Absolute QoS Differentiation in Optical Burst-Switched Networks

Qiong Zhang, *Student Member, IEEE*, Vinod M. Vokkarane, *Member, IEEE*, Jason P. Jue, *Senior Member, IEEE*, and Biao Chen, *Member, IEEE*

***Abstract*—A number of schemes have been proposed for providing quality-of-service (QoS) differentiation in optical burst-switched (OBS) networks. Most existing schemes are based on a relative QoS model in which the service requirements for a given class of traffic are defined relative to the service requirements of another class of traffic. In this paper, we propose an absolute QoS model in OBS networks which ensures that the loss probability of the guaranteed traffic does not exceed a certain value. We describe two mechanisms for providing loss guarantees at OBS core nodes: an early dropping mechanism, which probabilistically drops the nonguaranteed traffic, and a wavelength grouping mechanism, which provisions necessary wavelengths for the guaranteed traffic. It is shown that integrating these two mechanisms outperforms the stand-alone schemes in providing loss guarantees, as well as reducing the loss experienced by the nonguaranteed traffic. We also discuss admission control and resource provisioning for OBS networks, and propose a path clustering technique to further improve the network-wide loss performance. We develop analytical loss models for the proposed schemes and verify the results by simulation.**

***Index Terms*—Absolute quality-of-service (QoS), dense wavelength-division multiplexing (DWDM), optical burst switching, relative QoS.**

I. Introduction

THE explosive growth of the Internet demands a high-speed transmission technology for supporting rapidly increasing bandwidth requirements. Currently, dense wavelength-division multiplexing (DWDM) technology enables the multiplexing of 160–320 wavelengths into a single fiber, with a transmission rate of 10–40 Gb/s per wavelength. In order to efficiently utilize the raw bandwidth in DWDM networks, an all-optical transport method, which supports fast resource provisioning and asynchronous transmission, must be developed. Optical burst switching (OBS) is a promising bufferless DWDM switching technology that can potentially provide high wavelength utilization. OBS employs a signaling technique in which an out-of-band burst header packet (BHP) is first sent to reserve resources and configure network elements along the path of the data burst. After an offset time, the data burst is transmitted all-optically through the network. Offset time is necessary in order to configure the intermediate nodes before the data burst arrives. One example of such a signaling technique is just-enough-time (JET) [1].

An important issue in OBS is the scheduling of data bursts onto data channels at every core node. The scheduling algorithm must quickly and efficiently find an available outgoing data channel for each incoming burst in a manner which minimizes data loss. Several scheduling algorithms have been proposed for burst scheduling at an OBS node. These algorithms include Horizon or latest available unscheduled channel (LAUC) [2], LAUC-VF [3], Min-SV [4], and segmentation-based scheduling [5]. The most simple and practical scheduling algorithm is the LAUC algorithm. In LAUC, the node keeps track of the time at which the channel is scheduled to be available on every outgoing data channel. The arriving burst is scheduled onto an available channel, so that the gap between the ending time of the currently scheduled burst on that channel and the starting time of the arriving burst is minimum.

During scheduling, an arriving burst may contend with one or more scheduled bursts on the outgoing data channels. This contention results in the burst being dropped, leading to burst loss. In OBS, there are several contention resolution schemes that aim to minimize burst loss. The primary contention resolution schemes include wavelength conversion [6], [7], fiber delay line buffering [8], [9], deflection [10], and segmentation [12], [13].

Quality-of-service (QoS) support is another important issue in OBS networks. There are two models for QoS: *relative QoS* and *absolute QoS*. In the relative QoS model, the performance of each class is not defined quantitatively in absolute terms. Instead, the QoS of one class is defined relatively in comparison to other classes. For example, a high-priority class is guaranteed to experience lower loss probability than a low-priority class. However, the loss probability of the high-priority class still depends on the traffic load of the low-priority class; and no upper bound on the loss probability is guaranteed for the high-priority class.

The absolute QoS model provides a bound for loss probability of the guaranteed traffic. This kind of hard guarantee is essential to support applications with delay and bandwidth constraints, such as multimedia and mission-critical applications. Moreover, from the ISP's point of view, the absolute QoS model is preferred in order to ensure that each user receives an expected level of performance. Efficient admission control and resource provisioning mechanisms are needed to support the absolute QoS model.

Manuscript received November 4, 2003; revised August 2, 2004. This work was supported in part by the National Science Foundation (NSF) under Grant ANI-01-33899. Portions of this paper have appeared in IEEE HPSR'03, Torino, Italy, June 2003, in IEEE GLOBECOM'03, San Francisco, CA, Dec. 2003, and in IEEE GLOBECOM'04, Dallas, TX, Dec. 2004.

Q. Zhang, J. P. Jue, and B. Chen are with the Department of Computer Science, The University of Texas at Dallas, Richardson, TX 75083 USA (e-mail: qzhang77@utdallas.edu; jjue@utdallas.edu; bchen@utdallas.edu).

V. M. Vokkarane is with the Department of Computer and Information Science, University of Massachusetts Dartmouth, North Dartmouth, MA 02747 USA (e-mail: vvokkarane@umassd.edu).

Digital Object Identifier 10.1109/JSAC.2004.836289

QoS differentiation schemes may focus on providing loss differentiation, delay differentiation, or bandwidth guarantees. In OBS networks, bursts follow an all-optical path from source to destination. Thus, the delay incurred from source to destination is primarily due to propagation delay, and bandwidth guarantee is implicitly provided by supporting loss guarantee. Hence, the focus of QoS support in OBS networks is to provide loss differentiation.

In IP networks, many queueing disciplines have been developed in order to provide QoS differentiation. Priority queueing (PQ) is a relative differentiation scheme that stores the packets in prioritized queues at each hop, and the packets are scheduled onto an output port only if all packet queues of higher priority are empty. Weighted fair queueing [14] computes virtual finishing time for each packet at the head of each session queue, and transmits the packet with the smallest virtual finishing time. Weighted fair queueing can provide absolute QoS differentiation in the sense that it is able to guarantee a predictable amount of bandwidth and a maximum delay bound for a specific session. On the other hand, a proportional QoS differentiation model was proposed in [15] and [16] in order to provide relative QoS differentiation. Using this model, the relative QoS differentiation is refined and quantified in terms of queueing delay and packet loss probability. Further, in [17], a *dynamic class selection* framework was proposed to provide absolute QoS in which the proportional QoS differentiation approach controls the QoS spacing of each class at every hop, and the users dynamically search for an appropriate class to meet their absolute requirements. In [18], the authors gave an overview of recent research on the proportional QoS differentiation model for various QoS metrics, and proposed buffer management schemes for achieving absolute service bounds in the proportional QoS differentiation approach.

