3$ for different numbers of upstream nodes. The access delays for the VOB in mode II (i.e., $W_L<3$) as well as for OBS operation are smaller than 10 μs for all settings of the considered scenario and therefore are not shown here.

It is seen that the access delay for node $S_{1,1}$ is smaller than 150 μs at 70%–75% VOB load even when there are six upstream nodes. It should be noted that the access delay in the range of a few hundred μs is considered acceptable taking into account the propagation delay in a metro/core network that is usually in the range of tens of ms.

More specifically, we observe that the access delay in-

TABLE I
THROUGHPUT GAIN OF THE VOB NETWORK OVER THE OBS NETWORK AT 70% LINK LOAD IN PERCENT

	$B=2$		$B=3$		$B=4$		
	$W_L=2$	$W_L=1$	$W_L=3$	$W_L=2$	$W_L=4$	$W_L=3$	$W_L=2$
$n=2$	16.5	2.9	13.0	7.4	7.0	8.0	4.1
$n=3$	19.7	5.3	15.6	9.1	9.2	9.4	5.4
$n=4$	22.2	5.7	16.7	10.0	10.1	10.0	6.1
$n=6$	22.7	8.8	17.8	10.4	10.9	10.6	6.9

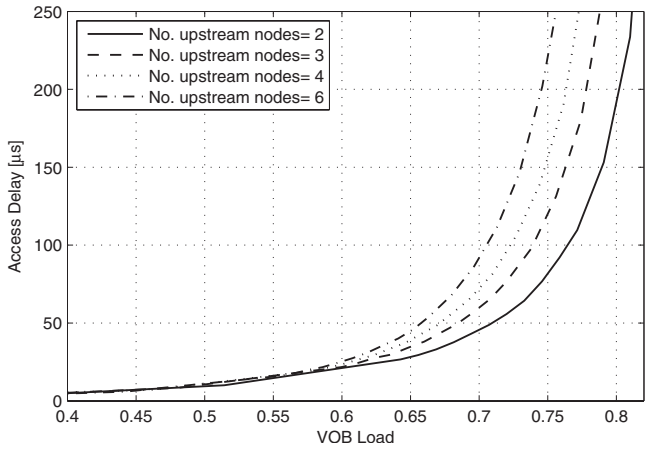

Fig. 7. Access delay experienced by node $S_{1,1}$ in accessing VOB 1 at $B=3$ and $W=3$. 90% confidence intervals are not shown since they are too small.

creases smoothly with the VOB load up to a knee point, which occurs at around 60%–65% load. At the load values beyond this point the access delays increase sharply. Although the number of upstream nodes has a large impact on the access delay after the knee point, this impact is not large when the VOB load is smaller than the load value of the knee point. Our further analysis—not depicted here—showed that the number of VOBs B does not affect the access delay to a large extent, which is not surprising since different VOBs are to a high extent isolated from each other as long as $W_L \geq B$.

In VOB mode II the access delay is smaller than 10 μs even at 70%–75% VOB load and with six upstream nodes. The reason is that for a given number of VOBs to keep the load on the link fixed while decreasing W_L the load per VOB has to be reduced accordingly. The reduction in the VOB load, in turn, pushes the access delay curve under the knee point, where the access delay value is very small.

Overall, it is seen that the VOB shaping delay for all the considered scenarios is acceptable as long as the load offered to the VOB is smaller than a given threshold, which is in the range of 70%–75% of the single channel capacity. This requirement can be enforced in the design process of the VOB network by properly setting the A_{max} value, which limits the load offered to any VOB in the network.

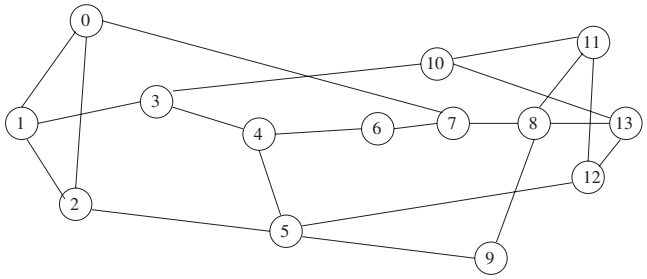

Fig. 8. NSFNET topology.

VI. VOB Network Design Examples

In this section we present two design examples of the VOB network and evaluate the performance of the designed networks. For this purpose, we consider two network topologies as explained below.

In the first example, we consider a network with 10 nodes, where nodes are connected according to a bidirectional ring topology. Each node is connected to any of the two adjacent nodes via two WDM fibers in different directions, i.e., there are 20 fiber links in the network. In the second example, the NSFNET backbone network with 14 nodes is considered, as depicted in Fig. 8. The network has 42 unidirectional fiber links and the minimum and maximum nodal degree of the nodes are 2 and 4, respectively.

