Available online at www.sciencedirect.com
ScienceDirect
Optical Switching and Networking 4 (2007) 20–32
Optical Switching and Networking
www.elsevier.com/locate/osn

Intermediate-node-initiation (INI): A generalized signaling framework for optical burst-switched networks☆

Vinod M. Vokkarane*

University of Massachusetts Dartmouth, Department of Computer and Information Science, North Dartmouth, MA 02747, United States

Received 15 June 2005; received in revised form 17 August 2006; accepted 22 September 2006
Available online 30 November 2006

Abstract

In this paper, we discuss different signaling techniques for optical burst-switched networks. We develop a generalized signaling framework for optical burst-switched networks, which provides guidelines about the performance of each signaling technique based on the different parameters in the framework. The two commonly used signaling techniques in optical burst switching are two-way based tell-and-wait (TAW) and one-way based just-enough-time (JET). TAW suffers from high end-to-end packet delay, while JET suffers from high packet loss. There is no signaling technique that offers flexibility in terms of both loss and delay. We propose a hybrid signaling technique called intermediate-node-initiated (INI) signaling for optical burst-switched networks. INI can provide different levels of loss and delay characteristics based on end-user application requirements. The granularity of INI ranges between the one-way based and the two-way based signaling techniques. In INI reservation of channels is initiated at an intermediate node, known as the initiating node, in both forward and backward directions at the same time. We show that by appropriately selecting the initiating node, we can simulate both TAW and JET using the INI signaling technique. Through simulations, we shown that INI performs better than TAW in terms of average end-to-end packet delay and better than JET in terms of burst loss probability. We extend the INI signaling technique to provide QoS differentiation in the OBS core, differentiated INI (DINI), by carefully choosing different initiation nodes depending on delay and loss requirements of end-user applications. Through extensive simulations, we show that the DINI technique outperforms the existing offset-based QoS technique.

Keywords: IP; WDM; OBS; QoS; Signaling

1. Introduction

To meet the explosive growth of the Internet and reduce costs, there has been a huge demand for higher transmission rates and faster switching technologies. IP over WDM is a promising framework that can support the bandwidth and flexibility requirement of next-generation networks. In order to efficiently utilize the amount of raw bandwidth in WDM networks, an all-optical transport method must be developed. IP over optical burst switching (OBS) is one such method for transporting traffic directly over a bufferless WDM network [1]. OBS has received an increasing amount of attention from industry and academia worldwide.

An OBS network consists of a collection of edge and core routers. The edge router assembles the electronic input packets into an optical burst which is sent over the OBS core. The ingress node aggregates incoming packets into bursts that are stored in the output buffer.

☆ Some portions of this paper was published in IEEE/OSA OFC'03, Atlanta, GA.

* Tel.: +1 508 910 6692; fax: +1 508 999 9144.
E-mail address: vvokkarane@ieee.org.

1573-4277/$ - see front matter
doi:10.1016/j.osn.2006.09.004

The assembled bursts are transmitted all-optically over OBS core routers without any storage at intermediate core nodes. The egress node, upon receiving the burst, disassembles the burst into packets and provides the packets to the upper layer. Basic architectures for core and edge routers in an OBS network have been studied in [2]. In an OBS network (Fig. 1), a data burst consisting of multiple IP packets is switched through the network all-optically. An out-of-band control header, known as the burst header packet (BHP) is transmitted ahead of the burst in order to configure the switches along the burst's route.

Several signaling techniques have been proposed for transmitting data all-optically in OBS networks. To accommodate the dynamic resource reservation requests, the source ingress node has to first compute a route to the destination egress node. Then, the signaling technique helps schedule data bursts on available wavelengths at each intermediate node along the route.

The most commonly studied distributed signaling techniques are tell-and-wait (TAW) and just-enough-time (JET). TAW is a two-way, acknowledgment-based signaling technique that uses immediate reservation and explicit release. JET is a one-way non-acknowledgment based signaling technique that uses delayed reservation and implicit release (see Section 2 for details). In order to implement all-optical data transfer, most OBS signaling techniques have an offset time duration between the BHP and the corresponding data burst. The offset time allows for the BHP to be processed at each intermediate node before the burst arrives at the intermediate node. The BHP may also specify the duration of the burst in order to let each intermediate node know when it may reconfigure its switch for the next arriving burst [3].

In an IP-over-OBS network, it is desirable to provide QoS support for various applications with diverse QoS demands, such as voice-over-IP, video-on-demand, and video conferencing. Several solutions have been proposed to support QoS in the OBS core network [4–8]. There is no single technique that offers flexibility to support both delay-sensitive and loss-sensitive traffic in the same OBS network.

In [4], an offset-based scheme was proposed. In this offset-based scheme, higher priority bursts are given a larger offset time than a lower priority burst. By providing a larger offset time, the probability of reserving the resources for the higher priority burst is increased, and thereby reducing higher priority burst loss. The limitations of the offset-based scheme are unfavorable end-to-end delay and unfair burst selection [5,9]. The offset-based scheme also suffers from high blocking probability of lower priority traffic.

In this paper, we propose a new flexible signaling technique called *intermediate-node-initiated (INI)* signaling, and also extend the proposed technique to provide QoS differentiation based on application requirements through *differentiated INI (DINI)* signaling. DINI provides loss and delay differentiation without explicitly introducing any additional offset time.

The remainder of the paper is organized as follows. Section 2 describes the generalized OBS signaling framework, and also discusses two specific OBS signaling protocols. Section 3 describes extensions to the generalized OBS signaling framework proposed earlier. Section 4 describes the proposed INI signaling technique. Section 5 describes the DINI signaling technique for providing QoS differentiation based on end-user application requirements. Section 6 analyzes the end-to-end delay equations for the different signaling techniques in OBS. Section 7 provides numerical results from simulation, and Section 8 concludes the paper.

2. Generalized OBS signaling framework

Signaling is a critical aspect that can significantly affect the performance of any network. For OBS networks, signaling is even more important, since the core is (usually) bufferless and any contention for resources during signaling can lead to data loss. In this section, we develop a generalized signaling framework, which can aid in the careful evaluation of all design parameters before opting for a particular signaling technique, given the requirements of the application data to be transmitted (see Fig. 2). We first explain the different design parameters that affect the performance of a signaling technique.

– *One-way or two-way:* The connection setup phase of a signaling technique can be either one-way or two-way. In one-way based signaling, the source sends out a BHP requesting the intermediate nodes along the path to allocate the necessary resources for the data burst. No acknowledgment message is sent back to the source notifying the success or failure of the resource reservation. The primary objective of one-way based signaling techniques is to minimizes the end-to-end data transfer delay. Unfortunately, this objective leads to high data loss due to contention of data bursts inside the OBS core.

  Two-way based signaling techniques are acknowledgment-based, where the request for a resource is sent from the source to the destination.