In OBS networks, several schemes have been proposed to support the relative QoS model. In [19], an extra-offset-based scheme that provides relative loss differentiation was proposed. In this extra-offset-based reservation scheme, higher priority class bursts are given a larger offset time than the lower priority class bursts. By providing a larger offset time, the probability of reserving the resources for the higher priority class bursts is increased, and therefore, the loss probability experienced by higher priority class bursts is decreased. The limitations of the extra-offset-based scheme are unfavorable end-to-end delay and unfairness [20], [21].

In [21], a proportional QoS scheme based on per-hop information was proposed to support burst loss probability and delay differentiation. The proportional QoS model quantitatively adjusts the QoS metric to be proportional to the differentiation factor of every traffic class. If p_i is the loss metric and s_i is the differentiation factor for Class i, then using the proportional differentiation model, the following will hold for every traffic class:

$$\frac{p_i}{p_j} = \frac{s_i}{s_j}. \tag{1}$$

In order to implement this model, each core node needs to maintain traffic statistics, such as the number of burst arrivals and the number of bursts dropped for every traffic class. Hence, the online loss probability of Class i, p_i, is the ratio of the number of Class i bursts dropped to the number of Class i burst arrivals during a fixed time interval. To maintain the differentiation factor between the traffic classes, an intentional burst dropping scheme is employed.

In [22], proportional QoS differentiation is provided by maintaining the number of wavelengths occupied by each class of bursts. Every arriving burst is scheduled based on a usage profile maintained at every node. Arriving bursts that satisfy their usage profiles preempt scheduled bursts that do not satisfy their usage profiles, so as to maintain the preset differentiation ratio.

Relative QoS differentiation schemes do not provide a hard guarantee for any of the supported QoS metrics, thus absolute QoS differentiation schemes are necessary. An intuitive approach to provide absolute QoS differentiation is to design a hybrid optical backbone network consisting of wavelength-routed lightpaths [23] to carry the guaranteed traffic, and a classical OBS network to carry the nonguaranteed traffic. This approach leads to inefficient usage of bandwidth over the wavelength-routed part of the network. In order to efficiently utilize bandwidth, we need to develop absolute QoS differentiation schemes in which all wavelengths in the network are available for statistical multiplexing and dynamic bandwidth allocation.

The primary objective of this paper is to develop mechanisms that are able to provide a guaranteed loss probability for the guaranteed traffic while also reducing the loss probability experienced by the nonguaranteed traffic. To our knowledge, no previous work has considered providing absolute loss guarantees in OBS networks. We propose two mechanisms to achieve the goal at a per-hop level, namely, *early dropping* and *wavelength grouping*. The early dropping mechanism probabilistically drops the bursts of lower priority class in order to guarantee the loss probability of higher priority class traffic. The wavelength grouping mechanism provisions wavelengths for the guaranteed traffic and schedules the bursts based on provisioning. The integration of these two mechanisms gives a very effective solution for providing absolute loss guarantee, while also significantly reducing the loss probability experienced by the nonguaranteed traffic.

To implement the per-hop mechanisms over an entire network, we propose a *path clustering* technique that prioritizes bursts based on the path hop-distance, so that the loss probability experienced by the nonguaranteed traffic is reduced, while providing end-to-end absolute loss for the guaranteed traffic. In this paper, our schemes assume that the signaling protocol is JET and the burst scheduling algorithm is LAUC. We also assume that bursts of Class i have higher priority than bursts of Class j, where $i < j$.

The paper is organized as follows. Section II discusses the network architecture to support the end-to-end absolute QoS model. Sections III and IV describe the early dropping and wavelength grouping mechanisms. Section V explains the integration of early dropping and wavelength grouping. Section VI describes the path clustering technique. The analytical loss models for the integrated scheme of early dropping and dynamic wavelength grouping as well as for the path clustering technique are developed in Section VII. Section VIII

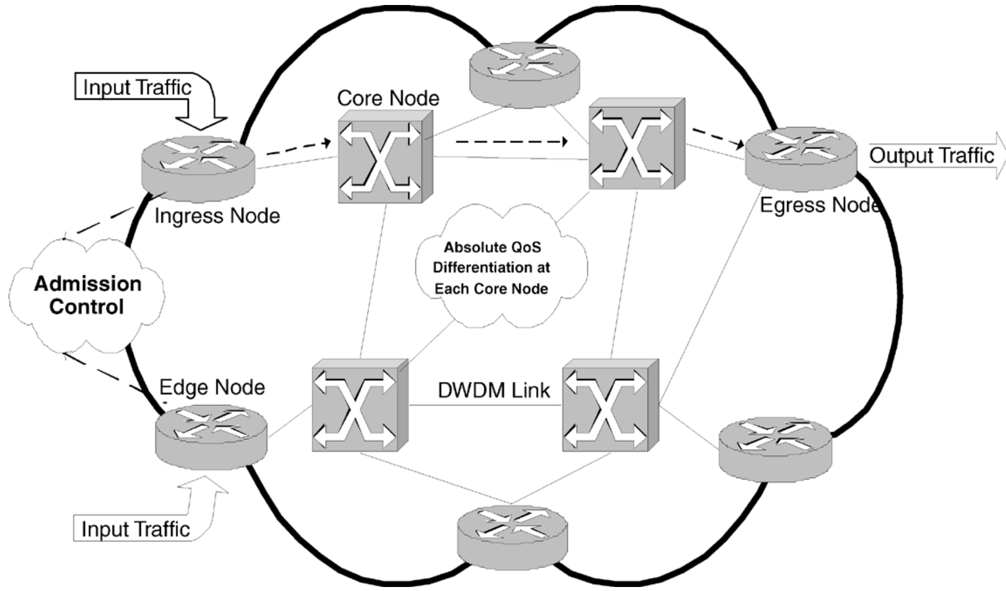


Fig. 1. OBS transport network.

discusses the performance of the proposed schemes based on the analytical model as well as simulation results. Section IX concludes the paper.