In both networks it is assumed that each WDM link supports one control channel and four data channels, where each data channel operates at the rate of 10 Gb/s. The architecture of the nodes is the same as the one described in Section II. In the ring network, each node is equipped with $T_n=R_n=8$ tunable lasers and receivers as well as $G_n=8$ internal wavelength converters. For the NSFNET, the number of available tunable lasers and receivers as well as that of wavelength converters in node i $(0 \leq i \leq 13)$ are set to $T_n=R_n=G_n=4\times ND_i$, where ND_i is the nodal degree of node i.

A. Solving the ILP

1) Ring Network: For the ring network described above we generate two different traffic matrices. The first one is a random traffic matrix as depicted in Table II. Each element in the matrix of Table II represents the average traffic de-

TABLE II
Random Traffic Matrix for the Ring Network Used in the Design Example

	0	1	2	3	4	5	6	7	8	9
0	0.000	0.180	0.043	0.173	0.341	0.167	0.312	0.094	0.219	0.372
1	0.318	0.000	0.385	0.364	0.249	0.020	0.156	0.141	0.119	0.310
2	0.124	0.092	0.000	0.073	0.140	0.361	0.097	0.328	0.298	0.195
3	0.211	0.365	0.310	0.000	0.205	0.378	0.162	0.006	0.076	0.174
4	0.066	0.061	0.327	0.058	0.000	0.196	0.039	0.017	0.275	0.179
5	0.241	0.330	0.347	0.054	0.030	0.000	0.053	0.068	0.073	0.123
6	0.105	0.215	0.034	0.348	0.096	0.135	0.000	0.260	0.147	0.203
7	0.262	0.398	0.160	0.232	0.049	0.360	0.382	0.000	0.250	0.204
8	0.276	0.031	0.104	0.220	0.074	0.148	0.230	0.259	0.000	0.327
9	0.299	0.177	0.320	0.058	0.096	0.044	0.024	0.180	0.032	0.000

mand between two corresponding nodes (i.e., a single O-D flow), as normalized to the capacity of a single wavelength. To generate the traffic matrix, we used a uniform random generator and scale the generated values such that the maximum O-D demand across the network is limited to 0.4. This selection provides possibilities for grouping O-D flows into VOBs. The average demand per O-D flow for the generated traffic matrix is equal to 0.187. The second traffic matrix is a uniform one, where the demands of all O-D flows are the same and equal to 0.187. That is, the average demand per O-D flow is the same for both the traffic matrices. In the preprocessing phase both possible paths[3] for each source–destination pair are included $\mathbb{P}$. The commercial solver CPLEX 9.0 [12] is used to solve the ILP problem formulated in Section IV. Our analysis of the VOB in Section V suggests that the parameter A_{max} should be set to a value around 0.7 in order that all the traffic sources associated with a VOB experience a reasonable access delay. Accordingly, the ILP problem is solved for each of the two traffic matrices and with three values of the parameter A_{max}: 0.6, 0.7, and 0.75. Summaries of the relevant characteristics of the resulting VOB networks and a comparison with the corresponding OBS network are depicted in Figs. 9 and 10. The details of solutions are not shown here because of the space limit and interested readers are referred to [11]. Further processing of the detailed solutions—see [11]—shows that for the considered scenarios the required number of insertion buffers per node in the network, on average, varies between 4.1 and 4.8, which is much less than the theoretical upper bound given in Section III.

From Figs. 9 and 10 we observe that the number of VOBs associated with the links in the VOB network for different settings is much smaller than the number of flows per link in the OBS network. More specifically, the average number of VOBs associated with a link in the network varies between 3.95 and 4.95, whereas in the OBS architecture 12.5 independent flows are associated with each link, on average. Recall that our goal is to minimize the number of VOBs (flows) per link in the network in order to reduce burst contentions. Also, comparing the loads per link for different settings in Fig. 10, we see that the VOB does not affect the average traffic load assigned to each link in the network noticeably, which implies that the routing of the traffic in the considered VOB network is not effectively different from the shortest path routing applied in OBS.

[3]There are two candidate paths for each source–destination pair on a bidirectional ring topology.

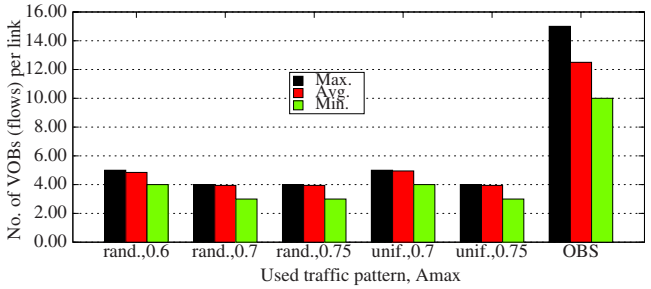

Fig. 9. (Color online) Number of VOBs per link resulting from solving the ILP for the ring network for different traffic matrices (random and uniform) and for different values of A_{max}. The results are compared with the classical OBS network for which the number of independent flows per link are shown.

