Available online at www.sciencedirect.com
ScienceDirect
Optical Switching and Networking 5 (2008) 177–187
Optical Switching and Networking
www.elsevier.com/locate/osn

A non-competing hybrid optical burst switch architecture for QoS differentiation

Kyriakos Vlachos*, Kostas Ramantas

Computer Engineering and Informatics Department and Research Academic Computer Technology Institute, University of Patras, GR26500, Rio, Patra, Greece

Received 12 March 2007; received in revised form 23 November 2007; accepted 1 December 2007
Available online 15 December 2007

Abstract

In this paper, we present a new hybrid optical burst switch architecture (HOBS) that takes advantage of the pre-transmission idle time during lightpath establishment. In dynamic circuit switching (wavelength routing) networks, capacity is immediately hard-reserved upon the arrival of a setup message at a node, but it is used at least a round-trip time delay later. This waste of resources is significant in optical multi-gigabit networks and can be used to transmit traffic of a lower class of service in a non-competing way. The proposed hybrid OBS architecture, takes advantage of this idle time to transmit one-way optical bursts of a lower class of service, while high priority data explicitly requests and establishes end-to-end lightpaths. In the proposed scheme, the two control planes (two-way and one-way OBS reservation) are merged, in the sense that each SETUP message, used for the two-way lightpath establishment, is associated with one-way burst transmission and therefore it is modified to carry routing and overhead information for the one-way traffic as well. In this paper, we present the main architectural features of the proposed hybrid scheme and further we assess its performance by conducting simulation experiments on the NSF net backbone topology. The extensive network study revealed that the proposed hybrid architecture can achieve and sustain an adequate burst transmission rate with a finite worst case delay.

Keywords: Hybrid optical switching; Optical burst switching; Service differentiation

1. Introduction

Optical burst switching combines the merit of both optical circuit and packet switching and has been proposed for dynamic optical networking for the on-demand use of capacity [1]. OBS is a feasible technology but when it relies on the one-way forwarding paradigm, it yields a high loss ratio especially for high loads and does not guarantee data delivery. In order to enable QoS provision and service differentiation in OBS networks various one-way schemes have been introduced, including the JIT *offset-time-based* scheme that uses time offsets to isolate different classes of traffic [2], the *composite-burst assembly* scheme that mixes traffic classes during burst assembly and provides QoS via prioritized burst segmentation [3], the *preemptive wavelength reservation* mechanism, where each class is associated with a predefined usage limit [4] and the *early dropping* mechanism that randomly drops bursts depending on their class [5]. An interesting hybrid signaling scheme that employs a two-way reservation up to an intermediate node followed by

* Corresponding author. Tel.: +30 2610 996990; fax: +30 2610 969007.
E-mail address: kvlachos@ceid.upatras.gr (K. Vlachos).

1573-4277/$ - see front matter
doi:10.1016/j.osn.2007.12.003

an (unacknowledged) one-way reservation process until the egress node has been proposed in [6]. This hybrid protocol provides a trade-off between burst loss and delay (by selecting the intermediate node) and thus enables QoS differentiation. However, only two-way burst reservation schemes [7,8], can guarantee lossless burst transmission in the core albeit they introduce an extra delay, associated with the round-trip time.

A very promising approach of optical burst switching is hybrid optical burst switching. General hybrid switching combines the merits of both circuit and burst/packet switching paradigms to increase the link utilization efficiency, decrease the required number of wavelengths and maintain to the bear minimum the traffic load processed by electronic IP routers. Furthermore, it allows the easy disjunction of classes of service, in the sense that data delivery in circuit switching networks is guaranteed, while that in OPS/OBS is not. Various hybrid switching architectures have been proposed so far, including the *hybrid optical switching* – HOS – [9–11], and the *hybrid optical transport network* (HOTNET) [12], where optical circuit and packet switching are integrated in a cooperative manner to transport a variety of traffic types efficiently by complementing each other. In the HOS approach, ingress nodes classify incoming traffic flows in small flows (best effort traffic) and large flows. Small flows are transported using OBS (through HOS-B modules) and large flows are transported using OCS (through HOS-C modules). HOS-B and HOS-C modules coexist in HOS core nodes, and compete for all wavelengths during the resource reservation process. HOTNET integrates TDM-WRN architecture (which refers to time-division multiplexed wavelength routing) with slotted OBS, and it can fulfil QoS requirements. In HOTNET the time-slot-based switching provides fine granularity, which results in a better utilization of resources. HOTNET ingress nodes first attempt to serve incoming traffic flows through pre-established optical circuits. If bandwidth is not sufficient, the remaining traffic is transmitted through slotted OBS. Other pure hybrid switching approaches include the light-trail [13], the light bus concept [14], the polarization-based scheme [15] and the ORION concept [16]. In the first two schemes all intermediate nodes between any source–destination pair have access on the established optical circuits, without the need for optical switch reconfigurations. Both schemes are particularly beneficial for transporting IP traffic, which exhibits burstiness, since they are able to achieve high statistical multiplexing. Finally, ORION architecture allows full sharing of all wavelengths on a link, as a means to cope with short-term temporal traffic imbalances [17]. ORION's basic principle of operation is that it allows the transparent insertion and removal of packets during the idle periods of established wavelength paths. As a result, ORION is able to obtain significant statistical multiplexing gains. Finally, in the polarization-based concept, the polarization state (SOP) is used to differentiate OCS traffic from best effort IP traffic.

In this paper, we present a radically different hybrid optical burst switch (HOBS) architecture that combines one-way with two-reservation under a single, unified control plane (hybrid signaling) for QoS differentiation. It takes advantage of the idle, round-trip time delay during lightpath establishment phase to transmit one-way data bursts of a lower class of service, while high priority data explicitly requests and establishes end-to-end connections (lightpaths), as in wavelength-routed OBS [7,8]. In high-speed, optical core networks, the pre-transmission, idle period of these lightpaths that use immediate reservation protocols, i.e. CR-LDP or RSVP, is significant and thus capacity is inefficiently used. For example, a medium-sized file of 200KB size over a distance of 1500 km at 10 Gb/s yields an efficiency of only 57%. In the proposed scheme one-way and two-way reservations do not compete for bandwidth resources, while both are serviced with the same signaling message. Thus, the SETUP message that is being used for setting up an end-to-end lightpath connection also acts as the control packet in the *tell-and-go* schemes (i.e. *just-enough-time* or *just-in-time*). In this communication, we present the main features of the proposed architecture, the signaling messages needed as well as a detailed performance evaluation. We have developed an extensive network simulator, where the basic features of the architecture were modeled. It is shown that the proposed scheme can achieve and sustain an adequate data transmission rate with a finite worst case delay, while hardware implementation cost is relatively low.