II. Absolute QoS Network Architecture

An OBS transport network consists of a collection of edge and core nodes as shown in Fig. 1. The input traffic from multiple clients is assembled at the ingress node and is transmitted as bursts through high-capacity DWDM links over the optical core. An egress node, upon receiving a burst, disassembles the burst, and delivers the data packets to the corresponding clients.

Absolute QoS differentiation relies on proper resource provisioning and admission control. One simple resource provisioning technique is to allocate resources for the traffic of each service class based on its QoS requirements. In this technique, each service class i is assumed to require a maximum network-wide loss guarantee, $P_{C_i}^{\mathrm{NET}}$. Given that each OBS node maintains the same loss guarantee, $P_{C_i}^{\mathrm{MAX}}$ for Class i traffic, we can calculate $P_{C_i}^{\mathrm{MAX}}$ at each node from the diameter of the network, D, and $P_{C_i}^{\mathrm{NET}}$ as follows:

$$P_{C_i}^{\mathrm{MAX}} = 1 - e^{ln\left(1-P_{C_i}^{\mathrm{NET}}\right)/D}. \tag{2}$$

Therefore, if the actual loss probability is guaranteed to be less than $P_{C_i}^{\mathrm{MAX}}$ at each node along the path, then the network-wide loss probability $P_{C_i}^{\mathrm{NET}}$ is guaranteed end-to-end.

In OBS networks, admission control can be implemented only at the edges nodes, since the edge nodes have the capability to electronically buffer incoming traffic and the core nodes do not have any buffers. Therefore, the maximum arrival rate between every source-destination pair can be controlled at the edge node during burst assembly. We assume that the burst arrivals in the OBS network follow a Poisson process. Based on the maximum arrival rate of the guaranteed traffic, the routing algorithm, and the network topology, we can obtain the maximum offered load of the guaranteed traffic on every link. For every link, let L_{C_i} be the maximum offered load of Class i traffic, and let W_{C_i} be the minimum number of wavelengths required in order to guarantee that the loss probability of Class i traffic is below $P_{C_i}^{\mathrm{MAX}}$. We can compute W_{C_i} for the guaranteed traffic of Class i using the standard Erlang-B formula

$$\frac{\frac{L_{C_i}^{W_{C_i}}}{W_{C_i}!}}{\sum_{n=0}^{W_{C_i}} \frac{L_{C_i}^{n}}{n!}} \le P_{C_i}^{\mathrm{MAX}}. \tag{3}$$

Hence, in order to guarantee the maximum end-to-end loss, each core node must provide at least W_{C_i} wavelengths and must guarantee the maximum per-hop loss probability, $P_{C_i}^{\mathrm{MAX}}$, for each Class i traffic. Note that (3) can apply to any burst length distribution. If the arrival process is not Poisson, then another method would be required to determine the minimum number of wavelengths needed to guarantee that the loss probability is below $P_{C_i}^{\mathrm{MAX}}$.

III. Early Dropping Mechanism

In this section, we present an early dropping mechanism that guarantees absolute loss probability for the higher priority class of traffic by intentionally dropping bursts of lower priority class. This mechanism is similar to the concept of random early detection (RED) gateways for congestion avoidance in packet-switched networks. In RED, the gateway detects congestion by computing the average queue size. When the average queue size exceeds a preset threshold, the gateway drops arriving packets with a certain probability, where the exact probability is a function of the average queue size [24]. Due to the bufferless nature of the OBS core nodes, the early dropping mechanism computes the intentional dropping probability based on measured online loss probability rather than the queue size.

In this mechanism, an *early dropping probability*, $p_{C_i}^{\mathrm{ED}}$, is computed for each Class i based on the online loss probability

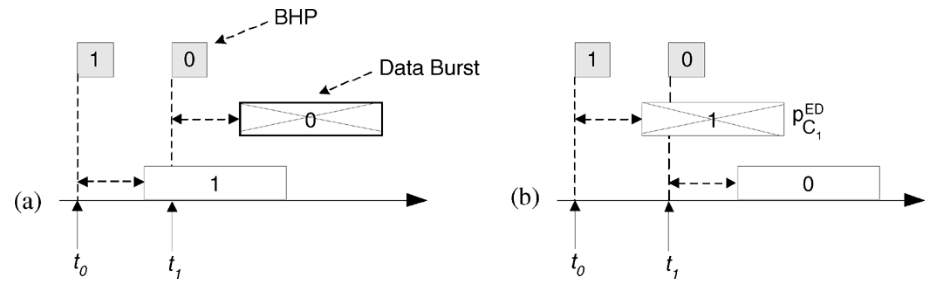


Fig. 2. (a) Standard dropping mechanism, and (b) early dropping mechanism.

and the maximum acceptable loss probability of the immediately-higher priority class. An *early dropping flag*, e_i, is associated with each Class i. e_i is determined by generating a random number between 0 and 1. If the random number is less than $p^{\mathrm{ED}}_{C_i}$, then e_i is set to 1, otherwise it is set to 0. Hence, e_i is 1 with probability $p^{ED}_{C_i}$, and is 0 with probability $(1 - p^{\mathrm{ED}}_{C_i})$. In order to determine whether or not to drop an arriving Class i burst, not only do we need to consider the early dropping flag of Class i, but also the early dropping flags of all higher priority classes. Thus, we generate an *early dropping vector*, ED_i, where $\mathrm{ED}_i = \{e_1, e_2, \ldots, e_i\}$ for the arriving Class i burst. The Class i burst is intentionally dropped if $e_1 \vee e_2 \vee \cdots \vee e_i = 1$, that is, the Class i burst is intentionally dropped with probability $(1 - \prod_{j=1}^{i} (1 - p^{\mathrm{ED}}_{C_j}))$. Note that we do not have an element e_0 for Class 0, since Class 0 has the highest priority.