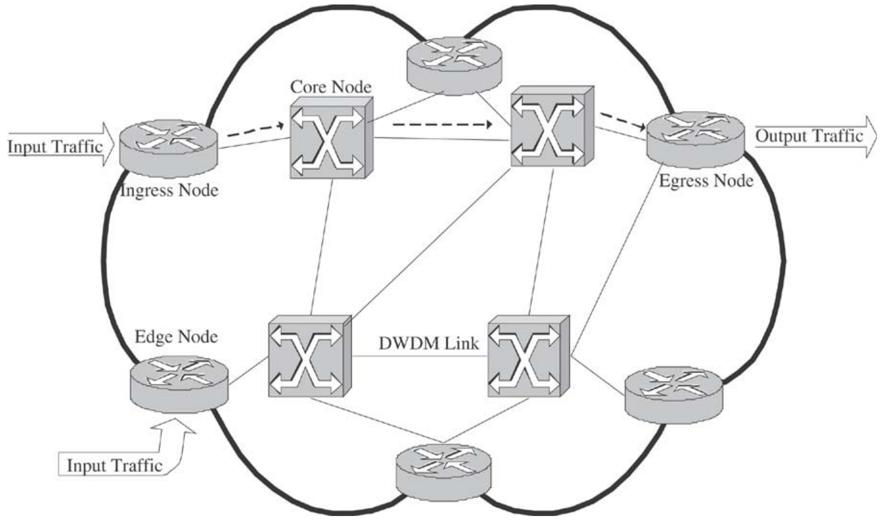

Fig. 1. OBS transport network.

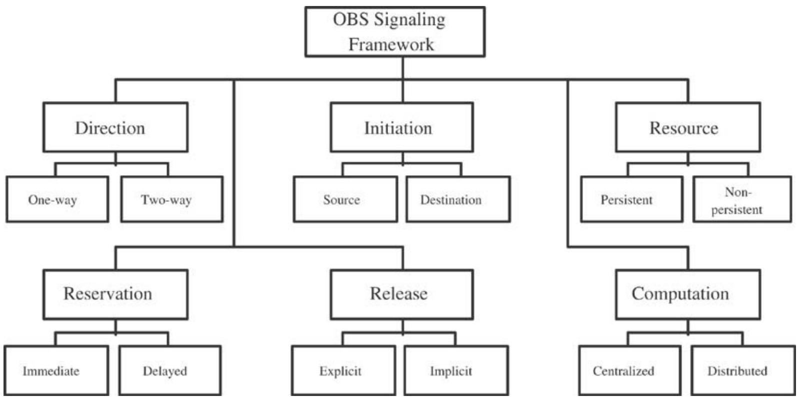

Fig. 2. Generalized OBS signaling framework.

The acknowledgment message confirming a successful assignment of requested resources is sent back from the destination to the source. The data burst is transmitted only after a connection is established successfully. If any intermediate nodes along the path is busy, then the request is blocked. That particular intermediate node takes suitable actions to release all the previously reserved links (if any), and also transmits a failure message back to the source. The source can choose to retry or drop the request. The primary objective of the two-way based technique is to minimize packet loss in the core network, but such an objective leads to high data transfer delay due to the round-trip connection setup.

– *Source initiation or destination initiation:* A signaling technique can initiate reserving the requested resources at the source or at the destination. In source initiation, resources are reserved in the forward path from the source to the destination. If the resource allocation is successful in the forward direction, an acknowledgment message containing the reserved wavelength may be sent back to the source. The source, upon receiving the resource confirmation, transmits the burst into the core network. In destination initiation, the source transmits a resource request to the destination node that collects wavelength availability information on every link along the path. Based on the collected information, the destination node will choose an available wavelength (if one exists), and sends a reservation request back to the source node, through the intermediate nodes, to reserve the chosen wavelength.

In general, source initiated techniques are greedy, in order to reduce the packet loss, the nodes in the forward direction may reserve more than the necessary wavelengths until the destination, and release the unnecessarily reserved wavelengths in the backward direction. This approach may lead to lower performance due to blocking of other requests due to *lack of resources*. On the other hand, destination initiated techniques first collect the wavelength availability information of all intermediate nodes and, based on that information, select a wavelength. In destination initiated techniques, a wavelength selected at the destination may be taken by some other request at any of the intermediate nodes along the path during the time when the status was collected and the time when the reservation message arrives at that node, also known as the *vulnerable period*. The primary cause of blocking (or data loss) in source initiation is due to the lack of free resources, while in destination initiation, the loss is due to outdated channel availability information stored at each core node [10,11].

– *Persistent or non-persistent:* One critical decision that each signaling technique needs to make is either to wait on a blocked resource (until it becomes free) or immediately indicate that there is a contention and initiate suitable connection failure mechanisms such as retransmission, deflection, or buffering [12]. In persistent signaling, the BHP waits on a blocked resource and assigns the wavelength when the resource becomes available. This approach leads to minimum loss, assuming that suitable buffers are provisioned at the nodes (edge and core) to store the incoming data bursts. In non-persistent signaling, the objective is to have an upper bound on the end-to-end data transfer delay, and hence each node declares the request to be a failure if the resource is not available immediately.
– *Immediate reservation or delayed reservation:* Based on when the reservation of a channel is started, the signaling techniques can support either immediate reservation or delayed reservation. In immediate reservation, the channel is reserved immediately from the instant the BHP is processed at a node. On the other hand, in delayed reservation, the channel is reserved from the actual arrival instant of the data burst at that node (or outgoing link). In order to employ delayed reservation, the BHP must carry the offset time between itself and its corresponding data burst. In general, immediate reservation is simple and practical to implement, but incurs higher blocking due to inefficient bandwidth allocation. On the other hand, implementation of delayed reservation is more involved, but leads to higher bandwidth utilization. Delayed reservation techniques also lead to the generation of idle voids between the scheduled bursts on the data channels. Scheduling algorithms used during reservation will need to store additional information about the voids. Based on that information, the scheduler must assign a wavelength to the reservation request. Delayed reservation and immediate reservation can be incorporated into any signaling technique, if the underlying node maintains the relevant information.

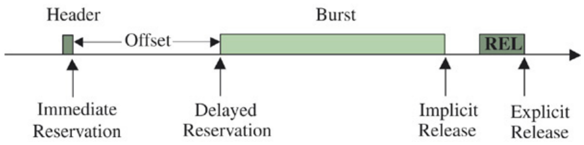

Fig. 3. Channel reservation and release mechanisms (for simplicity, BHP and data burst are shown to be on the same channel).

– *Explicit release or implicit release:* An existing reservation can be released in one of two ways, explicitly or implicitly. In explicit release, a separate control message is sent following the data burst, from the source towards the destination, in order to release (or terminate) an existing reservation. On the other hand, in implicit release, the BHP has to carry additional information such as burst length and offset time. We can see that the implicit release techniques results in better loss performance, due to the absence of any delay between the actual ending time of the burst and the arrival time of the release control message at each node. On the other hand, the explicit release technique results in lower bandwidth utilization and increased control messaging.