The rest of the paper is organized as follows. Section 2 presents the architectural concept of the proposed scheme, while, Section 3 presents a suitable node architecture. Section 4 presents a thorough performance evaluation at the network level, where the effect of various inter-dependencies is investigated.

2. Network concept

The proposed hybrid architecture has been designed to take advantage of the inefficiency of two-way reservations in OBS or other dynamic circuit switching

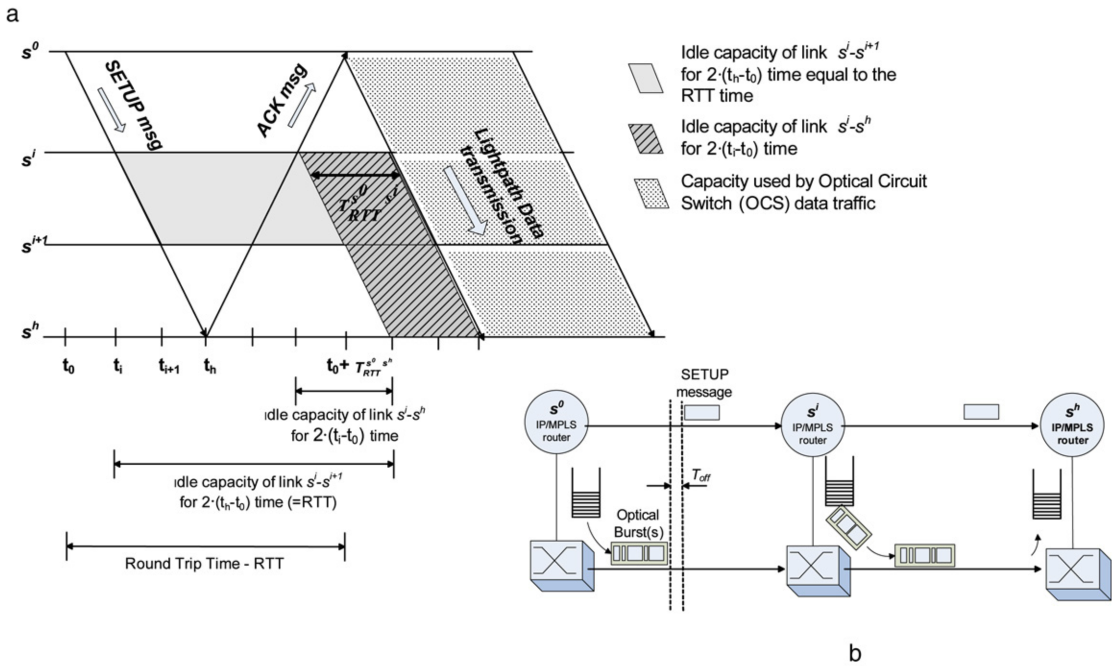

Fig. 1. (a) Timing considerations in the proposed hybrid switching scheme and idle capacity seen by node s^i. (b) Burst insertion in the established lightpath between $s^0 - s^i$ node. Data burst can be extracted at the destination node, s^h or any node across the path.

paradigm. Fig. 1a illustrates the reservation process in a circuit-switched optical network and the corresponding idle capacity for node s^i. Node s^0 generates a SETUP message, according to a bandwidth request, and transmit it along the $s^0 - s^h$ path to establish an optical circuit. When node s^i receives the SETUP packet at t_i time, it hard-reserves outgoing resources of link (L_i^{i+1}) immediately. However, s^i node is aware that data for this session will arrive at least a round-trip time ($T_{RTT}^{s^o-s^h}$) later. This applies to all nodes across the path and therefore all the nodes keep hard-reserved but idle, the capacity of their outgoing links for time equal to $T_{RTT}^{s^o-s^h}$. Therefore this unused capacity can be used to slip-in traffic heading for the next node s^{i+1}. If lightpath is established successfully and the SETUP packet reaches the destination node, an acknowledge message (ACK packet) is generated and sent back to the source. The ACK packet informs the intermediate nodes for the successful establishment of the optical lightpath. Upon reception of the ACK packet by node s^i, node s^i becomes aware that the end-to-end optical path, (up to s^h node), has been established. To this end, node s^i may use this capacity to transparently forward data bursts directly to their end destination. However, such a case requires tight traffic scheduling to avoid bursts transmitted from different nodes to overlap in time, while its efficiency is limited, [17], due to the shorter time that the unused capacity is available. In particular it is only available for a time period equal to $T_{RTT}^{s^o-s^i}$ that is the round-trip time between s^0 and s^i node. From that time onwards, data belonging to the established lightpath is expected to arrive and thus capacity cannot be further exploited.

In the proposed scheme, light synchronization of the one-way bursts is only needed for a large period equal to the source–destination round-trip time. One-way burst transmission can be lossless as long as the associated SETUP message is successfully forwarded, while critical information like burst size and relevant time arrivals are communicated to all nodes via the same signaling message. In a usual OBS network, each data burst is associated with a control packet that is transmitted prior to the burst to reserve resources. In the proposed hybrid scheme, this information is carried by the SETUP packet.

Fig. 1b displays an example of the operation. Edge router s^0 generates a SETUP, according to a lightpath request and retrieves from a local buffer a data burst with a matching destination. The burst overhead information (size, time offset, etc) is encoded in the SETUP message. It then transmits the message as well as the associated burst (or bursts) with a time offset, T_{off} as in the usual case of a one-way protocol. Intermediate node s^i receives the SETUP message and process its fields, while transparently forwarding its associated burst. It may also add (if any) a new burst at the end. In the case that the reservation fails, a REJECT packet is created and is sent back to free resources. In that case, burst losses are inevitable, so as to rule out possible interplay between the different classes of traffic. However, resources on the individual links remain reserved up to the time that the REJECT message is processed by the corresponding nodes. It must be noted here that the transmission of bursts is guaranteed to be lossless, since capacity is reserved but remains idle. There is no case, where a lightpath is accepted and a burst is not, as long as there is enough capacity during the idle, pre-transmission time.

Based on the above analysis, we have defined three different cases for routing/forwarding one-way burst traffic:

- *Burst(s) transmission with a matching destination from the edge node.*This is the simple case described above. For each lightpath request, a single burst with a matching destination is associated and subsequently transmitted.
- *Burst(s) transmission with a matching destination from any node across the path*. In this case, each intermediate node across the network path, adds a new burst with a matching destination with the SETUP message. The new bursts are added at the end, and therefore a train of bursts is formed. The SETUP message is updated at each node to communicate to the next nodes the burst additive size.
- *Burst(s) transmission from all-to-all nodes across the path.*This is the most complex case, where each node adds (or extracts) optical bursts heading for (or coming from) any node across the path. The implementation of this case requires a tighter scheduling, so that each core node receives its own bursts, while transparently forwarding the rest. The way that the burst train is formed as well as how the individual bursts are scheduled in the burst train, is encoded in the SETUP message (see Section 3.2).