Let us consider a two-class example to illustrate the early dropping concept. In Fig. 2(a), the BHP of a Class 1 burst (low priority) arrives at time t_0 and reserves the channel. The BHP of a Class 0 burst (high priority) arrives at time t_1, where $t_1 > t_0$, and contends with the Class 1 burst, resulting in the Class 0 burst being dropped. In order to reduce the likelihood of this scenario, a burst of Class 1 is intentionally dropped when $e_1 = 1$, prior to the BHP arrival of the Class 0 burst [Fig. 2(b)]. e_1 is set to 1 with probability $p^{\mathrm{ED}}_{C_1}$; The key is to decide when to trigger the early dropping mechanism, and how to compute the early dropping probability.

In order to measure the online loss probability, each OBS core node must monitor the traffic statistics for each guaranteed class. For each output port of an OBS node, let a_{C_i} be the burst arrival counter, and let d_{C_i} be the burst drop counter. We calculate $p_{C_i} = d_{C_i}/a_{C_i}$, as the online loss probability for Class i traffic. For this purpose, a_{C_i} and d_{C_i} can be measured within a fixed time window.

Note that the early dropping mechanism needs wavelength provisioning in order to meet the loss requirement of the guaranteed traffic. If there are not enough wavelengths for providing absolute loss probability for the guaranteed traffic, then even by dropping all of the nonguaranteed traffic, the loss requirement of the guaranteed traffic can not be satisfied.

We now describe the following *early drop by threshold* and *early drop by span* schemes to compute the early dropping probability, $p^{\mathrm{ED}}_{C_i}$, for Class i bursts.

A. Early Drop by Threshold (EDT)

The basic idea of early drop by threshold (EDT) is to drop the arriving Class i bursts when the online loss probability of Class $(i-1)$, $p_{C_{i-1}}$ reaches the maximum acceptable loss probability, $P^{\mathrm{MAX}}_{C_{i-1}}$. This early dropping of bursts of the lower priority classes is a simple way to provide loss guarantee for the higher priority class. The early dropping probability of Class i bursts is given by

$$p^{\mathrm{ED}}_{C_i} = \begin{cases} 0, & p_{C_{i-1}} < P^{\mathrm{MAX}}_{C_{i-1}} \\ 1, & p_{C_{i-1}} \geq P^{\mathrm{MAX}}_{C_{i-1}} \end{cases} \tag{4}$$

where $i \geq 1$.

In the EDT scheme, bursts of each class with lower priority than Class $(i-1)$, suffer from high loss probability when $p_{C_{i-1}}$ exceeds $P^{\mathrm{MAX}}_{C_{i-1}}$. Since there is a single trigger point, the scheme takes extreme steps in order to regulate $p_{C_{i-1}}$.

B. Early Drop by Span (EDS)

In order to alleviate the side effect of EDT, we introduce an early drop by span (EDS) scheme that linearly increases $p^{\mathrm{ED}}_{C_i}$ as a function of $p_{C_{i-1}}$. Here, a span (range) of acceptable loss probabilities, $\delta_{C_{i-1}}$, for Class $(i-1)$ is chosen. The EDS scheme is triggered when the online loss probability of Class $(i-1)$, $p_{C_{i-1}}$, is higher than $P^{\mathrm{MIN}}_{C_{i-1}}$, where $P^{\mathrm{MIN}}_{C_{i-1}} = P^{\mathrm{MAX}}_{C_{i-1}} - \delta_{C_{i-1}}$. Thus, the early dropping probability of Class i bursts is given by

$$p^{\mathrm{ED}}_{C_i} = \begin{cases} 0, & p_{C_{i-1}} < P^{\mathrm{MIN}}_{C_{i-1}} \\ \frac{p_{C_{i-1}} - P^{\mathrm{MIN}}_{C_{i-1}}}{\delta_{C_{i-1}}}, & P^{\mathrm{MIN}}_{C_{i-1}} \leq p_{C_{i-1}} < P^{\mathrm{MAX}}_{C_{i-1}} \\ 1, & p_{C_{i-1}} \geq P^{\mathrm{MAX}}_{C_{i-1}} \end{cases} \tag{5}$$

where $i \geq 1$.

The span ($\delta_{C_{i-1}}$) can be chosen as a percentage value of $P^{\mathrm{MAX}}_{C_{i-1}}$. We observe that, if $\delta_{C_{i-1}}$ is too high, EDS will be triggered prematurely, leading to high loss probability for bursts of lower priority classes; while, if $\delta_{C_{i-1}}$ is too low, $p^{\mathrm{ED}}_{C_i}$ will be high, also resulting in high loss probability for bursts of lower priority classes.

IV. Wavelength Grouping Mechanism

In this section, we propose another mechanism, known as *wavelength grouping* for supporting absolute loss probability in OBS networks. In the wavelength grouping mechanism, traffic is classified into different groups, and a label is assigned to each group. Each group is provisioned a minimum number of wavelengths. One approach to group the traffic is to assign all traffic of the same service class to the same unique group. Thus, on each link l, Class i bursts are assigned the same unique local label Li. We obtain W_{C_i} and $P^{\mathrm{MAX}}_{C_i}$ for each guaranteed Class i traffic from (2) and (3). Link l must provide W_{C_i}

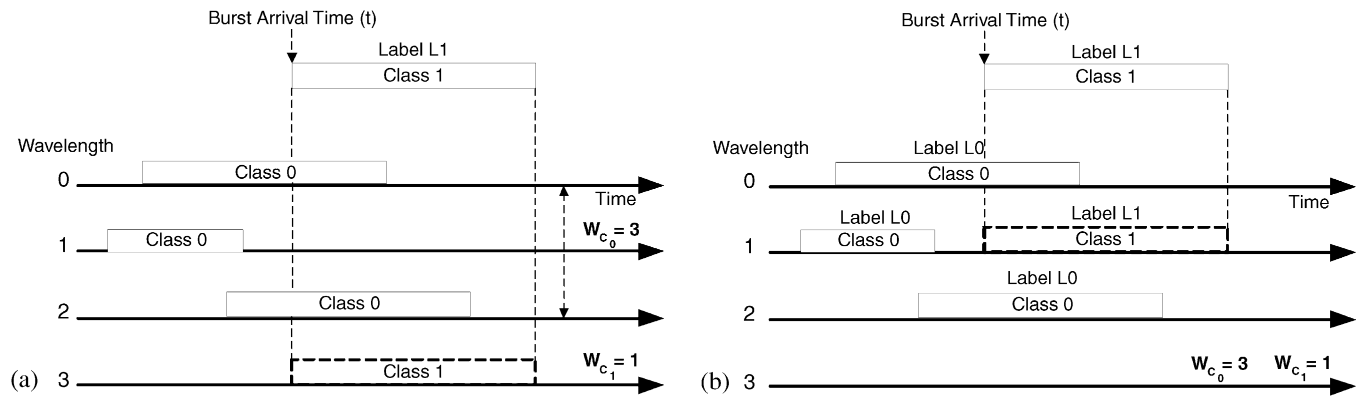


Fig. 3. Illustration of (a) SWG, and (b) DWG schemes.