Based on the reservation and release mechanisms (Fig. 3), the signaling techniques can be categorized into four categories, *Immediate Reservation/Explicit Release*, *Immediate Reservation/Implicit Release*, *Delayed Reservation/Explicit Release*, and *Delayed Reservation/Implicit Release* [13,14]. Immediate reservation and explicit release indicates that an explicit control message is sent in order to perform the intended functionality, such as reserving a channel or releasing a connection. In delayed reservation, the BHP needs to carry the offset time, and in the case of implicit release, the duration of the data burst (in addition to offset time). We can easily observe from Fig. 3 that techniques employing delayed reservation and implicit release result in higher bandwidth utilization, while the techniques employing immediate reservation and explicit release are simple to implement at the expense of lower bandwidth utilization.

– *Centralized or distributed:* In centralized signaling, as proposed by [15], a dedicated centralized request server is responsible for setting up the route and assigning the wavelength on each route for every data burst for all source–destination pairs. The centralized technique may perform more efficiently when the network is small and the traffic is non-bursty. On the other hand, in distributed signaling, each node has a burst scheduler that assigns an outgoing channel for each arriving BHP in a distributed manner. The distributed approach is suitable for large optical networks and for bursty data traffic.

The objective of developing a generalized OBS signaling framework is that we can understand the performance of the signaling technique based on the parameters selected. Two prominent signaling techniques for an OBS network are Tell-and-Wait (TAW) and Just-Enough-Time (JET). In both of these techniques, a BHP is sent ahead of the data burst in order to configure the switches along the burst's route. We now describe these two signaling techniques.

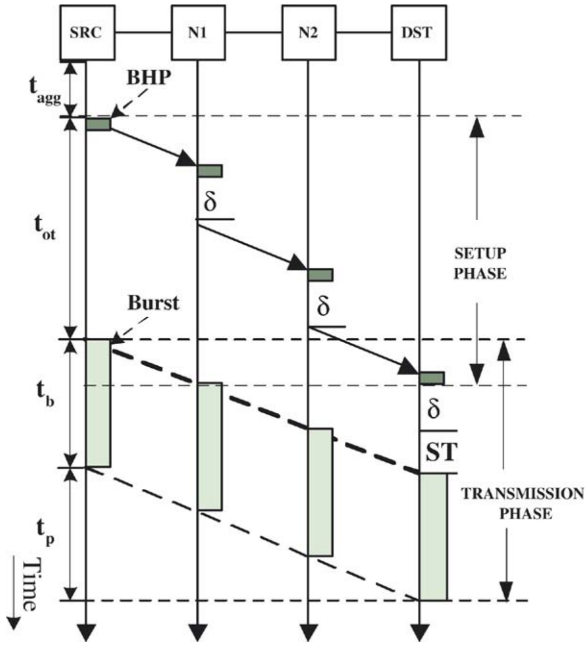

Fig. 4. Just-enough-time (JET) signaling technique.

2.1. Just-enough-time (JET)

Fig. 4 illustrates the JET signaling technique. As shown, a source node first sends a BHP on a control channel toward the destination node. The BHP is processed at each subsequent node in order to establish an all-optical data path for the corresponding data burst. If the reservation is successful, the switch will be configured prior to the burst's arrival. Meanwhile, the burst waits at the source in the electronic domain. After a predetermined offset time, the burst is sent optically on the chosen wavelength [1]. The offset time is calculated based on the number of hops from source to destination, and the switch reconfiguration time of a core node. Offset time is calculated as $\text{OT} = h \cdot \delta + \text{ST}$, where h is the number of hops between the source and the destination, δ is the per-node burst header processing time, and ST is the switching reconfiguration time. If at any intermediate node, the reservation is unsuccessful, the burst will be dropped. The unique feature of JET when compared to other one-way signaling mechanisms is delayed reservation and implicit release.

The information necessary to be maintained for each channel of each output port of every switch for JET consists of the starting and the finishing times of all scheduled bursts, which makes the system rather complex. On the other hand, JET is able to detect situations where no transmission conflict occurs, although the start time of a new burst may be earlier than the finishing time of an already accepted burst, i.e. a new burst can be scheduled in between two already reserved bursts. Hence, bursts can be accepted with a higher probability in JET.

Fig. 4 illustrates the different phases in JET. After a burst is assembled and the corresponding BHP is generated at the ingress node, we begin the setup phase. In the setup phase, the BHP attempts to set up an all-optical data path from the source to the destination along a pre-determined route (using delayed reservation). After the preset offset time, the transmission phase is initiated. In the transmission phase, the data burst cuts through the core network all-optically to reach the destination. Finally, resources at each node are immediately released after burst transmission (using implicit release).

At this point, it is important to note that there are other one-way based OBS signaling techniques, such as Just-In-Time (JIT) [16,17] and Tell-And-Go (TAG) [18]. JIT is similar to JET except that JIT employs immediate reservation and explicit release instead of delayed reservation and implicit release. Fig. 5(a) and (b) compares a similar signaling scenario using JET and JIT, respectively. An architectural framework for implementing various JIT schemes is presented in [19]. The primary benefit of using these one-way techniques is that the end-to-end delay is minimized for data transmission over an optical backbone network, at the cost of high data loss due to burst contentions for resources at the bufferless core network.

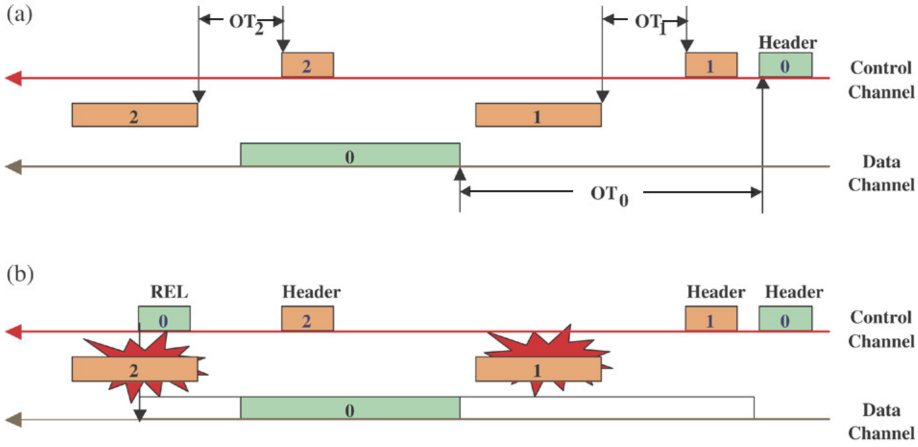

Fig. 5. Comparison of (a) JET and (b) JIT signaling.

In the TAG approach, the data burst must be delayed at each node in order to allow time for the BHP to be processed and for the switch to be configured, instead of pre-determining this duration at the source and incorporating the delay in the offset time. In TAG, each core node has input FDLs to provide the necessary burst header processing delay to the bursts in the data plane. TAG employs immediate reservation and implicit release.