Finally, it should be noted here that the proposed scheme can be modified to operate with two-way reservation mechanisms, that hard-reserve resources with the acknowledge message during backward propagation. With respect to the example of Fig. 1b, node s^i may transmit on-way bursts directly to the destination (or other node across the path), upon receiving the acknowledge message and for time equal to $T_{RTT}^{s_0-s_i}$. Alternatively upon receiving the setup message, it may transmit one-way bursts to only the next node for $T_{RTT}^{s_0-s_h}$ time or until an acknowledge message for another session arrives. In what follows, we will focus on reservation mechanisms that hard-reserve capacity in a downstream propagation mode for reasons of simplicity.

3. Switch architecture and SETUP message format

3.1. Node architecture

In this section, a suitable node architecture is presented, capable of handling both one-way and two-way OBS traffic. Emphasis is given to the additional hardware needed to handle optical bursts in order to show, how the proposed scheme can be effectively integrated into an existing circuit switching network. Fig. 2 shows its architectural design. The only hardware requirements for the hybrid node implementation are a set of 1×2 and 2×1 switches commissioned to extract or insert bursts from/to the establishing optical paths. A set of receivers and tunable transmitters (denoted as OBS Rx and OBS Tx in Fig. 2) are used for this task. Lightpath data are received by IP/MPLS router via the dropped ports of the switch fabric, as in the normal case of an optical circuit switching network. Thus, the node maintains its full functionality when setting up or tearing down optical circuits. Most of the work is being performed in the control plane, where the node employs a dedicated agent that acts as a timer, calculating the time at which the first data of the established lightpath will arrive. Specific modules of this agent are:

Traffic scheduler: The traffic scheduler processes the SETUP messages and in particular, it reads the burst information stored in it. Upon the reception of such a message, the IP/MPLS electronic router proceeds in configuring the switch fabric, while the traffic scheduler acts as follows:

- turns the extraction switch ON to forward the subsequent burst(s) to the electronic domain, if this node is their destination node. The time arrival of the burst(s) destined for this node is stored in the SETUP message as well.
- Turns the extraction switch OFF, while keeping the insertion switch OFF as well, to forward the rest of the bursts to the node output
- Turns the insertion switch ON to insert new burst(s) at the end of the burst train. Insertion of a new burst is allowed only when the total size of the burst train is smaller than the pre-transmission idle time for the specific routing path.

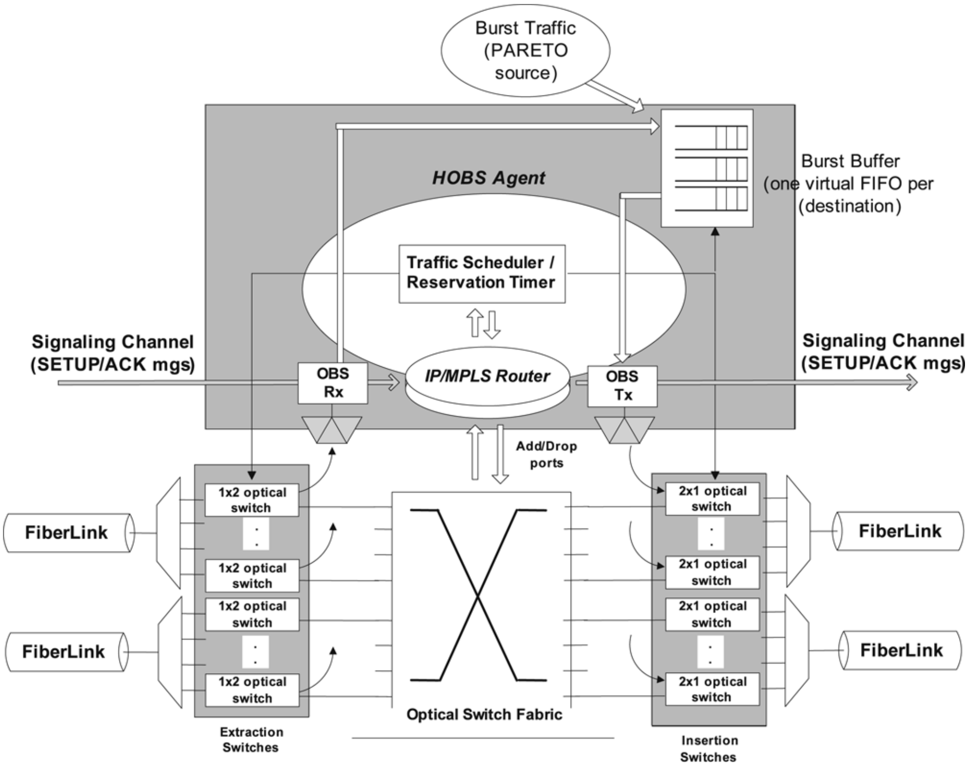

Fig. 2. Hybrid OBS node architecture employing a dedicated agent that acts as a timer, calculating the time at which the first data will arrive. Upon the reception of a SETUP message, traffic scheduler extracts data bursts heading for this node and signals the burst FIFO and the 2 × 1 switch to retrieve and insert additional burst at the burst train.

Burst buffer: This is an electronic buffer that is used to store one-way OBS traffic coming from the access network. The buffer consists of a single per destination FIFO and it is accessed by traffic scheduler to retrieve packets (in the form of a single burst).

Burst transmitters/receivers: These are a dedicated set of transceivers commissioned to transmit and receive OBS traffic. In principle, a separate transceiver is needed for each wavelength from each incoming/outgoing link.

The architecture shown in Fig. 2 is for a core switch. For the edge, the corresponding design is similar but simpler, employing only the burst "*insertion*" part (right part of Fig. 2) since the edge routers are the termination nodes and thus receive OBS data as normal traffic, without the need of an additional extraction switch. The traffic scheduler of the edge node, upon the creation of a new SETUP message, extracts a burst from the buffer with a matching destination and modifies the SETUP message, with its corresponding size and time offset. It then transmits the optical burst, subsequently after the SETUP message. In the case of multiple burst transmission (routing policy #3), the edge node extracts all the bursts with an end-destination across the routing path, modifies the SETUP message again and transmits the bursts one after the other.