wavelengths for bursts in the group with assigned Label Li in order to guarantee $P_{C_i}^{\mathrm{MAX}}$. If we run out of wavelengths, then the requirements of the remaining guaranteed classes of traffic cannot be satisfied with the given network capacity. On the other hand, if wavelengths are still available after provisioning wavelengths for all the guaranteed traffic, these remaining wavelengths can be used to carry the nonguaranteed traffic. We propose two schemes for wavelength grouping, namely, *static wavelength grouping* (SWG) and *dynamic wavelength grouping* (DWG).

A. Static Wavelength Grouping (SWG)

In SWG, a fixed set of wavelengths is dedicated for the traffic within a given group. If W_{C_0} wavelengths on link l are required for bursts in the group with assigned Label $L0$, the first W_{C_0} wavelengths $(w_0, w_1, \ldots, w_{(W_{C_0}-1)})$ are reserved for bursts in this group. Furthermore, bursts labeled $L0$ can only use these $w_0, w_1, \ldots, w_{(W_{C_0}-1)}$ wavelengths on the link l. In the case that more than one class of traffic is guaranteed, the process is repeated until the necessary wavelengths have been reserved for all of the guaranteed traffic. The remaining unreserved wavelengths are used to carry the best-effort traffic. For the scenario shown in Fig. 3(a), when a burst labeled $L1$ arrives at time t, it can only be scheduled on Wavelength 3, which is statically preassigned to the bursts labeled $L1$.

B. Dynamic Wavelength Grouping (DWG)

In DWG, a fixed number of wavelengths, but not necessarily a fixed set of wavelengths, is reserved for the traffic within a given group. To ensure that the number of wavelengths occupied by bursts of a given group does not exceed the number of wavelengths provisioned for that group, the OBS node must keep track of the number of wavelengths currently occupied by bursts of each group. A burst with a given label can be dynamically scheduled onto an available wavelength if the number of wavelengths currently occupied by bursts of the same label is less than the number of wavelengths provisioned for that group. In Fig. 3(b), suppose the number of wavelengths that bursts labeled $L1$ can use is, $W_{C_1} = 1$. When a burst labeled $L1$ arrives at time t, Wavelength 1 and Wavelength 3 are available and no bursts labeled $L1$ are currently scheduled. Hence, the arriving burst is scheduled on Wavelength 1, which is the latest available unscheduled channel.

Comparing SWG and DWG, we note that SWG is less complex and simpler to implement. However, DWG has the advantage of being able to dynamically schedule a burst onto the best wavelength based on the channel allocation status of each link, thereby improving network performance.

V. Integrated Schemes

Without the help of an early dropping mechanism, the wavelength grouping mechanism schedules the bursts of a given class only on a limited number of wavelengths, even when the loss probabilities of other classes of traffic are much lower than their required maximum loss probabilities. This restriction results in inefficient wavelength utilization. Therefore, we integrate the early dropping mechanism with the wavelength grouping mechanism to achieve better performance. In the early dropping mechanism, EDS has significantly better loss performance than EDT based on simulation results (Fig. 6); hence, we integrate EDS with the wavelength grouping schemes. In the integrated schemes, EDS assigns a local label to each burst based on the class of the burst and the current value of the corresponding early dropping vector. The wavelength grouping mechanism provisions a minimum number of wavelengths for each group of traffic with the same label and schedules each burst based on the provisioning.

We now describe an approach to assign labels and to provision the necessary wavelengths for the integrated schemes, using a two-class example. Fig. 4 presents the burst scheduling process in the integrated schemes. EDS is implemented by an *EDS Labeler*, and wavelength grouping is implemented by a *WG Scheduler*. Initially, the EDS labeler labels each burst according to the class of the burst and the value of the corresponding early dropping vector, $\mathrm{ED}_1 = \{e_1\}$. As shown in Table I, a burst is assigned a Label $L0$ either if the burst is of Class 0, or if the burst is of Class 1 and e_1 is 0. A burst is assigned a Label $L1$ if the burst is of Class 1 and e_1 is 1. The labeled burst is then sent to the WG scheduler, which schedules the burst solely based on its label. Table II gives the number of wavelengths provisioned for each group of bursts with a given label. A burst labeled $L0$ is allowed to be scheduled on any of the W wavelengths. Therefore, all of the wavelengths can be utilized when the early dropping scheme is not triggered and all the arriving bursts are labeled $L0$. A burst labeled $L1$ can only

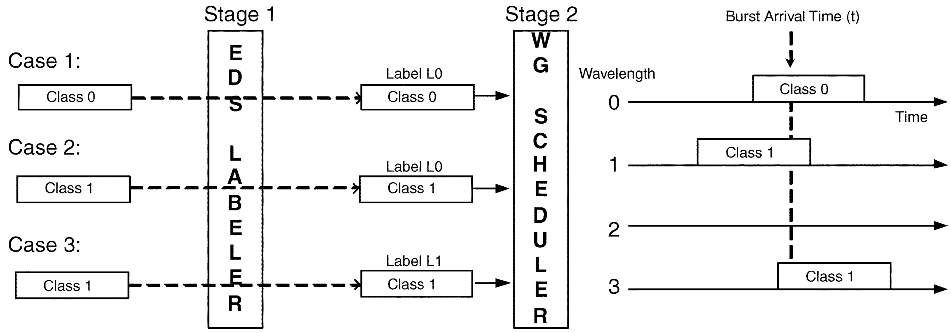


Fig. 4. Illustration of the integrated schemes.