2.2. Tell-and-wait (TAW)

Fig. 6 illustrates the four phases of the TAW signaling technique. In the setup phase, a BHP is sent along the burst's route to collect channel availability information at every node along the path. At the destination, a channel assignment algorithm is executed, and the reservation period on each link is determined based on the earliest available channel times of all the intermediate nodes. Next, in the confirm phase, a BHP is sent in the reverse direction (from destination to source) to reserve the channel at each intermediate node. If the BHP reaches the source successfully, then the burst is sent into the core network at the beginning of the transmission phase. At any node along the path, if the required channel is already occupied, a BHP is sent toward to the destination to release all the previously reserved resources. Finally, in the release phase, based on the information contained in the initial BHP during the setup phase, an implicit/explicit release is implemented after data transmission.

TAW is similar to wavelength-routed networks, in the sense that the channel can be reserved in the forward direction as in source initiated reservation (SIR) or in the reverse direction from the destination back to the source as in destination initiated reservation (DIR) [11, 10]. TAW in OBS is different from wavelength-routed WDM networks in the sense that in TAW resources are reserved at any node only for the duration of the burst. Also, if the duration of the burst is known during reservation, then an implicit release scheme can be followed to maximize bandwidth utilization.

In comparison, the primary disadvantage of TAW is the high round-trip setup time; however the data loss is very low. Therefore, TAW is predominantly suited for loss-sensitive and delay-tolerant traffic. On the other hand, the primary disadvantage of JET is the high data loss, however the end-to-end data transfer delay is minimal. TAW takes approximately three times the one-way propagation delay from source to destination for the burst to reach destination, whereas JET takes the sum of the one one-way propagation delay and an offset time. Therefore, JET (or JIT or TAG) is predominantly suited for delay-sensitive and loss-tolerant traffic. Hence, in OBS, there is no signaling technique that offers flexibility in terms of both delay and loss.

3. Generalized OBS signaling framework extensions

In this section, we extend the generalized signaling framework to include certain hybrid reservation schemes. By using the signaling framework, we can carefully evaluate various design parameters before opting for a particular signaling technique, given the requirements of the data to be transmitted. The following are the additions to the generalized signaling framework as shown in Fig. 7.

– *Hybrid (part two-way and part one-way) direction:* In the hybrid signaling technique, the signaling is two-way from the source to the initiating node

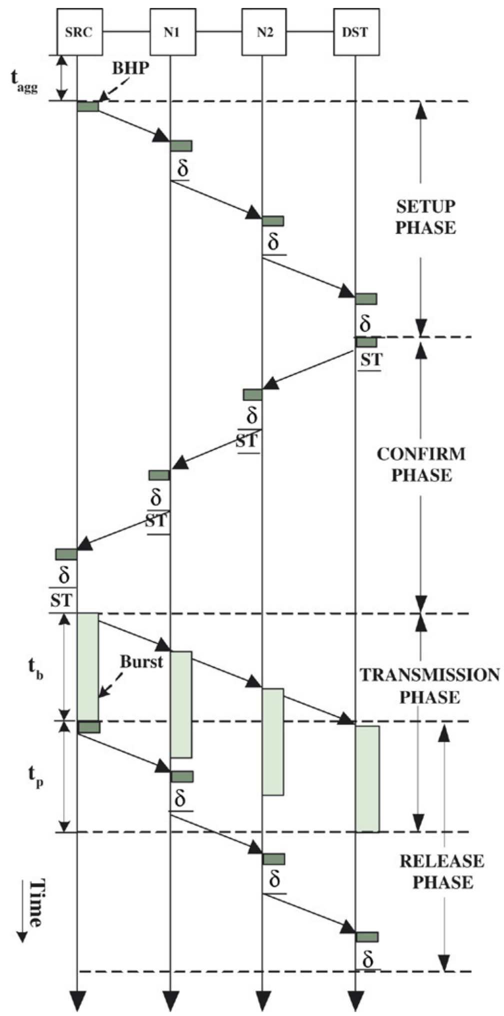

Fig. 6. Tell-and-wait (TAW) signaling technique.

(IN), and one-way from the initiating node to the destination. If the initiating node is closer to the source, performance is similar to pure one-way based techniques, such as JET, and if the initiating node is closer to the destination, performance is similar to pure two-way based techniques, such as TAW. Based on the position of the initiating node, different loss and delay characteristics can be obtained.

– *Intermediate initiation:* In intermediate initiation, typically the resources are reserved similar to destination initiation technique from the source to the intermediate node, and similar to source initiation technique from the intermediate node to the destination.

The remaining parameters in the generalized signaling framework such as *Resource, Reservation, Release*, and *Computation* remain the same.

In the next section, we describe a new OBS signaling technique called *intermediate-node-initiated (INI)* signaling, which captures the advantages of both TAW and JET, and supports flexible delay and loss application requirements. The reservation request is initiated at an intermediate node, called the initiating node (IN). In the first part of the path, i.e., from source to the initiating node, the INI signaling technique works with an acknowledgment similar to TAW. In the later part of the path, i.e., from the initiating node to destination, the INI signaling technique works without an acknowledgment similar to JET.

4. Intermediate-node-initiated (INI) signaling

To overcome the limitations of TAW and JET, we propose the intermediate-node-initiated signaling technique. In the INI signaling technique, a node between source and destination along the path is selected as the initiating node. An initiating node is an intermediate node between the source and the destination at which a channel reservation algorithm is executed to determine the earliest time that the burst

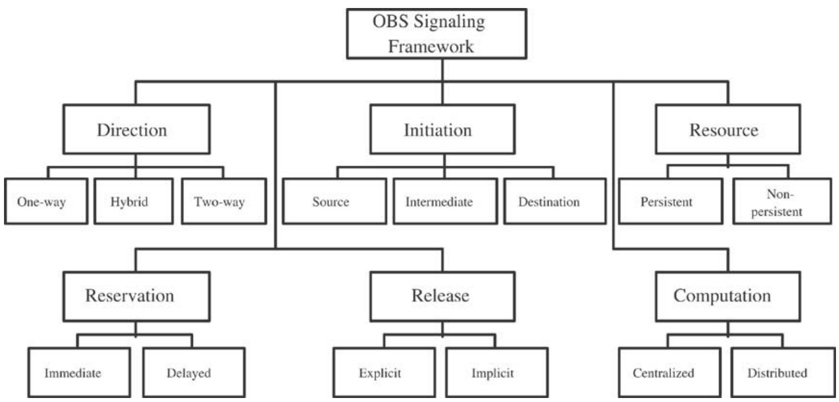

Fig. 7. Generalized OBS signaling framework with INI extensions.