3.2. SETUP message format

For the implementation of the proposed hybrid switch architecture, a modified SETUP message is required for signaling purposes. This message should communicate to all nodes across the network path, critical information for all the bursts associated with it, such as their arrival time, size and destination. This information is encoded in the SETUP message with a triple field (s, B, TO) together with other fields that concern the lightpath setup. B field denotes the total data size that is associated with the setup message and may correspond to more than one burst, while TO field is its relevant time offset from the SETUP message.

Fig. 3 shows its formats for the case of (a) bursts transmission with matching destination (either from the initiated node or from any node across the path) and (b) in the case of multiple burst transmission. The path of the SETUP packet is specified as a sequence of link identifiers $L_1, L_{2,} \ldots, L_h$. Message fields that concern one-way bust transmission are organized based on the

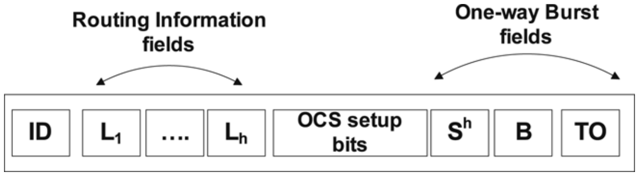

Fig. 3a. The different fields of the SETUP packet for the case of burst transmission with a matching destination from the initiated node only or any node across the path.

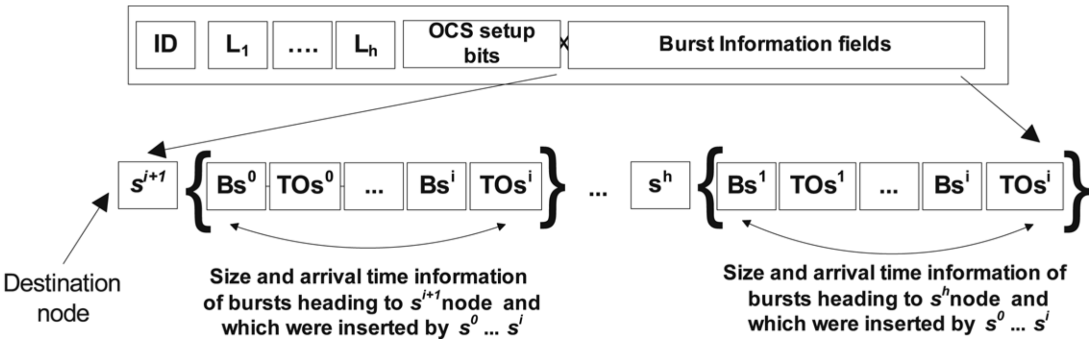

Fig. 3b. The different fields of the SETUP packet for the case of multiple burst transmission from all-to-all nodes across the network path.

destination node and depend on their routing policy. In the first policy (see Fig. 3a), no field is processed, and the burst is forwarded transparently to the next node. In that case, field s^h that defines the destination node of the burst traffic is the same with L_h that defines the destination for the circuit data. Thus, there is no need for any processing, and upon the reception of such a SETUP packet, any of the intermediate nodes transparently forwards the associated bursts to the next node of the path. In the second routing policy, where each node across the path may add bursts with a matching destination, the burst size and the time-offset fields must be processed in order to determine when and where to insert the new bursts in the path. For example, if node s^i, wishes to add a new burst of size B^i, then upon reception of a SETUP message with a matching destination, it process *TO* and *B* fields (denoted as *TO* and B^{i-1}) to identify the end of the burst train and places the new burst at its trailing edge. It then updates *B* field to $B^{i-1} + B^i$ to match the new accumulated size and forwards the message to the next node. *TO* field does not need to be updated.

In the third routing policy, where bursts from all nodes to all nodes across the path can be transmitted, each burst must be represented with its own set of (s, *B, TO*) fields. Thus, bursts added from node s^i and which are heading for any of the subsequent node s^{i+1}, $s^{i+2} \dots s^h$ are denoted in the message field as follows: $(B_{s^i}^{s^{i+1}}, TO_{s^i}^{s^{i+1}}), (B_{s^i}^{s^{i+2}}, TO_{s^i}^{s^{i+2}}), \dots, (B_{s^i}^{s^h}, TO_{s^i}^{s^h})$, where the low and the high index denote the source and the destination node. $TO_{s^i}^{s^j}$ field denotes the time arrival (relative to the setup message) of the burst inserted by node s^i and heading for node s^j. This information is needed for configuring the insertion/extraction switches. To facilitate message processing all the fields are organized in a linked list based on the destination node as follows (see Fig. 3b):

$$
\begin{aligned}
& s^{i+1}[(B_{s^0}^{s^{i+1}}, TO_{s^0}^{s^{i+1}}), (B_{s^1}^{s^{i+1}}, TO_{s^1}^{s^{i+1}}), \dots \\
& (B_{s^i}^{s^{i+1}}, TO_{s^i}^{s^{i+1}})], \\
& s^i[(B_{s^0}^{s^i}, TO_{s^0}^{s^i}), (B_{s^1}^{s^i}, TO_{s^1}^{s^i}), \dots (B_{s^{i-1}}^{s^i}, TO_{s^{i-1}}^{s^i})] \\
& \dots \\
& s^2[(B_{s^0}^{s^2}, TO_{s^0}^{s^2}), (B_{s^1}^{s^2}, TO_{s^1}^{s^2})] \\
& s^1[(B_{s^0}^{s^1}, TO_{s^0}^{s^1})].
\end{aligned}
\tag{1}
$$

For example, when node s^i receives such a setup message it process only the following line $s^i[(B_{s^0}^{s^i}, TO_{s^0}^{s^i}), (B_{s^1}^{s^i}, TO_{s^1}^{s^i}), \dots (B_{s^{i-1}}^{s^i}, TO_{s^{i-1}}^{s^i})]$ and based on this information, it then configures its insertion/extraction switches to receive these data, while transparently forwarding the rest. For this purpose, it uses the relative sizes and time-offset fields as described above. In the case that node s^i wishes to add new burst in the train it updates the message by adding the following fields: $(B_{s^{i+1}}^{s^{i+2}}, TO_{s^{i+1}}^{s^{i+2}})$, $(B_{s^{i+1}}^{s^{i+3}}, TO_{s^{i+1}}^{s^{i+3}}), \dots (B_{s^{i+1}}^{s^h}, TO_{s^{i+1}}^{s^h})$ to the corresponding lines of Eq. (1) The bursts are added at the end of burst train, provided there exists idle capacity that can be exploited.