TABLE I
INTEGRATED SCHEMES: LABEL ASSIGNMENT

e_1	Class 0 Burst Label	Class 1 Burst Label
0	L0	L0
1	L0	L1

TABLE II
INTEGRATED SCHEMES: WAVELENGTH PROVISIONING

Burst Label	Arrival Rate	Wavelengths provisioned
L0	λ_{L0}	W
L1	λ_{L1}	$W - W_{C_0}$

be scheduled on W_{C1} wavelengths, where $W_{C_1} = W - W_{C_0}$. This restriction ensures that there are a required number of W_{C0} wavelengths on which the bursts labeled $L0$ can be scheduled.

The computational overhead of the integrated schemes includes incrementing counters for monitoring traffic statistics, updating the early dropping vector, and label matching for each arriving burst. Updating the early dropping vector involves the generation of a random number and a comparison between the random number and the early dropping probability. These functions do not result in much computational overhead. Label matching also does not result in much computational overhead, since the static label assignment table has very few entries. In order to handle the burst quickly, the statistic counters can be incremented and the early dropping probabilities can be calculated after a burst is scheduled on a wavelength.

A. Integrated EDS and SWG

Using SWG, a burst labeled $L0$ can be scheduled on any available wavelength. while a burst labeled $L1$ can only be scheduled on the statically preassigned W_{C_1} wavelengths. In Fig. 4, we illustrate three possible burst arrival scenarios. The current wavelength allocation is shown on the right hand side of the figure. In Case 1, when a Class 0 burst labeled $L0$ arrives, the burst is scheduled on Wavelength 2. In Case 2, when a Class 1 burst labeled $L0$ arrives, the burst is also scheduled on Wavelength 2. While in Case 3, when a Class 1 burst labeled $L1$ arrives, the burst cannot be scheduled on Wavelength 2, since a burst labeled $L1$ can be scheduled only on the statically provisioned Wavelength 3.

B. Integrated EDS and DWG

Following LAUC, the DWG scheduler records the label of the latest-scheduled burst on every wavelength. When a burst labeled $L0$ arrives, DWG can schedule the burst on any available wavelength. On the other hand, when a burst labeled $L1$ arrives, the burst is scheduled on any of the available wavelengths, as long as the number of bursts labeled $L1$ already scheduled at the arrival time of the arriving burst is less than W_{C_1}. In Fig. 4, suppose the label of the latest scheduled burst recorded on each of Wavelengths 0, 1, and 3 is $L0$. With the DWG scheduler, for all three burst arrival scenarios, the arriving burst can be scheduled on Wavelength 2.

The integrated schemes provide better resource allocation compared to each of the stand-alone schemes for the following reasons. First, in the wavelength grouping schemes, Class 1 bursts can be scheduled only on W_{C_1} wavelengths, while, in the integrated schemes, the Class 1 bursts labeled $L0$ can be scheduled on any wavelength. Second, compared to early dropping schemes, the integrated schemes reduce the unnecessary intentional dropping of Class 1 bursts, since the Class 1 bursts labeled $L1$ can use a maximum of W_{C_1} wavelengths.

VI. PATH CLUSTERING

In order to implement the proposed per-hop mechanisms over an entire network, an approach is necessary to ensure that the combined loss probability on each hop along a path satisfies the end-to-end loss requirements. A simple technique (as described in Section II), is to have the same loss guarantee, $P_{C_i}^{\mathrm{MAX}}$, at every hop, so that the end-to-end loss probability of the maximum hop-distance path is guaranteed. In this technique, each class of traffic would be assigned its own unique priority level. The limitation of this approach is that the loss probability experienced by bursts that are traversing shorter hop-distance paths will be much lower than the required end-to-end loss requirement. This reduced loss probability for the bursts traversing shorter hop-distance path leads to increased loss probability for the nonguaranteed traffic (higher intentional dropping).

Another simple (but extreme) technique is to set different P_{Ci}^{MAX} for each specific hop-distance path at every node. Such an approach, though optimal in performance, is not scalable. For example, if the number of different hop-distances of paths is six and the number of traffic classes is two, then the number of priority levels supported at the core nodes must be equal to seven. Class 0 traffic would have six priority levels corresponding to the six different hop-distance paths, and Class 1 traffic would have a single priority level. Hence, if a network supports multiple service classes, the scheduling at each node becomes impractical. In order to provide absolute loss guarantee over an entire network in a practical manner, we propose a *path clustering* technique, which aims to achieve a balance between the number of priority levels that needs to be maintained at each node and the amount of intentional dropping of lower priority class bursts in the entire network.

In path clustering, the source-destination pairs are divided into different clusters based on their path hop-distance. In the following discussion, we assume that the maximum number of clusters supported in the network is two. Consider a network that has paths of hop-distance from $\{1, 2, \ldots, k\}$. For each service class i, let the per-hop loss guarantee for the maximum hop-distance path be $P_{C_i}^{\mathrm{MAX}}$, and let a possible clustering combination be $G_{C_i}^0 = \{\lfloor k/2 \rfloor + 1, \ldots, k\}$, and $G_{C_i}^1 = \{1, 2, \ldots, \lfloor k/2 \rfloor\}$. Each node must maintain two different traffic statistics for each guaranteed service class i, one for each cluster. This can be achieved by assigning a unique priority to all guaranteed traffic belonging to a specific cluster. If there are n guaranteed service classes and one best effort service class, the number of priority levels required is $(2n+1)$. The traffic belonging to Cluster $G_{C_i}^0$ is assigned a higher priority, $(2i)$, than the traffic belonging to Cluster $G_{C_i}^1$, which is assigned a lower priority, $(2i+1)$. The per-hop loss guarantee for traffic of each priority is computed based on the maximum hop-distance in the corresponding cluster using (2). Then, paths of hop-distance belonging to Cluster $G_{C_i}^0$, must provide a per-hop loss guarantee, $P_{P_{2i}}^{\mathrm{MAX}}$ for priority $(2i)$ traffic, and paths of hop-distance belonging to Cluster $G_{C_i}^1$, must provide a per-hop loss guarantee, $P_{P_{2i+1}}^{\mathrm{MAX}}$ for priority $(2i+1)$ traffic. We observe that the traffic traversing paths belonging to Cluster $G_{C_i}^1$ can still satisfy the end-to-end loss guarantee, $P_{C_i}^{NET}$, with a relaxed per-hop maximum of $P_{P_{2i+1}}^{\mathrm{MAX}}$, since this traffic traverses fewer hops.