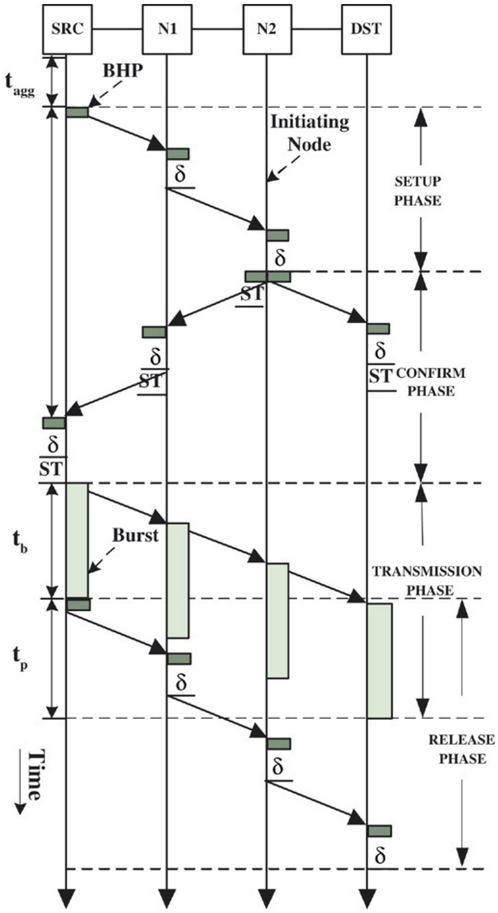

Fig. 8. Intermediate node initiated (INI) signaling technique.

can be sent from the source node and the corresponding earliest times at which the nodes between the source and the initiating node can be scheduled to receive the burst. At the initiating node, the actual reservation of the channels starts in both directions, i.e., from the initiating node to the source as well as from the initiating node to the destination. The selection of the initiating node is critical in INI.

Fig. 8 illustrates the different phases of the INI signaling technique. In the setup phase, when a burst is created at the edge node, a BHP containing the destination as well as the initiation node (IN) is sent through the OBS core to the IN. The BHP collects the details of channels at every node along the path until it reaches the initiating node. Next, in the confirm phase, a channel assignment algorithm is executed at the initiating node to determine the time duration that the channels will need to be reserved at each intermediate hop between the source and initiating node. A confirm BHP is then sent to the source node, which reserves channels along the path from the initiating node to the source. The IN simultaneously sends another BHP towards the destination, for reserving the channels between the IN and the destination. If a channel is busy at any node between the IN and the source, a release BHP is sent back to the initiating node to release any previously reserved resources. If the confirming BHP reaches the source successfully, then the transmission phase is initiated where the burst is sent into the core at the scheduled time. If, at any node between the initiating node and the destination node, the BHP fails to reserve the channel, the burst is dropped at that node.

We note that all the nodes along the routing path in the OBS network should support INI signaling. Based on the location of the IN, every node along the path will have a pre-defined role in the signaling process. The nodes before the IN (upstream) have to implement signaling functions similar to TAW, the nodes after the IN (downstream) have to implement signaling functions similar to JET/JIT, and the IN has to implement signaling functions similar to both TAW and JET/JIT.

In TAW, there is an acknowledgment from the destination before the burst is sent from the source, and in JET, there is no acknowledgment. In INI, there is an acknowledgment from the initiating node, thereby decreasing the probability of blocking compared to JET. Also, since the burst waits at the source for a time less than the roundtrip propagation delay from the source to the destination, INI decreases the end-to-end delay compared to TAW. In the INI signaling technique, if the initiating node is set to be the source node, then the signaling technique is identical to JET, and if the initiating node is set to be the destination node, then the signaling technique is identical to TAW. For the INI signaling technique, TAW and JET are the two extremes, so by appropriately selecting the initiating node, we can implement TAW and JET by using INI. Also, note that in INI we can use both regular reservation and delayed reservation. In general, as we have discussed in earlier sections, with delayed reservation the signaling technique has improved performance. In our simulations, we used delayed reservation based INI signaling.

Table 1 compares the OBS signaling techniques in terms of their signaling parameters, end-to-end delay, and loss probability.

Illustration: Consider the path 2–4–5–7 in Fig. 9, with Node 2 as the source and Node 7 as the destination. Here we have four possible initiating nodes including the source and destination nodes. If we choose the source, i.e. Node 2, as the initiating node, then the INI signaling technique emulates JET. If we choose

Table 1
Comparison of OBS signaling techniques

Signaling	Direction	Initiation	Reservation	Release	Delay	Loss
TAW	Two-way	Src./Dest.	Immediate	Explicit	High	Low
TAG	One-way	Source	Immediate	Implicit	Low	High
JET	One-way	Source	Delayed	Implicit	Low	High
JIT	One-way	Source	Immediate	Explicit	Low	High
INI-RR	Hybrid	Intermediate	Immediate	Explicit	Flexible	Flexible
INI-DR	Hybrid	Intermediate	Delayed	Implicit	Flexible	Flexible

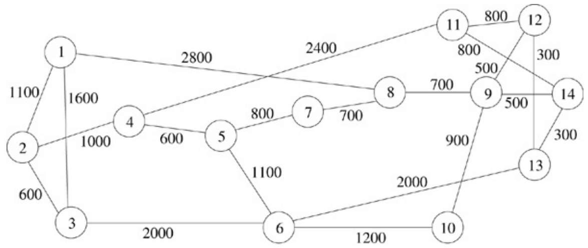

Fig. 9. 14-node NSF backbone network topology (distance in km).

the destination, i.e. Node 7, as the initiating node, then the INI signaling technique emulates TAW. Other possibilities of initiating nodes are Node 4 and Node 5. Let us consider Node 5 to be the initiating node and observe how the INI signaling technique works. Node 2 sends the BHP to the next hop, Node 4, along with the channel availability information of the Link 2–4. Node 4 adds the channel availability information of Link 4–5 and forwards the BHP to the next node, Node 5. When the initiating node, Node 5 gets the BHP, it runs a channel reservation algorithm to determine the earliest times at which the burst can be scheduled on the intermediate nodes along the path between source and initiating node. A BHP that reserves the channels at the intermediate nodes at the pre-determined time instances is sent from the initiating node to the source. As soon as the reply BHP reaches the Source 2, the burst is transmitted into the core. The BHP sent from the initiating node (Node 5) to the destination reaches Node 7 and configures Node 7 to receive the incoming burst at the appropriate time.

5. Differentiated intermediate node initiated (DINI) signaling

The INI signaling technique can be extended to provide QoS at the optical layer. In theory, it is possible to implement multiple signaling techniques in the same network to provide differentiated services in order to support both loss and delay sensitive traffic, i.e., we can use TAW for loss sensitive traffic, and JET for delay sensitive traffic. This approach of having a hybrid core network with two (or more) different signaling schemes can only provide a coarse QoS guarantee. In order to provide a finer level of QoS differentiation, we modify the INI scheme.

Using INI, we can satisfy both the loss and delay constraints of each specific application by carefully selecting the initiating node. In general, for applications with delay constraints we need to choose the initiating node to be closer to the source node, such that the end-to-end delay is less than the application-specified constraint. For applications with loss constraints, we need to choose the initiating node to be closer to the destination node, such that the majority of the path is two-way acknowledged.