4. Evaluation of the architecture

In order to evaluate the proposed hybrid switch architecture, we have developed a discrete-event network simulator based on the ns-2 platform. For reasons of simplicity, we will denote the one-way traffic as OBS traffic and the two-way traffic as OCS traffic, which can be of any kind i.e. packets, bursts etc. We have implemented the signaling scheme and the switch design detailed above and investigated the efficiency of the architecture. The prime target was to measure the efficient throughout that can be achieved in terms of bit- and burst-rate as well as the delay and burst loss ratio bounds. We have assumed that two-way OBS traffic independently requests and setups end-to-end lightpaths before transmission, while one-way traffic is transmitted without service guarantees (loss ratio and delay requirements) and is assembled into bursts and transmitted during the idle, pre-establishment time.

Performance of the proposed hybrid architecture depends on the network load, namely on the arrival rate of the lightpath requests and the number of wavelengths per fiber. To this end, in what follows, we first evaluate the performance for a specific lightpath request arrival rate and compare the different routing policies proposed, while in Section 4.2, we investigate the effect of the network load in terms of wavelengths per fiber and lightpath arrival rate. Performance evaluation was carried out on the NSF network topology, where all links were assumed to be bidirectional, with 10 Gbps capacity per wavelength. Full wavelength conversion was implemented, while the routing algorithm employed was fixed-shortest-path routing. The traffic between each source–destination pair was modeled with two separate traffic sources: (a) one that generates lightpath requests (denoted as two-way OCS traffic), according to a Poisson process with a mean of λ_{OCS} and exponential mean duration, $1/\mu$, equal to 100 ms, and (b) one that generates data bursts (one-way traffic) at a rate of λ_{OBS} bursts/s and a Pareto distributed size with 1MB mean value. The burst arrival rate was varied so as to perform measurements for various workloads, p according to:

$$\frac{1}{\lambda} = \frac{\text{burstsize } \times 8}{\text{access rate } \times p}.$$

It must be noted here that burst loss ratio in the core follows the blocking ratio of the lightpath requests, which has been selected 1%. To this end, if the SETUP message is blocked, the burst or the burst train is blocked as well. However, most of the one-way traffic is expected to be lost due to either buffer overflow at the edge or delay expiration, when one-way data wait for an appropriate setup message to arrive. In our experiment the buffer size was set to 1 GB per node, while the maximum delay tolerance of the bursts was set to 0.7 s. Destinations were evenly distributed for all types of traffic, while the opto-electronic (O/E) conversion and message processing delays were set equal to 10 ms. This delay also determines the minimum separation time of the bursts and the time offset needed. In the following analysis, it has been assumed that all wavelength channels can carry one-way traffic and thus each wavelength per fiber has its own insert/extract switches and relevant transceivers (see Fig. 2).

4.1. Burst loss ratio, delay and efficient throughput

Fig. 4a and b displays the burst loss ratio and average queuing time of the packets that comprise the bursts, versus burst arrival rate λ_{OBS}, in the case of 8 wavelengths/fiber and $\lambda_{OCS} = 30$ requests/s (two-way traffic). Results for all the three different policies are displayed. The combination of wavelength/fiber and lightpath request rate was chosen so as to yield a 1% blocking ratio. From Fig. 4 it can be seen that policy 1 is the worst performing one. It exhibits a higher loss ratio (17.5%) and delay than the other two. This is due to the fact that bursts are queued for a longer period of time before finding a matching setup message. Policies 2 and 3 perform better, providing a minimum burst loss ratio of 12% at the maximum load. However, their prime performance difference against policy 1 is packet queuing time, which is almost half. To this end, we may argue that all policies are capable of servicing same amount of traffic but with different service times, since policies 2 and 3 are capable of servicing more (and thus faster) bursts with a single SETUP message.

We have also measured the yielding throughput of all the policies. What is of interest to investigate is the efficient bit-rate over a specific source–destination pair as well as the actual burst sending rate of a single source towards all the destinations. Fig. 5a and b shows the corresponding results versus burst arrival rate λ_{OBS}. From Fig. 5a, it can be seen that policies 2 and 3 are capable of supporting, on average, a bit-rate of up to 270 Mbits per source–destination pair at the maximum traffic load (burst arrival rate $\lambda_{OBS} = 500$ burst/s), while policy 1 only 200 Mbps. Similarly, the burst transmission rates are 430, 380 and 320 bursts/s (see Fig. 5b). In other words, using policy 3, it is possible to

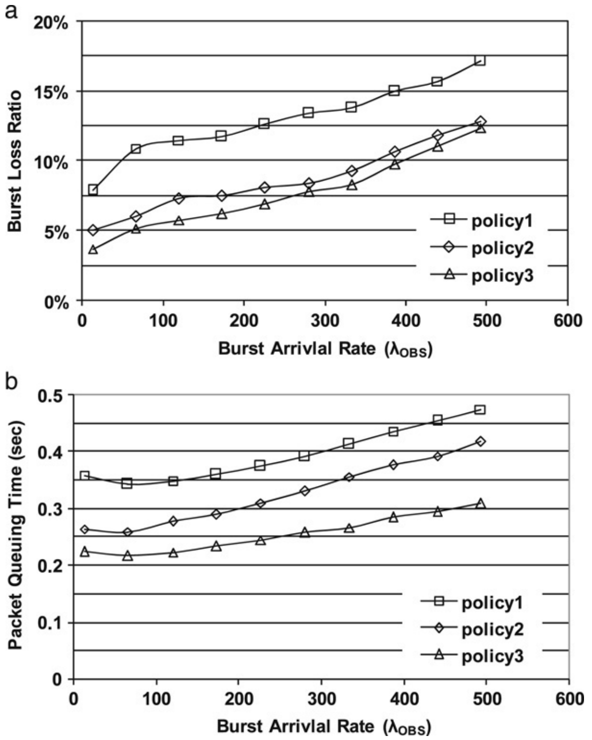

Fig. 4. (a) Burst loss ratio and (b) packet queuing time versus burst arrival rate for the three different policies defined.