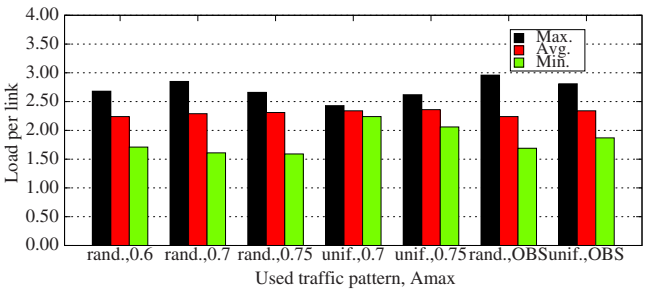

Fig. 10. (Color online) Load per link resultng from solving the ILP for the ring network for different traffic matrices (random and uniform) and for different values of A_{max}. The results are compared with the classical OBS network. Load values are normalized to the capacity of a single wavelength channel.

Now let us consider the differences among the solutions at different values of A_{max}. First, we observe that under the random traffic matrix, increasing A_{max} from 0.6 to 0.7 results in a reduction of the number of VOBs per link; however, increasing A_{max} beyond 0.7 does not affect the number of VOBs per link. This indicates that the impact of A_{max} on the average number of VOBs per link is a stepwise relation, i.e., there is a range of values of A_{max} over which the average number of VOBs per link remains unchanged. On the other hand, although setting A_{max} to any value within this range does not affect the number of VOBs per link, we expect that it impacts the performance of the network. For example, in the considered scenarios, the maximum number of VOBs per link in the network is 4 at both values of $A_{max}=0.7$ and 0.75 for the random traffic matrix, and as each link in the network supports 4 wavelength channels, the loss rate would be zero in the network at both values of A_{max}. However, we expect the access delays to be higher for $A_{max}=0.75$ as the VOBs are more packed in this case. As a conclusion, it would be better—in terms of the access delay—to design the VOB layout of a network at the minimum value of A_{max} in the interval where the average number of VOBs per link remains fixed.

In contrast to the case with random traffic, under the uniform traffic matrix, varying A_{max} around 0.7 causes larger variation in the number of VOBs per link. In fact, under the uniform traffic matrix all the demands are the same and therefore the playground for the optimization is not as large as that under random traffic. For example, for the uniform traffic, the maximum number of O-D flows that can be supported on a single VOB on a link is limited to 4 and 3 for $A_{max}=0.7$ and 0.75, respectively, and, if the number of flows is increased beyond these values, the corresponding A_{max} will be violated. That is, at $A_{max}=0.7$ a VOB could be loaded at most to $3\times0.187=0.561$ of the channel capacity; thus the number of required VOBs to support the traffic demand at $A_{max}=0.7$ would be much larger than that at $A_{max}=0.75$.

2) NSFNET: The traffic matrix used for the NSFNET topology is a measurement-based estimation of traffic in the NSFNET backbone as given in [13]. The estimation presented in [13] shows the aggregate number of bytes between each pair of nodes in a 15 min interval. We use these numbers as an approximate indication of relative intensities of

demands between pairs of nodes in the network and accordingly scale the numbers in a way that the maximum element of the demand matrix is equal to 7 Gb/s. Note that if the demand between any two nodes in the network is larger than around 70% of the capacity of a single wavelength, it might make more sense to establish a direct point-to-point lightpath between the two nodes. The resulting demand matrix is depicted in Table III, where each value is normalized to 10 Gb/s.

In contrast to the ring topology, here the total number of possible paths that can be generated by the k-shortest path routing algorithm in the preprocessing phase is not a small number. Therefore, in order to evaluate possible impacts of the total number of paths on the results of optimization, we solve the ILP with four sets of VOB candidates $\mathbb{P}$, where each has a different size P. Taking into account that in the NSFNET with 14 nodes, the set $\mathbb{P}$ should at least contain 182 paths, i.e., one path for each source–destination pair, we solve the ILP formulation for $P=182,364,546,728$, which correspond to 1, 2, 3, and 4 paths per source–destination pair, respectively. For solving the ILP in all the considered cases, we set $A_{max}=0.7$. Summaries of the results obtained through solving the corresponding ILPs are depicted in Figs. 11 and 12. Also, details of the solutions are presented in [11]. As for the routing of the flows, the detailed solutions presented in [11] show that, as expected, the asymmetry between the routes for flows with the same end points increases with k, such that at $k=4$ around 56% of bidirectional flows take different paths on opposite directions. Furthermore, processing of the solutions demonstrates that, on average, only 3.93 to 5.43 insertion buffers are required per node for different values of k.