Suppose we have to support three classes of traffic, say P1, P2, and P3, with P1 being delay sensitive, P2 being both delay and loss sensitive, and P3 being loss sensitive. We can use the source node as the initiating node for P1, the center node as the initiating node for P2, and the destination node as the initiating node for P3, thus providing differentiated services in the same OBS network.

At this point, we would like to mention that the primary limitation of DINI is that the maximum number of service classes that can be simultaneously supported is directly proportional to the hop-length of the route between each source–destination pair. To overcome this limitation, DINI can be implemented in conjunction with any of the other existing OBS QoS techniques, such as offset-based [4], prioritized burst segmentation [6], and early-drop and wavelength grouping [7,8]. We do not address these topics in this paper; these topics would be candidates for future work.

6. End-to-end delay analysis

In this section, we develop analytical equations for evaluating the delay characteristics of each OBS signaling technique. Without loss of generality, we investigate a network with a single wavelength per fiber. Our model can be directly extended to a network with multiple wavelengths per fiber. Due to the absence of wavelength converters, multiple wavelengths in each fiber can be thought of as multiple layers of the network, with one layer for each wavelength. We also assume that no optical buffering (FDLs) is supported at core nodes. In the following analysis, we ignore the delay incurred in BHP creation, collecting channel availability information at each node, and the execution of the channel selection algorithm. We define the following notation:

- R_{sd}: route from Source s to Destination d.
- t_{bhp}: burst header packet (BHP) processing delay at each OBS node (core and edge). We assume that the processing delays of different types of BHP at all the nodes is identical. t_{bhp} is in the hundreds of ns range.
- t_{sw}: switching time required to reconfigure the optical cross-connect at each OBS node. t_{sw} is in the tens of μs range [20,21].
- t_{agg}: burst aggregation delay based on the assembly technique adopted at the ingress OBS node.
- t_b: data burst transmission time.
- t_{ot}: offset time, the fixed initial time between the BHP and the data burst at the ingress node.
- t_p^{ij}: propagation delay on the fiber link between nodes i and j. t_p^{ij} is 5 μs/km.

We first calculate the average end-to-end packet delay, T_{SIG}, incurred by each signaling technique. T_{SIG} is the duration from the instant the first packet arrives at the ingress node to the instant the burst is completely received at the destination and the connection is completely released (if applicable). Consider a route R_{sd} with h hops to the destination.

(a) *Tell-and-wait (TAW)*

In TAW, the end-to-end delay is given by the sum of the burst aggregation time, the round trip connection setup time, the burst transmission time, and the data burst propagation time. The initial BHP collects the channel information from each core link along the path to the destination. At the destination, the channel assignment is computed and the reservation of resources is initiated on the reverse direction. Hence, the round trip connection setup time is the sum of the round trip propagation delay, the BHP processing time at each of the nodes, and one switching time (refer Fig. 6). The data transmission phase incurs a one-way propagation delay between the source and the destination plus the burst transmission time. Finally, the release phase incurs the BHP processing time at each of the nodes. Therefore, the end-to-end packet delay of TAW, is given by:

$$T_{\text{TAW}} = t_{\text{agg}} + t_b + t_{\text{sw}} + 3 \sum_{l^{ij} \in R_{sd}}^{h} (t_p^{ij} + t_{\text{bhp}}). \tag{1}$$

(b) *Just-enough-time (JET)/just-in-time (JIT)/tell-and-go (TAG)*

In JET, the end-to-end delay is given by the sum of the burst aggregation time, the offset time, the burst transmission time, and the data burst propagation time (release time not applicable).

$$T_{\text{JET}} = t_{\text{agg}} + t_{\text{ot}} + t_b + \sum_{l^{ij} \in R_{sd}}^{h} t_p^{ij}, \tag{2}$$

where,

$$t_{\text{ot}} = h\, t_{\text{bhp}} + t_{\text{sw}}. \tag{3}$$

In JIT, there is no preset offset time between the BHP and the data burst. Also, additional delay is incurred during the explicit release phase as compared to JET. Therefore, the end-to-end delay is given by:

$$T_{\text{JIT}} = t_{\text{agg}} + t_b + t_{\text{sw}} + \sum_{l^{ij} \in R_{sd}}^{h} (t_p^{ij} + 2\, t_{\text{bhp}}). \tag{4}$$

In TAG, at each node, the input FDL provides the necessary BHP processing delay and switch reconfiguration delay. Hence, the end-to-end delay is same as Eq. (2) (JET) with $t_{\text{ot}} = 0$. Instead, the equivalent delay is provided by the FDLs at each node along the path.

(c) *Intermediate-node-initiated (INI)*

In INI, the end-to-end delay is given using a combination of the delay equation of TAW and JET. The end-to-end delay in INI depends upon the location of the initiation node (IN), the burst aggregation time, the burst transmission time, and the data burst propagation time. Let l is the number of hops between the source node and IN, and m is the number of hops between IN and the destination node.

$$T_{\text{INI}} = t_{\text{agg}} + t_b + t_{\text{sw}} + 3 \sum_{l^{ij} \in R_{si}}^{l} (t_p^{ij} + t_{\text{bhp}}) + \sum_{l^{ij} \in R_{sd}}^{m} (t_p^{ij} + t_{\text{bhp}}). \tag{5}$$

If we implement INI with regular reservation, the nodes from the IN to the destination node can immediately reserve data channels for each reservation request. On the other hand, if we implement INI

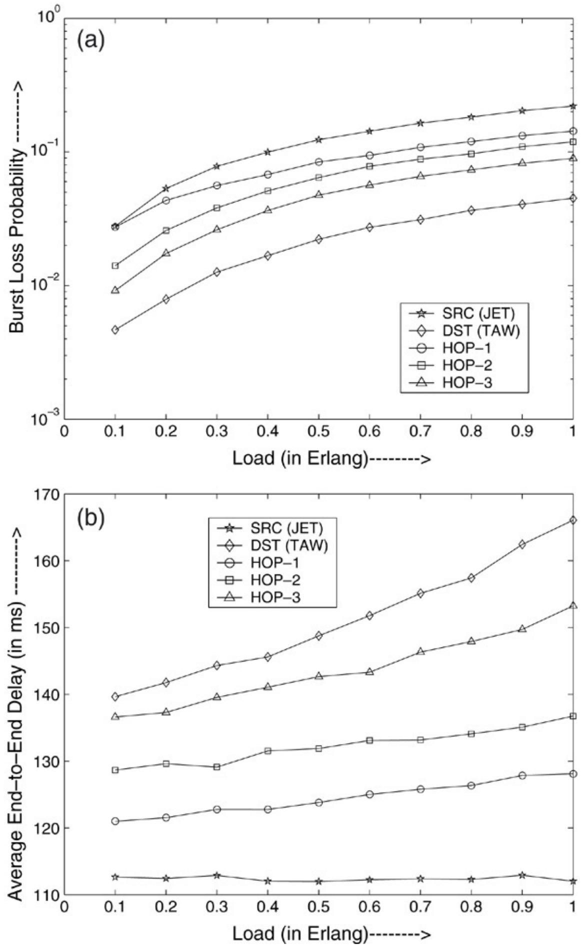

Fig. 10. (a) Burst loss probability versus load, and (b) average end-to-end delay versus load, when the initiating nodes are source, first-hop, second-hop, third-hop, and destination.