We now describe the procedure for provisioning the required number of wavelengths for each guaranteed class of traffic, the procedure for scheduling using the integrated EDS and DWG scheme, and the procedure for finding the optimal path clustering. Without loss of generality, we consider a network with two classes of traffic that can support two clusters.

A. Provisioning Minimum Number of Wavelengths

This section describes how to provision the minimum number of wavelengths required for each guaranteed class of traffic. We need to compute the arrival rates for the guaranteed traffic on every link based on the clustering, the traffic arrival distribution, the routing algorithm, and the network topology.

Given a network with two classes of traffic, in which Class 0 traffic is guaranteed an absolute loss probability, and Class 1 is the best-effort traffic, the network must support at least three traffic priorities in order to handle two clusters. Each ingress node assigns either Priority 0 or Priority 1 to Class 0 bursts based on the clustering, and assigns Priority 2 to all Class 1 bursts.

Let us consider the following notation.

- λ_{sd} Traffic arrival rate between source s and destination d.
- $\lambda_{C_0}^{sd}$ Arrival rate of Class 0 traffic between source s and destination d.
- $\lambda_{C_1}^{sd}$ Arrival rate of Class 1 traffic between source s and destination d, where $\lambda_{C_1}^{sd} = (\lambda_{sd} - \lambda_{C_0}^{sd})$.
- l_{ij} Link between node i and node j.
- r_{sd} Route from source s to destination d based on routing algorithm.
- h_{sd} Hop-distance of route (path) r_{sd}.
- $\lambda_{P_0}^{ij}$ Arrival rate of Priority 0 traffic on link l_{ij}.
- $\lambda_{P_1}^{ij}$ Arrival rate of Priority 1 traffic on link l_{ij}.
- $\lambda_{P_2}^{ij}$ Arrival rate of Priority 2 traffic on link l_{ij}.

The arrival rates for the prioritized traffic on link l_{ij} are as follows:

$$\lambda_{P_0}^{ij} = \sum_{\{\forall(s,d) | l_{ij} \in r_{sd}, h_{sd} \in G_{C_0}^0\}} \lambda_{C_0}^{sd} \tag{6}$$

$$\lambda_{P_1}^{ij} = \sum_{\{\forall(s,d) | l_{ij} \in r_{sd}, h_{sd} \in G_{C_0}^1\}} \lambda_{C_0}^{sd} \tag{7}$$

$$\lambda_{P_2}^{ij} = \sum_{\{\forall(s,d) | l_{ij} \in r_{sd}\}} \lambda_{C_1}^{sd}. \tag{8}$$

On every link, the loss guarantee of Priority 0 traffic, $P_{P_0}^{\mathrm{MAX}}$, and the loss guarantee of Priority 1 traffic, $P_{P_1}^{\mathrm{MAX}}$, are computed based on (2), where D is equal to the maximum hop-distance of the cluster. The node must provision the minimum number of wavelengths, namely, W_{P_0} and W_{P_1}, for Priority 0 and Priority 1 traffic. W_{P_0} and W_{P_1} are provisioned using (3).

B. Scheduling Using Integrated EDS and DWG

This section describes the scheduling scheme in path clustering based on the integrated EDS and DWG. The label assignment is similar to the integrated scheme. Table III shows the label assignment under different traffic scenarios when path clustering is implemented. The first two columns represent the early dropping flags, e_1 and e_2, for bursts of Priority 1 and Priority 2, respectively. e_1 is set to 0 with probability $(1 - p_{P_1}^{\mathrm{ED}})$ and is set to 1 with probability $p_{P_1}^{\mathrm{ED}}$, while, e_2 is set to 0 with probability $(1 - p_{P_2}^{\mathrm{ED}})$ and is set to 1 with probability $p_{P_2}^{\mathrm{ED}}$. The third, fourth, and fifth columns indicate the labels assigned to the arriving bursts of Priority 0, Priority 1, and Priority 2, respectively. Let us now consider the four different scenarios. For the case in which e_1 is set to 0 and e_2 is set to 0, an arriving burst of Priority 0, Priority 1, or Priority 2 is labeled *L*0. For the scenario in which e_1 is set to 1 and e_2 is set to 0, an arriving burst of Priority 0 is labeled *L*0, while an arriving burst of Priority 1 or Priority 2 is labeled *L*1. When e_1 is set to 0 and e_2 is set to 1, an arriving burst of Priority 0 or Priority 1 is labeled *L*0, and an arriving burst of Priority 2 is labeled *L*2. In the last scenario, when e_1 is set to 1 and e_2 is set to 1, an arriving burst of Priority 0 is labeled *L*0, an arriving burst of Priority 1 is labeled *L*1, and an arriving burst of Priority 2 is labeled *L*3.

Table IV gives the required number of wavelengths on which a burst with a given label can be scheduled. Let W_{P_2}

TABLE III
PATH CLUSTERING: LABEL ASSIGNMENT

e_1	e_2	Priority 0 Burst Label	Priority 1 Burst Label	Priority 2 Burst Label
0	0	L0	L0	L0
1	0	L0	L1	L1
0	1	L0	L0	L2
1	1	L0	L1	L3

TABLE IV
PATH CLUSTERING: WAVELENGTH PROVISIONING

Burst Label	Arrival Rate	Wavelengths Provisioned
L0	λ_{L0}	W
L1	λ_{L1}	W_{P1}
L2	λ_{L2}	$W - W_{P1}$
L3	λ_{L3}	W_{P2}

be the number of wavelengths provisioned for Priority 2 traffic (nonguaranteed traffic). Since Priority 0 and Priority 1 traffic belong to the same Class 0 traffic, W_{P_2} is provisioned based on the number of wavelengths required by Class 0 traffic, that is, $W_{P_2} = W - W_{C_0}$. We allow all bursts labeled *L*0 to be scheduled on any available wavelength. All bursts labeled *L*1 are only scheduled on W_{P_1} wavelengths since we need to provide the loss guarantee of Priority 1 traffic. All bursts labeled *L*2 are scheduled on $(W - W_{P_1})$ wavelengths, since this restriction ensures that a minimum of W_{P_1} wavelengths are reserved for bursts labeled *L*1. All bursts labeled *L*3 are scheduled on W_{P_2} wavelengths.