First observe in Figs. 11 and 12 that, as expected, applying the clustering significantly reduces the number of independent flows per link in the VOB network as compared with the classical OBS network. For instance, at $k=4$ the maximum number of independent flows per any link in the VOB network is equal to 4, whereas it is equal to 18 in the OBS network, i.e., around 78% reduction is observed. In addition, it is seen that even establishment of VOBs based on the set of paths obtained from the first shortest path routing, i.e., $k=1$, considerably reduces the number of flows per link.

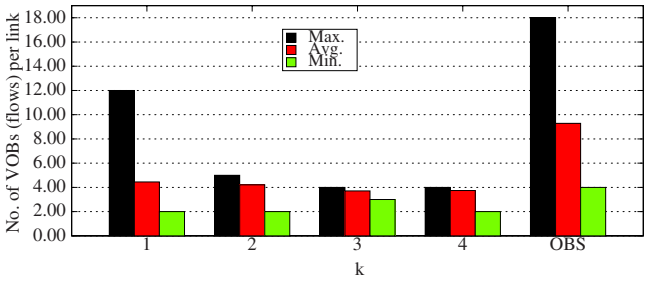

Fig. 11. (Color online) Number of VOBs per link resulting from solving the ILP for the NSFNET topology at $A_{max}=0.7$ and for different values of k. The results are compared with the classical OBS network for which the number of independent flows per link are shown.

As for the impact of k on the optimization results, it is seen that at the beginning, increasing k from 1 to 2 leads to a sharp reduction in the number of VOBs per link. Nevertheless, beyond $k=2$, further improvements of the results through increasing k slows down. This is in fact because the optimum value of the objective function is approached. More specifically, the optimum value of the objective function is achieved at $k=3$ and increasing k will not improve the solution any further. Nevertheless, increasing k from 3 to 4 leads to the slightly smaller number of total VOBs in the network since it results in the selection of longer VOBs by the solver. This is also reflected in the increased value of the average number of VOBs per link at $k=4$ as compared with that at $k=3$.

B. Performance Results

To quantify the performance of the VOB networks explained in the last section we use a simulation model developed in OMNeT++. The simulation model can be set to

TABLE III
TRAFFIC MATRIX OF THE NSFNET

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.00	0.08	0.03	0.01	0.06	0.00	0.02	0.07	0.01	0.01	0.09	0.03	0.06	0.01
1	0.21	0.00	0.18	0.09	0.17	0.08	0.12	0.46	0.03	0.06	0.24	0.31	0.23	0.16
2	0.03	0.14	0.00	0.14	0.03	0.11	0.03	0.26	0.03	0.01	0.15	0.02	0.06	0.04
3	0.02	0.02	0.04	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.02	0.01
4	0.37	0.48	0.06	0.01	0.00	0.01	0.03	0.19	0.07	0.05	0.22	0.35	0.06	0.04
5	0.01	0.05	0.01	0.02	0.01	0.00	0.01	0.01	0.00	0.01	0.02	0.01	0.02	0.00
6	0.11	0.19	0.31	0.01	0.07	0.02	0.00	0.34	0.06	0.07	0.46	0.28	0.49	0.07
7	0.04	0.70	0.63	0.03	0.08	0.01	0.29	0.00	0.13	0.10	0.27	0.21	0.27	0.06
8	0.25	0.06	0.11	0.02	0.07	0.02	0.07	0.18	0.00	0.12	0.33	0.44	0.19	0.14
9	0.01	0.13	0.03	0.01	0.07	0.03	0.01	0.17	0.02	0.00	0.11	0.08	0.04	0.04
10	0.03	0.11	0.17	0.02	0.03	0.04	0.06	0.11	0.06	0.07	0.00	0.18	0.11	0.10
11	0.09	0.39	0.06	0.04	0.13	0.02	0.05	0.17	0.12	0.09	0.63	0.00	0.20	0.08
12	0.24	0.68	0.16	0.07	0.27	0.09	0.10	0.27	0.09	0.00	0.39	0.41	0.00	0.19
13	0.12	0.17	0.06	0.02	0.03	0.00	0.01	0.07	0.35	0.11	0.21	0.24	0.16	0.00

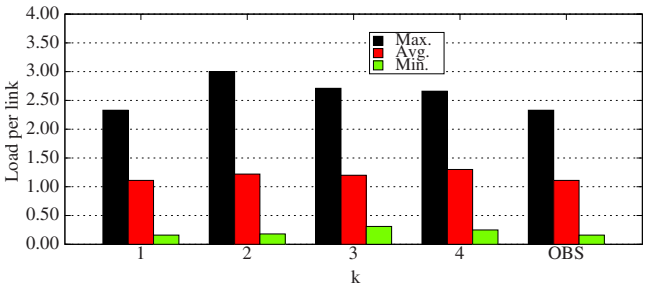

Fig. 12. (Color online) Load per link resulting from solving the ILP for the NSFNET topology at $A_{max}=0.7$ and at different values of k. The results are compared with the classical OBS network. Load values are normalized to the capacity of a single wavelength channel.