with delayed reservation, the nodes from the IN to the destination node should have accurate information about offset time and length of the burst. The offset time for INI with DR is given by,

$$t_{\mathrm{ot}} = t_{\mathrm{sw}} + \sum_{l^{ij} \in R_{si}}^{l} (2\,t_p^{ij} + t_{\mathrm{bhp}}). \tag{6}$$

We can easily observe that when $l = h$ ($m = 0$), then delay is same as TAW (Eq. (1)), and if $l = 0$ ($m = h$), then delay same as JET (Eq. (2)). In conclusion,

$$T_{\mathrm{JET}} \le T_{\mathrm{INI}} \le T_{\mathrm{TAW}}. \tag{7}$$

7. Numerical results

In order to evaluate the performance of the new INI and DINI signaling techniques, a simulation model is developed. Burst arrivals into the network are assumed to be Poisson with an exponentially distributed burst length. The average burst length is set to 0.1 ms. The link transmission rate is 10 Gb/s. Each arriving packet is 1250 bytes in length. The switching reconfiguration time is 0.01 ms. There is no FDL buffering or wavelength conversion at any core node. Retransmission of the lost bursts is not considered [22]. Fig. 9 shows the 14-node NSF network topology on which the simulation is implemented.

Fig. 10(a) and (b) plot the burst loss probability and average end-to-end delay versus load when the initiating nodes are taken as source (SRC), first-hop (Hop-1), second-hop (Hop-2), third-hop (Hop-3), and destination (DST) respectively. In Fig. 10(a) and (b), only paths that are more than or equal to four hops are considered to show the effect of the INI signaling technique. We observe that as expected the loss probability decreases as the initiating node moves away from the source. If the initiating node is chosen to be closer to the source, a greater part of the path is unacknowledged leading to higher data loss. On the other hand, if the initiating node is chosen to be closer to the destination, a greater part of the path is acknowledged leading to lower data loss. We also observe that the delay increases proportionally to the increase in distance between the initiating node and the source, since the path from the source to the initiating node is acknowledged, and hence incurs a higher round-trip delay. Also, the values of loss and delay when the initiating node is at the source and the destination are consistent with JET and TAW respectively. Similarly, Fig. 11(a) and (b) plot the burst loss probability and average end-to-end delay versus load for JET, INI (center hop), and TAW. We observe that the results are consistent with the previous plots.

Fig. 12(a) and (b) plot the burst loss probability and average end-to-end delay versus load for the three priority traffic. We observe that P3 suffers the least loss, while P1 incurs the least delay, and P2 experiences loss and delay between the values of P1 and P3. For comparable values of offset time, we found that INI outperforms the traditional offset-based QoS scheme [4]. In the offset-based scheme, the source has to estimate the additional-offset to provide differentiated services, while in INI, the initiating node has the channel availability information of all nodes between itself and the source. Also, in INI the data burst does not enter the network until resources have been reserved between the source node and the initiating node.

8. Conclusion

In this paper, we developed a generalized signaling framework for optical burst-switched networks. Each of

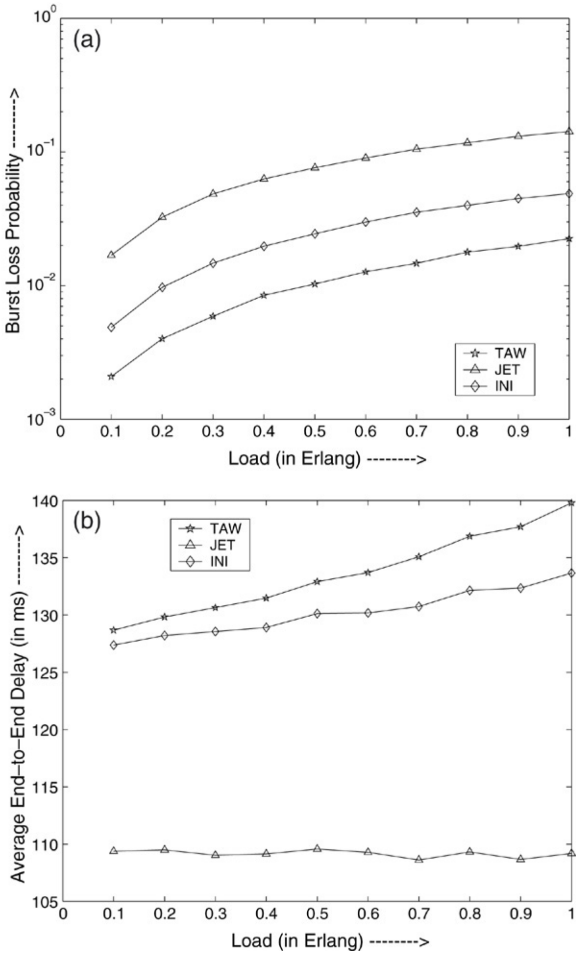

Fig. 11. (a) Burst loss probability versus load, and (b) average end-to-end delay versus load, for JET, TAW, and INI with the initiating node at the center hop.

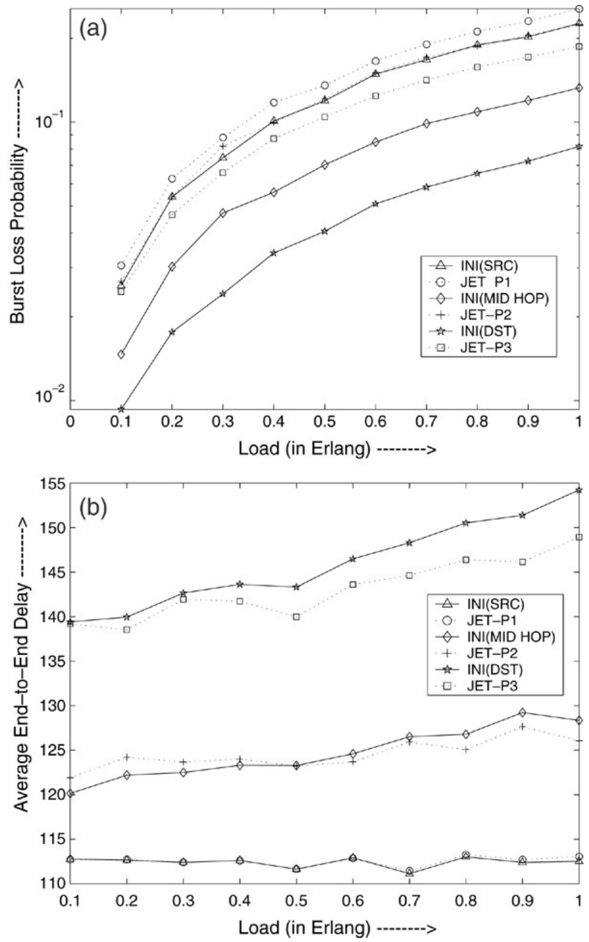

Fig. 12. Burst loss probability versus load, and (b) average end-to-end delay versus load, when the initiating nodes are source, middle-hop and destination in the same network to provide differentiation through signaling.