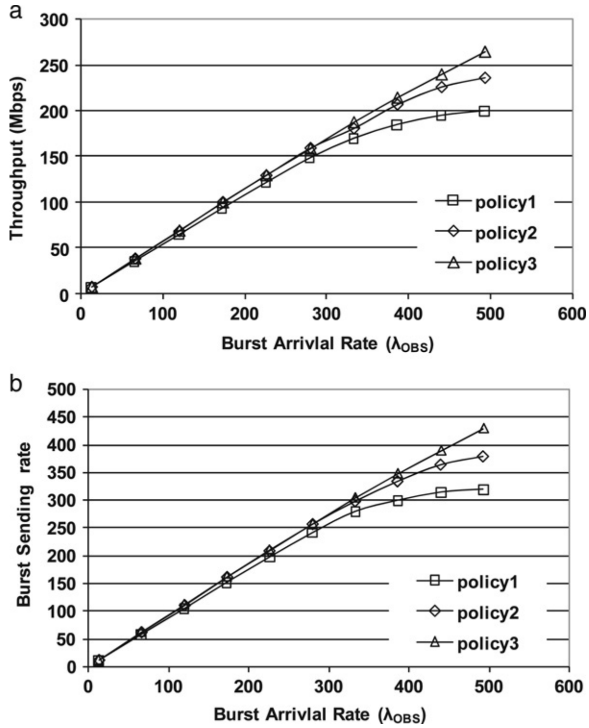

Fig. 5. (a) Efficient bit-rate over a specific source–destination pair and (b) Burst sending rate from a single source to all destinations for the three different policies defined.

salvage up to 86% of the 500 bursts arrived. However at such burst arrival rates, the corresponding loss ratio is high (17.5% for the first and 12.5%, for the second and third routing policy respectively) and could be unacceptable even for best effort traffic. Therefore, operation should be optimized at lower arrival rates. In particular, the burst sending ratio increases to a maximum of 99% (transmission of 12.8 bursts/s) for an arrival rate of $\lambda_{OBS} = 13$ burst/s (load $p = 0.1$).

The similar and higher performance of policies 2 and 3 can be attributed to the large data sizes that are transmitted per SETUP message. Intuitively policy 3 outperforms policy 2 only marginally, because the idle capacity is already highly utilized with bursts added by any node across the path, heading for the same end-destination with the setup message. Thus, adding bursts heading to intermediate nodes (policy 3) is redundant. To prove this concept, it was found that for policies 2 and 3, the mean burst-train size transmitted was 14 and 15 MB respectively, while for policy 1, only 10 MB. With the selected 1MB average burst size; these data correspond to 10, 14 and 15 bursts. These figures must be compared with the idle pre-transmission time per SETUP message, which for the NSF network topology correspond, on average for all source–destination pairs, to 32 MB of data. It is therefore clear that the proposed hybrid switch architecture is capable of utilizing the pre-transmission capacity up to 50%. To further strengthen this viewpoint, we have measured the number of bursts per SETUP message, for the three different policies. Figs. 6 and 7 show the cumulative and probability density function for a burst arrival rate of $\lambda_{OBS} = 500$ burst/s ($p = \sim 1$). It can be seen that 90% of the SETUP messages, in all cases carry up to 30 bursts. This is a significant amount of traffic and proves the advantageous use of the proposed architecture. Further, from Fig. 7, it can be seen how narrower the distribution of policy #1 is (see Fig. 7a), with respect to that of policy #3 (see Fig. 7b). In particular, probability density function of policy 3 bears a maximum of 61 MB data transmission with a probability of 0.4%. It must be noted here however, that upon selection of other mean burst size or other arrival process, burst size distribution will change, but in any case the yielding throughput will be the same. Figs. 6 and 7 only show how efficiently the idle, pre-transmission time can be utilized. In addition, albeit the exploitation of the idle capacity is on average 50%, and still there are bursts lost, this is because these two metrics depend indirectly to each other via the arrival rate and the destination distribution of the lightpath requests. In any case, utilization of the idle

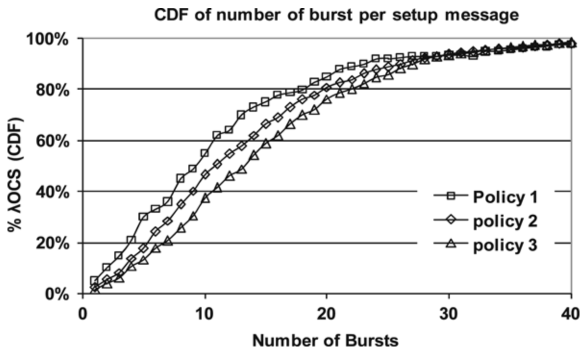

Fig. 6. Cumulative density function of number of bursts transmitted per SETUP message.

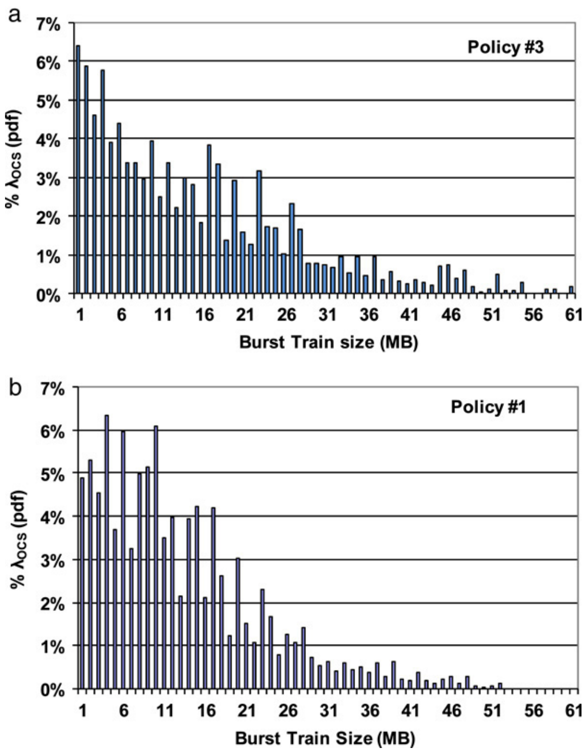

Fig. 7. Probability density function of the number of bursts transmitted per SETUP message for (a) policy #1 and (b) policy #3.

capacity can be further increased if one-way traffic is transmitted not only during message intervals but also during all channel's idle times as in [16].

4.2. Performance inter-dependencies

In what follows, we have investigated the effect of various network parameters such as the available number of wavelengths and the lightpath request rate, λ_{OCS}. Results shown refer only to policy 3, but may also apply to policy 2, due to their similar performance.