work under VOB or OBS. In either case, the just enough time (JET) protocol [3] is used for channel reservation in the network. Packets for each O-D flow are generated according to the Poisson process with a deterministic packet size of 10 KB. The assumptions on the traffic are in line with the traffic analysis presented in [8]. For the ring network the arrival rates of the packets are set in a way that the average load generated for each flow is in agreement with the corresponding value in Table II for the random traffic scenario and is equal to 0.187 for the uniform traffic matrix. Similarly for the NSFNET scenario, the arrival rates are set according to load values given in Table III. For the ring network we consider two scenarios: without FDL buffers ($F=0$) and with a single FDL buffer per output port ($F=1$). In case $F=1$, the length of the delay line is set in a way that it provides delay equal to the transmission time of a single packet over the channel. For the NSFNET scenario, we only consider the case without FDL buffers. Other settings at the nodes are the same as those explained in Section V. For the sake of simplicity, in our simulations we assume the propagation delays in the network to be negligible. Additionally, in case of OBS, the shortest path routing is used for routing packets through the network.

Tables IV and V present the results of simulating the ring network and the NSFNET, respectively. Shown in the tables are the average values the of the packet drop rate, throughput, and access delay as well as the maximum value in the set of the average access delays of all the O-D flows for each scenario. The results are compared with those obtained from the simulation of OBS.

We observe that the packet drop rate improves greatly when the VOB architecture is employed in comparison with OBS. Specifically, for the case in which the maximum number of VOBs per link is equal to the available wavelength channels on the link, the loss rate is completely suppressed and no instance of packet loss is observed during the long simulation periods, as expected. Also, for the case in which the number of wavelength channels are smaller than that of VOBs, a large reduction—up to 2 orders of magnitude—in the loss rate is observed.

The elimination/reduction of the loss rate, which has been the first objective of designing the VOB architecture, gives

TABLE IV
PERFORMANCE EVALUATION RESULTS FOR THE RING NETWORK[a]

			Average Packet Drop Rate	Average Network Throughput (Mb/s)	Average Access Delay Over All Flows (μs)	Max Access Delay Over All Flows (μs)
Random traffic	$F=0$	VOB $A_{max}=0.6$	3×10^{-2} $\pm1.5\times10^{-5}$	163,225 ±18	5.91 ±0.1	16.83 ±0.13
		VOB $A_{max}=0.7$	0 —	168,645 ±5.02	12.3 $\pm6.5\times10^{-3}$	52.7 ±0.77
		VOB $A_{max}=0.75$	0 —	168,645 ±4.7	17.01 $\pm5\times10^{-2}$	111 ±5.1
		OBS	9.4×10^{-2} $\pm2.4\times10^{-5}$	152,763 ±8.9	1.7 $\pm3.7\times10^{-4}$	28 ±0.1
Uniform traffic	$F=0$	VOB $A_{max}=0.7$	3.6×10^{-2} $\pm2.4\times10^{-5}$	162,287 ±26.2	6.2 $\pm7.5\times10^{-3}$	12.2 ±2.2
		VOB $A_{max}=0.75$	0 —	168,657 ±32	21.1 $\pm4.9\times10^{-2}$	62.7 ±11.5
		OBS	1.2×10^{-1} $\pm6.6\times10^{-5}$	148,250 ±25.2	1.2 $\pm7.5\times10^{-3}$	1.8 ±1.3
	$F=1$	VOB $A_{max}=0.7$	5.9×10^{-5} $\pm6.4\times10^{-7}$	168,561 ±7.1	6.6 $\pm1.6\times10^{-3}$	12.6 ±2.3
		OBS	3.8×10^{-3} $\pm1.7\times10^{-5}$	167,732 ±30.3	2.1 $\pm3.5\times10^{-3}$	4.2 $\pm6.5\times10^{-1}$

[a]The average values are shown with 90% confidence level.

TABLE V
PERFORMANCE EVALUATION RESULTS FOR THE NSFNET[a]

	Average Packet Drop Rate	Average Network Throughput (Mb/s)	Average Access Delay Over All Flows (μs)	Max Access Delay Over All Flows (μs)
VOB ($k=1$)	1.08×10^{-2} $\pm2.42\times10^{-5}$	217,022 ±27.19	3.74 $\pm4.9\times10^{-3}$	20.11 ±0.16
VOB ($k=2$)	6.6×10^{-3} $\pm7.47\times10^{-6}$	217,947 ±27.52	5.86 $\pm4.5\times10^{-3}$	43.69 ±0.73
VOB ($k=3$)	0 —	219,400 ±28.18	6.75 $\pm8.8\times10^{-3}$	49.65 ±0.92
VOB ($k=4$)	0 —	219,400 ±28.19	6.93 $\pm1\times10^{-2}$	42.07 ±0.43
OBS	1.75×10^{-2} $\pm2.01\times10^{-5}$	215,567 ±17.4	2.18 $\pm1.7\times10^{-3}$	9.41 $\pm2.7\times10^{-2}$