the signaling parameters are thoroughly discussed and evaluated. TAW and JET are discussed in detail, and the advantages and disadvantages of each technique are also discussed. We identify that there is a significant void in the OBS literature for providing flexible signaling. We introduced a new and flexible OBS signaling technique called intermediate-node-initiated (INI) signaling for an OBS network. The INI signaling technique provides flexibility during channel reservation based on the requirements of application data to be transmitted. We described the working principle of the INI signaling technique, and its advantages over existing techniques like TAW and JET. Through extensive simulations and delay analysis, we show that the packet loss probability of INI is less than that of JET and the end-to-end data transfer delay is less than that of TAW. Hence, the proposed hybrid technique is a flexible solution suitable for handling the varying traffic demands on the next-generation optical network.

We extend the INI signaling technique to provide QoS differentiation in the OBS core, differentiated intermediate-node-initiated signaling (DINI), by carefully choosing different initiation nodes depending on delay and loss requirements of the end-user application. We also illustrate how the DINI signaling technique can be used to provide QoS and validate the performance of the technique through simulations. For comparable values of offset-time, we found that DINI out-performs the traditional offset-based QoS technique.

An area of future work is to study the performance of the INI signaling technique with wavelength conversion in a multi-wavelength network and combine it with deflection of BHPs during signaling to improve channel utilization and to monitor the delay trade-off. Also, DINI can be extended to work in conjunction with other OBS QoS techniques, such as offset-based [4], prioritized burst segmentation [6], and early-drop and wavelength grouping [7,8].

Acknowledgments

The author would like to thank Mr. Ravikiran Karanam (UT Dallas) for performing certain simulations and Prof. Jason Jue (UT Dallas) for several useful discussions on the topic.

References

[1] C. Qiao, M. Yoo, Optical burst switching (OBS) — a new paradigm for an optical internet, Journal of High Speed Networks 8 (1) (1999) 69–84.
[2] Y. Xiong, M. Vanderhoute, H.C. Cankaya, Control architecture in optical burst-switched WDM networks, IEEE Journal on Selected Areas in Communications 18 (10) (2000) 1838–1854.
[3] C. Qiao, M. Yoo, Choices, features and issues in optical burst switching, SPIE Optical Networks Magazine 1 (2) (2000) 36–44.
[4] 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 (10) (2000) 2062–2071.
[5] Y. Chen, M. Hamdi, D.H.K. Tsang, Proportional QoS over OBS network, in: Proceedings, IEEE Globecom, vol. 3, November 2001, pp. 1510–1514.
[6] V.M. 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.
[7] Q. Zhang, V.M. Vokkarane, B. Chen, J.P. Jue, Absolute QoS differentiation in optical burst-switched networks, IEEE Journal on Selected Areas in Communications 22 (9) (2004) 1781–1795.
[8] C.-H. Loi, W. Liao, D.-N. Yang, Service differentiation in optical burst switched networks, in: Proceedings, IEEE Globecom, vol. 3, November 2002, pp. 2313–2317.
[9] F. Poppe, K. Laevens, H. Michiel, S. Molenaar, Quality-of-service differentiation and fairness in optical burst-switched networks, in: Proceedings, SPIE OptiComm, vol. 4874, July 2002, pp. 118–124.
[10] K. Lu, G. Xiao, I. Chlamtac, Analysis of blocking probability for distributed lightpath establishment in WDM optical networks, IEEE/ACM Transactions on Networking (2004).
[11] X. Yuan, R. Melhem, R. Gupta, Y. Mei, C. Qiao, Distributed control protocols for wavelength reservation and their performance evaluation, Photonic Networks and Communications 1 (3) (1999) 207–218.
[12] S. Yao, B. Mukherjee, S.J.B. Yoo, S. Dixit, A unified study of contention–resolution schemes in optical packet-switched networks, IEEE/OSA Journal of Lightwave Technology (March) (2003).
[13] L. Xu, H.G. Perros, G. Rouskas, Techniques for optical packet switching and optical burst switching, IEEE Communications Magazine 39 (1) (2001) 136–142.
[14] T. Battestilli, H. Perros, An introduction to optical burst switching, IEEE Communications Magazine 41 (8) (2003) 510–515.
[15] M. Dueser, P. Bayvel, Analysis of a dynamically wavelength-routed optical burst switched network architecture, IEEE/OSA Journal of Lightwave Technology 20 (4) (2002) 574–586.
[16] J.Y. Wei, J.L. Pastor, R.S. Ramamurthy, Y. Tsai, Just-in-time optical burst switching for multi-wavelength networks, in: Proceedings, IFIP TC6 International Conference on Broadband Communications, November 1999, pp. 339–352.
[17] A.H. Zaim, I. Baldine, M. Cassada, G.N. Rouskas, H.G. Perros, D. Stevenson, The JumpStart just-in-time signaling protocol: A formal description using EFSM, Optical Engineering 42 (2) (2003) 568–585.
[18] A. Detti, M. Listanti, Application of tell and go and tell and wait reservation strategies in a optical burst switching network: A performance comparison, in: Proceedings, IEEE International Conference on Telecommunications, ICT 2001, June 2001.
[19] I. Baldine, G.N. Rouskas, H.G. Perros, D. Stevenson, Jumpstart: A just-in-time signaling architecture for WDM burst-switched networks, IEEE Communications Magazine 40 (2) (2002) 82–89.
[20] L. Rau, S. Rangarajan, D.J. Blumenthal, H.-F. Chou, Y.-J. Chiu, J.E. Bowers, Two-hop all-optical label swapping with variable length 80 Gb/s packets and 10 Gb/s labels using nonlinear fiber wavelength converters, unicast/multicast output and a single EAM for 80- to 10 Gb/s packet demultiplexing, in: Proceedings, Optical Fiber Communication Conference, OFC, March 2002, pp. FD2–1–FD2–3.
[21] A. Neukermans, R. Ramaswami, MEMs technology for optical networking applications, IEEE Communications Magazine 39 (1) (2001) 62–69.
[22] Q. Zhang, V.M. Vokkarane, Y. Wang, J.P. Jue, Evaluation of burst retransmission in optical burst-switched networks, in: Proceedings, IEEE Broadnets 2005, Optical Networking Symposium, October 2005.