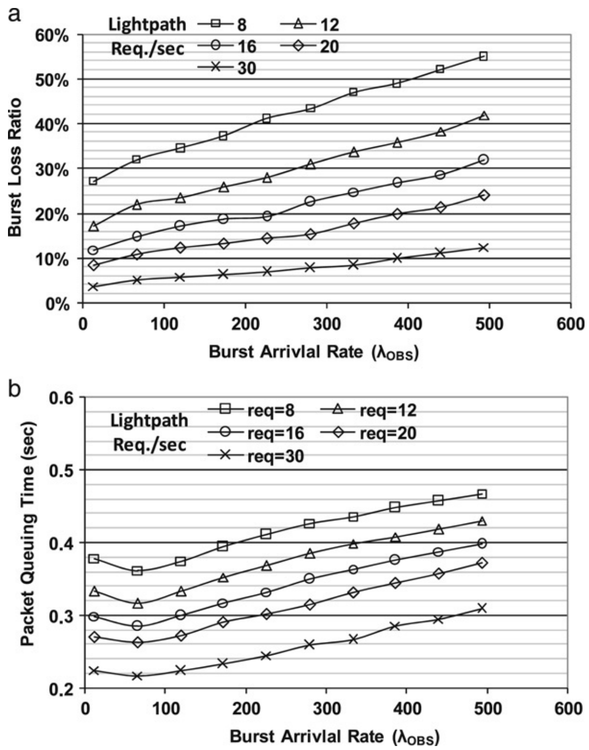

Fig. 8. (a) Burst loss ratio and (b) packet queuing time for a lightpath request rate of $\lambda_{OCS} = 8$, 12, 16, 20, and 30 and for WL = 8 wavelengths per link. Results shown here correspond to policy 3.

Fig. 8a and b shows the burst loss ratio and average packet delay respectively for $\lambda_{OCS} = 8$, 12, 16, 20 and 30 lightpath requests per second and for WL = 8. It can be seen that there is a significant increase in both the burst loss ratio and packet delay. In particular, loss ratio increases to 54%, while packet queuing increases to 0.46 s for $\lambda_{OBS} = 500$, when lightpath requests decrease from 30 to 8. The loss increase per λ_{OCS} decrease was found to be almost constant to 10%. This performance was expected, since the increase of the lightpath requests indirectly increases the number of idle SETUP periods, and thus increases the available capacity for one-way burst transmissions. It must be noted here that in this set of experiments, blocking of lightpath requests was not constant but was decreasing with the decrease of the arrival rate.

Finally, we have also measured the effect of the number of wavelengths (WL). Fig. 9 shows the corresponding results for WL = 2, 4, 6, 8 and 12 available wavelengths per link. It can be seen that the effect of WL is relatively weaker and this is because it affects performance indirectly through the acceptance (or not) of more lightpath requests. However, in such a case the blocking ratio of lightpath requests tends to zero, and thus the number of the successful SETUP

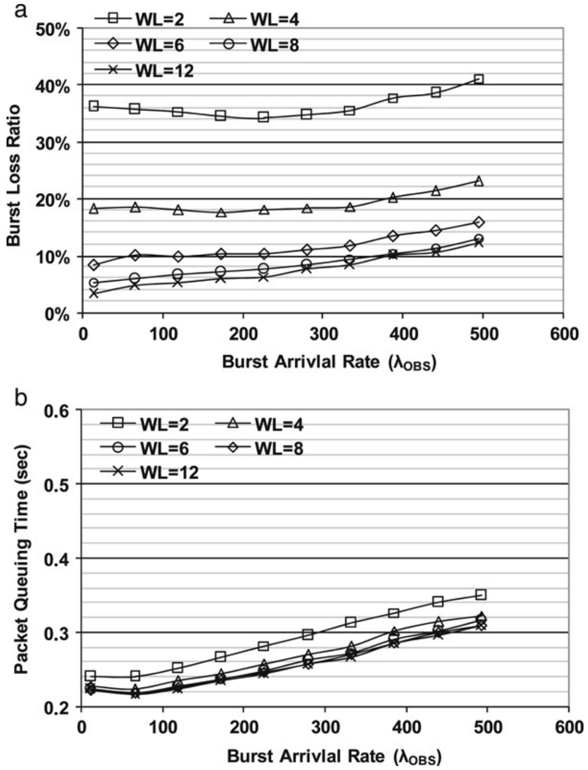

Fig. 9. (a) Burst loss ratio and (b) packet delay achieved for WL = 2, 4, 6, 8, and 12 wavelengths per link. Lightpath request rate, λ_{OCS}, is set equal to 30. Results shown here correspond to policy 3.

messages remains constant. To this end, burst loss ratio exhibits a higher gain when moving from WL $= 2 \rightarrow 4$, rather than when WL $= 4 \rightarrow 12$. This is clear from Fig. 9a, where we can observe that loss curves for WL = 6, 8 and 12 diverge slightly and only for burst arrival rates higher than $\lambda_{OBS} > 300$ (load $p > 0.8$). The same effect can be observed in the packet delay performance.

For the above set of experiments efficient throughput performance varies accordingly. To this end, burst sending rate and bit-rate per source–destination pair increase significantly with the increase of lightpath arrival rate (λ_{OCS}) and only slightly with the increase of wavelengths. Fig. 10a and b shows the corresponding results for $\lambda_{OBS} = 500\,(p = 1)$. From Fig. 10a, it can be seen that both metrics increase almost linearly, to reach a maximum of 270 Mpbs and 430 bursts/s respectively for $\lambda_{OCS} = 30$ requests/s. In contrast, when the number of wavelengths is changed from WL = 2 to 12, both curves increase fast when WL $= 2 \rightarrow 6$ but negligibly when WL $= 6 \rightarrow 12$ (see Fig. 10b).

Based on the above experimental results, we may conclude that the proposed hybrid architecture is capable of exploiting the unused capacity during the lightpath establishment process. For the NSF network topology this corresponds to an average transmission of 12 MB data per bandwidth request. This gain is significant and proves the economic viability of the design architecture. In addition, overall performance can be further improved, allowing burst transmission during any idle interval. However, it must be noted that performance of the scheme depends on the load of the two-way traffic, and in particular, it depends directly on the arrival rate of the lightpath requests and indirectly on the number of available wavelengths. We may argue that the proposed scheme can be used either to salvage traffic during temporal traffic overloads or operate independently for QoS differentiation. In the latter case, a constant bit-rate per source–destination can be guaranteed with certain upper bound of losses and queuing times.