[a]The average values are shown with 90% confidence level.

rise to the improvement of the overall throughput of the network. For the considered ring network the throughput improvement is up to 13.57%, which is associated with the uniform traffic with $A_{max}=0.75$ and without a FDL buffer. Also for the case of NSFNET around a 1.7% increase in the overall network throughput is achieved. As observed, the improvements in the throughput for the NSFNET scenario as well as the ring network at $F=1$ are marginal. The reason is that for these cases the fraction of packets that are dropped in the network under OBS is of the order of 10^{-2}–10^{-3}, i.e., there is marginal room for improvement in the throughput. It should be, however, noted that although a good throughput is achieved for these cases under classical OBS, the drop rate is not acceptable taking into account that each—dropped—burst is the result of aggregating several packets, which are usually consecutive packets of a flow.

Table V also includes the results achieved at different values of k, which in turn reflects the impact of routing of O-D flows on the performance of the VOB network. For instance, at $k=1$ the same routing is applied in the VOB network as in OBS. From the table it is clear that the routing of flows has an important role in the VOB architecture.

The cost of improvement in the packet loss rate and the network throughput is the access delay penalty that is introduced by the flow coordination in the VOB approach. Nevertheless, it is seen in Tables IV and V that the delay penalty is negligible for all the investigated scenarios—taking into account that propagation delays in a typical metro/core network are in the range of tens of ms. For instance, the average length of the links in NSFNET is 1081 km, which leads to 5.4 ms propagation delay per hop, on average. The acceptable access delays indicate that the values of A_{max} are set properly in the VOB layout design phase.

To sum up, it is observed that in the VOB network design there is a trade-off between the edge shaping delay, i.e., the VOB access delay, and the loss rate inside the network. In the extreme case where there is no clustering in the network, we expect that the performance of the network approaches that of the classical OBS. In this case, the VOB MAC protocol will perform very similar to OBS since there will be no need for checking the insertion buffer before injecting the packets into the network. When clustering is introduced in the network, on one hand, burst losses are reduced, and on the other hand, access delays increase, where the amount of these changes in the loss rate and access delay depends, among other things, on how dense the clusters are. In fact, allowing larger shaping delays at the edge of the network, realized by increasing the A_{max}, results in the formation of denser clusters and thereby further reducing burst losses. Moving in this direction—depending on the load distribution and the available network resources—a network operation point is reached where the maximum number of VOBs per link is equal to the number of channels per link and therefore the loss rate inside the network is fully suppressed. Beyond this point, making the VOBs denser results merely in increased shaping delay.

C. Sensitivity of the VOB Network to Load Variations

In the design and performance evaluation of the VOB networks in the previous sections we assumed that the traffic matrix is fixed. However, in reality the traffic matrix changes over time, which might necessitate redesign of the VOB network layout in order to achieve the desired performance. Therefore, in this section we discuss how changes in the traffic matrix can affect the performance of the VOB network. For this purpose, we consider the ring network that was introduced earlier in this section. The demand matrix given in Table II is taken as the initial traffic matrix. Based on this, new traffic matrices are produced that differ from the initial one—in terms of the load intensity—between −20% to +20%. More specifically, the initial traffic matrix is multiplied by 0.8, 0.9, 1.1, and 1.2 to generate traffic matrices with −20%, −10%, +10%, and +20% difference, respectively. For each new traffic matrix, the performance of the ring network is evaluated under three settings: i) the VOB network is designed for the new traffic matrix, ii) the VOB network is designed for the initial traffic matrix, and iii) the network operates under classical OBS. Due to the space

limit, we only present the conclusions drawn from these experiments and detailed results are not shown here.

First let us consider the case that the load increases. In this case, the burst loss rate is smaller in the network designed for the initial traffic matrix than in the network designed for the new traffic matrix. For instance, at $A_{max}=0.7$ with +20% load change, the loss rate is zero in setting II, while it increases to 7.1×10^{-2} in setting I. On the other hand, for the same scenario the maximum VOB access delay in setting II is around 1 order of magnitude larger than that in setting I. To explain this behavior we recall that in the network designed for the initial traffic matrix, the maximum number of VOBs per link is equal to the number of wavelength channels per link, which means no loss rate in the network. Accordingly, when the load increases in this network we still have the zero loss rate, since the VOB layout does not change in the network. Instead, the effect of the load increase is manifested in the delay because load per VOB violates the A_{max} value that the network has been designed for. Consequently, when the desired A_{max} value is enforced by redesigning the network based on the new traffic matrix, the maximum number of VOBs per link increases to 5, which results in a reduction of the VOB access delay at the cost of introducing some burst loss in the network. That is, as discussed earlier, there is a trade-off between the VOB access delay and the burst loss rate in the network.