5. Conclusions

In this paper, we have presented a new hybrid optical burst switch architecture and its main features. The proposed scheme combines two-way with one-way reservations by transmitting burst traffic during lightpath establishment in the idle pre-transmission times. In the proposed architecture, a new signaling message is defined to allow the lightpath SETUP message to carry routing and overhead information for the one-way traffic as well. The additional control plane overhead is relatively low, since only a tight, timely schedule of data traffic is needed, while network is alleviated from processing multiple control messages. Further, we have presented a detailed evaluation in order to assess the yielding throughput, burst loss ratio and average queuing delay. For this, we modeled in ns-2, a suitable node design, capable of handling data from two traffic generators and implemented the unified signaling. We have also proposed three different policies for associating bursts to signaling messages and evaluated their performance. We may argue that the proposed architecture is relatively simple to implement and requires few hardware enhancements. Thus, it can cost effectively integrate one-way OBS traffic into a circuit-switched network, allowing data channel to operate independently for inserting/extracting one-way traffic in/out the lightpaths. Its overall operation relies on the control plane that is responsible for traffic scheduling of both one-way and two-way traffic. The performance studies revealed that the proposed hybrid switching scheme can transmit a significant portion of traffic and can utilize capacity far more efficiently. It can support and sustain a constant bit-rate per

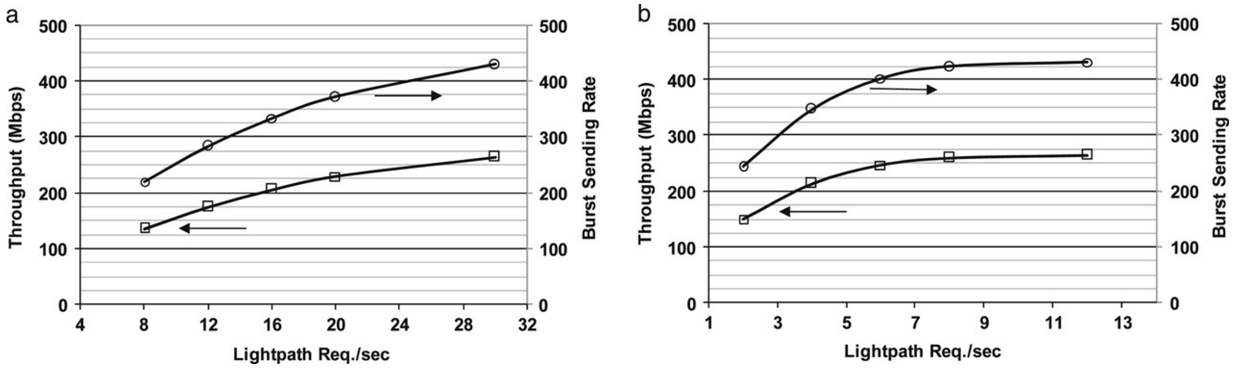

Fig. 10. Efficient bit-rate over a specific source–destination pair and burst sending rate from a single source to all destinations versus (a) lightpath arrival rate (λ_{OCS}) and (b) available number of wavelengths per link. Results shown correspond to policy 3 and to a burst arrival rate of $\lambda_{OBS} = 500$ ($p = 1$).

source–destination pair with a constant upper bound of delay.

Acknowledgments

This work has been supported by European Commission through the NoE E-Photon/ONe+ project via the Virtual Department on *Optical Switching Systems* (VD-S) and Joint Project on *Optical networking for grids and e-science* (JP-G).

References

[1] C. Qiao, M. Yoo, Optical burst switching (OBS) — a new paradigm for an optical Internet, Journal of High Speed Networks (JHSN) on WDM Networks 8 (1) (1999) 69–84 (special issue).

[2] M. Yoo, C. Qiao, S. Dixit, QoS performance of optical burst switching in IP-over-WDM networks, IEEE Journal on Selected Areas in Communications 18 (2000) 2062–2071.

[3] V. Vokkarane, J.P. Jue, Prioritized burst segmentation and composite burst-assembly techniques for qos support in optical burst-switched networks, IEEE Journal on Selected Areas in Communications 21 (7) (2003) 1198–1209.

[4] W. Liao, C. Loi, Providing service differentiation for optical-burst-switched networks, IEEE/OSA Journal of Lightwave Technology 22 (7) (2004) 1651–1660.

[5] Q. Zhang, V.M. Vokkarane, J.P. Jue, Biao Chen, Absolute QoS differentiation in optical burst-switched networks, IEEE Journal on Selected Areas in Communications 22 (9) (2004) 1781–1795.

[6] R. Karanam, V.M. Vokkarane, J.P. Jue, Intermediate node initiated (ini) signaling: a hybrid reservation technique for optical burst-switched networks, in: Proceedings, IEEE/OSA Optical Fiber Communication Conference 2003, Atlanta, GA, March 2003.

[7] K. Christodoulopoulos, K. Vlachos, L. Stampoulidis, E. Kehayas, E.A. Varvarigos, EBRP: A hybrid signaling protocol for efficient burst-level reservations and QoS differentiation in OBS networks, Journal of Optical Networking 5 (3) (2006) 147–158.

[8] M. Dueser, P. Bayvel, Analysis of a dynamically wavelength-routed optical burst switched network architecture, IEEE/OSA Journal of Lightwave Technology 20 (2002) 574–585.

[9] Hai Le Vu, A. Zalesky, E.W.M. Wong, Z. Rosberg, S.M.H. Bilgrami, M. Zukerman, R.S. Tucker, Scalable performance evaluation of a hybrid optical switch, IEEE/OSA Journal of Lightwave Technology 23 (10) (2005) 2961–2973.

[10] C. Xin, C. Qiao, Y. Ye, S. Dixit, A hybrid optical switching approach, in: Proc. IEEE Global Telecommunications (GLOBECOM) 7 Dec. 2003, pp. 3808–3812.

[11] G.M. Lee, B. Wydrowski, M. Zukerman, J.K. Choi, C.H. Foh, Performance evaluation of an optical hybrid switching system, in: Proc. IEEE Global Telecommunications, GLOBECOM, vol. 5, Dec. 2003, pp. 2508–2512.

[12] Hyeong-II Lee, Seung-Woo Seo, Hybrid optical transport Network (HOTNET): An optical network with hybrid switching technologies for integrated services, IEICE Transactions on Communications E88-B (10) (2005) 3874–3885.

[13] A. Gumaste, I. Chlamtac, Light-trails: An optical solution for IP transport, Journal of Optical Networking 3 (2004) 261–281.

[14] S. Balasubramanian, Ahmed Kamal, A.K. Somani, Medium access control protocols for light trail and light bus networks, in: Proc. of 8th IFIP Working Conference on Optical Network Design and Modeling, ONDM, Ghent, Belgium, 2–4 February 2004.

[15] S. Bjornstad, D.R. Hjelme, N. Stol, A highly efficient optical packet switching node design supporting guaranteed service, in: Proceedings of ECOC 2003, paper Mo 4.4.5, September 2003, pp. 110–111.

[16] E. Van Breusegem, J. Cheyns, D. de Winter, D. Colle, M. Pickavet, P. Demeester, J. Moreau, A broad view on overspill routing in optical networks: A real synthesis of packet and circuit switching? Elsevier Journal of Optical Switching and Networking 1 (2005) 51–64.

[17] K. Christodoulopoulos, E. Van Breusegem, K. Vlachos, M. Pickavet, E. Varvarigos, D. Colle, Performance evaluation of overspill routing in optical networks, IEEE International Conference on Communications 6 (2006) 2538–2543.