Following the same line of reasoning, three different situations can be identified when the load in the network scales down. The first situation occurs when the burst loss rate is zero under the initial traffic matrix. In this situation the reduction of the load in the network designed for the initial load has no impact on the loss rate, but reduces the VOB access delay. Therefore, redesigning the VOB layout is not necessary in this case. The second situation occurs when there are already burst losses in the network under the initial traffic matrix and the reduction in the load is so large that redesigning the VOB layout of the network can result in a smaller number of VOBs per link and thereby can fully suppress the burst loss in the network. In this case, further operation of the network under the layout designed for the initial traffic results in burst losses, which would be avoided if the VOB network is redesigned based on the new traffic matrix. Therefore, in this case, redesigning the network is recommended. The third case is the same as the second one, except that, if the network is redesigned, the change in load reduction is not large enough to reduce the maximum number of VOBs per link. Again, it is not necessary in this case to redesign the network.

Overall, we can argue that for the new traffic matrices, the operation of the VOB network, which has been designed for the initial traffic matrix, is always better than classical OBS. Though, as pointed out before, this comes at the cost of increased access delay. This, together with the above discussion, suggests that the VOB network can safely handle small variations—in the range of ±10%–20%—in load without requiring the designed layout to be changed.

VII. Related Work

Although there is a large number of publications on the contention resolution issue of OBS networks, only a few of them consider proactive measures to reduce the packet collisions rate. In this section we outline the most relevant approaches and compare them with the VOB.

The authors in [14] present a new packet-based optical transport architecture based on a light trail (LT), which is a generalized form of a lightpath (optical circuit) where intermediate nodes along the path of the circuit may also inject traffic into it. Therefore, a LT could also be considered as an optical bus that facilitates traffic grooming in the network. Our approach is essentially different from the LT. Specifically, in the LT approach the whole capacity of a wavelength channel must be dedicated to each LT on any link in the network being part of that LT and there is no switching in the intermediate nodes. That is, the LT approach allows only for intrawavelength multiplexing. With the VOB approach, however, a single wavelength on a link may be shared by all the VOBs on that link, i.e., both intrawavelength and interwavelength multiplexing are realized. On the other hand, in order to transmit any packet using LT one needs to establish a LT in advance, whereas in the VOB establishing a virtual bus can take place in the background while transmission of packets is in progress.

In [15] the authors present a proactive burst scheduling algorithm aiming at reducing the collision rate in the network. In this approach, every node injects its local packets into the network in a way that they do not overlap in time with each other if the packets are supposed to traverse the same path. Our approach based on VOB differs from proactive burst scheduling in that the latter one only does a local coordination at every ingress node through implementing a new MAC protocol assuming that routings of the O-D flows are already fixed. In the VOB-based approach, however, network-wide traffic engineering is applied, which also involves finding appropriate paths for the flows to reduce the potential of collisions in the network.

A dual-bus optical ring (DBORN) is an architecture for optical metro transport networks that is introduced in [16]. A metro network based on a DBORN is composed of two WDM rings, where one is used as the working ring and the other one as the backup and WDM channels on each link are divided into upstream and downstream channels. All nodes attached to the ring are passive except one node that is called the hub and is the only node that can transmit the traffic over the downstream channels and receives or removes traffic over the upstream channels. All the passive nodes have access to the upstream channels using an optical carrier sense multiple access with collision avoidance (CSMA/CA) MAC protocol. In order that node i transmit a data packet to node j, where i and j are both attached to the ring, node i has to first transmit the packet to the hub over one of the upstream channels and then the hub node would forward the packet to node j on one of the downstream nodes. In DBORN both the transmission and the packet processing of the whole network are limited by those of the hub node. Also, the resources in the network are not utilized efficiently since, for example, to transmit a packet from a given node to its adjacent node the packet has to travel the whole circumference of the ring network first. In contrast to the DBORN, the VOB can achieve high network utilization since source–destination pairs communicate directly and without using a hub. In addition, DBORN is a ring-based

approach and is not extended to an arbitrary meshed topology, whereas the VOB architecture can be applied to any meshed topology.

The authors in [17] present time-domain wavelength-interleaved networking (TWIN) as a new optical transport architecture. TWIN consists of smart edge nodes that are equipped with ultrafast tunable lasers and passive fixed-routing optical core nodes. A unique wavelength is assigned to each destination node in the network, which by the way serves as the address of that node. In order to transmit a packet from node i to node j, node i has to tune its tunable laser to the wavelength associated with node j for the duration of the packet transmission. All the passive intermediate nodes are configured in advance to guide incoming packets towards their destinations based on the wavelength channel that the packets are arriving on. In this way, a multipoint-to-point tree is formed to every egress node in the network. The key design issue of the TWIN architecture is to devise a scheduler that can handle a collision-free transmission of packets on a given destination tree. Designing an appropriate scheduling algorithm that can work efficiently under highly dynamic traffic is a challenging issue. In fact, a fixed scheduling for all nodes of a tree has to be calculated before any packet transmission can take place on that tree. This is, however, not required in the VOB architecture that instead of trees uses buses, which are much easier to coordinate. Additionally, the number of wavelength channels in TWIN must always be at least equal to the number of nodes in the network, whereas in the VOB there is no constraint on the required number of wavelengths in the network.

VIII. Conclusion

We have presented the VOB as a new architecture for dynamic subwavelength allocation of bandwidth in optical transport networks. The VOB is based on establishing coordination among a set of ingress nodes in the network in a way that no collision occurs among packets belonging to these nodes. We formulated the design of a VOB-based network, which consists of grouping all aggregate flows into VOBs, as an ILP problem with the objective of minimizing inter-VOB collisions.

Through a simulation-based performance study, it is shown that the VOB network outperforms classical OBS in terms of loss rate and throughput. The achieved gain comes at the cost of a marginal increase in the access delays. Also, two design examples are presented that further quantify the merits of the VOB network architecture. Our numerical results demonstrate that this approach can greatly alleviate the contention problem that is central to the design of OBS networks. As one of the design examples, we presented a bidirectional ring network with 10 nodes and 4 wavelength channels—10 Gb/s each—in each direction. The presented network achieves 168.6 Gb/s throughput and has zero packet loss inside the network. The same network under OBS operation has 148.2 Gb/s throughput—13.57% less—and suffers from a 12% packet loss rate. Through another design example similar performance improvements are demonstrated for NSFNET topology.

There are some issues that can be considered as the future direction to the presented architecture. Among others, an appropriate signaling scheme has to be developed for setting up VOBs inside the network. Also, the impact of VOB architecture on the design of protection mechanisms is an issue that deserves further study.

Acknowledgments

This work was supported in part by the European Commission (7th ICT Framework Programme) within the Network of Excellence project BONE (Building the Future Optical Network in Europe) and within the Collaborative project 4WARD.

References

[1] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: a new area in optical networking research," *IEEE Network*, vol. 18, no. 3, pp. 16–23, 2004.

[2] A. Rostami and A. Wolisz, "Impact of edge traffic aggregation on the performance of FDL-assisted optical core switching nodes," in *IEEE Int. Conf. on Communications (ICC)*, 2007, pp. 2275–2282.

[3] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1838–1851, 2000.

[4] H. R. van As, "Media access techniques: the evolution towards terabit/s LANs and MANs," *Comput. Networks ISDN Syst.*, vol. 26, nos. 6–8, pp. 603–656, 1994.

[5] H. Akimaru and K. Kawashima, *Teletraffic: Theory and Applications*. Springer, 1999.

[6] M. Pióro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan Kaufmann, 2004.

[7] D. Eppstein, "Finding the k shortest paths," *SIAM J. Sci. Comput.*, vol. 28, no. 2, pp. 652–673, 1999.

[8] A. Rostami, A. Wolisz, and A. Feldmann, "Traffic analysis in optical burst switching networks: a trace-based case study," *Eur. Trans Telecommun.*, vol. 20, no. 7, pp. 633–649, 2009.

[9] OMNeT++ User Manual. Available: http://omnetpp.org/documentation.

[10] H. Buchta, "Analysis of physical constraints in an optical burst switching network," Ph.D. thesis, Technical University of Berlin, Germany, 2005.

[11] A. Rostami, "Virtual optical bus: a novel packet-based architecture for optical transport networks," Telecommunication Networks Group (TKN), Technical University of Berlin, Tech. Rep., Feb. 2010.

[12] CPLEX 9.0 User's Manual. ILOG, SA, 2003.

[13] B. Mukherjee, *Optical WDM Networks*. Springer, 2006.

[14] A. Gumaste and I. Chlamtac, "Light-trails: an optical solution for IP transport," *J. Opt. Netw.*, vol. 3, no. 5, pp. 261–281, 2004.

[15] J. Li and C. Qiao, "Schedule burst proactively for optical burst switched networks," in *IEEE Global Communications Conf.*, 2003, pp. 2787–2791.

[16] G. Hu, C. M. Gauger, and S. Junghans, "Performance of MAC layer and fairness protocol for the Dual Bus Optical Ring Network (DBORN)," in *Int. Conf. on Optical Networking Design and Modeling*, 2005, pp. 467–476.

[17] I. Widjaja and I. Saniee, "Simplified layering and flexible bandwidth with TWIN," in *ACM SIGCOMM Workshop on Future Directions in Network Architecture*, 2004, pp. 13–20.