C. Finding Optimal Path Clustering

There are two parameters that define a path clustering: the number of clusters and the elements in each cluster. The number of clusters depends on how many priority levels the network can support. The assignment of elements into each cluster determines the per-hop loss guarantee of each cluster and the arrival rates of different priority traffic. The optimal path clustering can be found offline as follows.

Let us consider a network with paths of hop-distances from one to six. Given that the network can support two clusters, the possible cluster combinations for the six hop-distance network will be {1}{2,3,4,5,6}, {1,2}{3,4,5,6}, {1,2,3}{4,5,6}, {1,2,3,4}{5,6}, and {1,2,3,4,5}{6}. For each cluster combination, we initially provision the wavelengths on every link in the network. For example, with a clustering of {1} and {2,3,4,5,6}, all Class 0 bursts along a path with a hop-distance of one will be assigned Priority 1 (low), and all Class 0 bursts along a path with a hop-distance of two, three, four, five, or six will be assigned Priority 0 (high). Each core node would satisfy two different per-hop maximum loss probabilities, $P_{P_0}^{\mathrm{MAX}}$ and $P_{P_1}^{\mathrm{MAX}}$. All Class 1 bursts will be assigned Priority 2. We can then compute the per-hop loss of each priority traffic using the analytical model in the following section. The end-to-end loss probability of Class 1 traffic can be obtained by rearranging (2). This procedure is repeated for each of the cluster combinations for the six hop-distance network, {1}{2,3,4,5,6}, {1,2}{3,4,5,6}, {1,2,3}{4,5,6}, {1,2,3,4}{5,6}, and {1,2,3,4,5}{6}. The clustering with the least end-to-end loss probability of Class 1 traffic, which can also support the maximum loss probability of the guaranteed Class 0 traffic, is the optimal clustering. If the traffic is uniformly distributed among all source-destination pairs, we only need to analyze the per-hop loss performance of a bottleneck link under each cluster combination. In this case, the optimal clustering is the cluster combination with the least per-hop loss probability of Class 1 traffic.

VII. ANALYTICAL MODEL

In this section, we develop analytical loss models for the integrated EDS and DWG scheme and path clustering. Without loss of generality, we model a two-class network. The proposed model can be extended to a multi-class network. We assume that the burst arrival process is Poisson with rate λ, and the burst length is exponentially distributed with an average burst length of $1/\mu$. For analytical tractability, we assume that the bursts arrive in the same order as their BHPs at each node. Although bursts may in fact arrive in a different order as their BHPs, this assumption is still reasonable because as long as the difference in offset time between two bursts is small compared to their inter-arrival time, the probability that bursts arrive out of order is low. The difference in offset time between two bursts is θh, where θ is the BHP processing time at a single node, and h is the difference in hop-distance. The probability that bursts arrive out of order is less than the probability that the difference in offset time between two bursts is greater than their interarrival time, that is

$$\text{Prob (bursts arrive out of order)} < \text{Prob (burst interarrival time} < \theta h) \quad (9)$$

and

$$\text{Prob (burst interarrival time} < \theta h) = 1 - e^{-\lambda\theta h}. \quad (10)$$

A. Loss Model for Integrated EDS and DWG

Let W be the total number of wavelengths on each link, and let W_{C_0} be the minimum number of wavelengths provisioned for Class 0 traffic. The Class 1 traffic that is labeled *L*1 by EDS can be scheduled on a maximum of $W_{C_1} = W - W_{C_0}$ wavelengths. The bursts of Class 0 have an arrival rate of λ_{C_0}, and the bursts of Class 1 have an arrival rate of λ_{C_1}. In EDS, Class 1 bursts are labeled *L*1 with probability $p_{C_1}^{\mathrm{ED}}$, which is calculated based on the measured online loss probability p_{C_0}. Let $P_{C_1}^{\mathrm{ED}}$ be the mean of $p_{C_1}^{\mathrm{ED}}$ at steady state. Let λ_{L0} be the arrival rate of bursts that are labeled *L*0 by EDS and can be scheduled on W wavelengths; let λ_{L1} be the arrival rate of bursts that are labeled *L*1 by EDS and can be scheduled on W_{C_1} wavelengths. The arrival rate of bursts labeled *L*0 is

$$\lambda_{L0} = \lambda_{C_0} + (1 - P_{C_1}^{\mathrm{ED}})\lambda_{C_1} \quad (11)$$

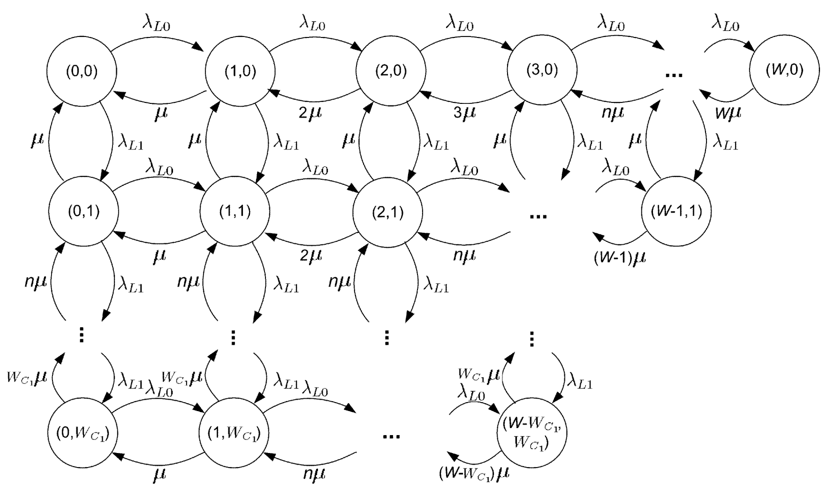


Fig. 5. Markov chain for the integrated EDS and DWG scheme.

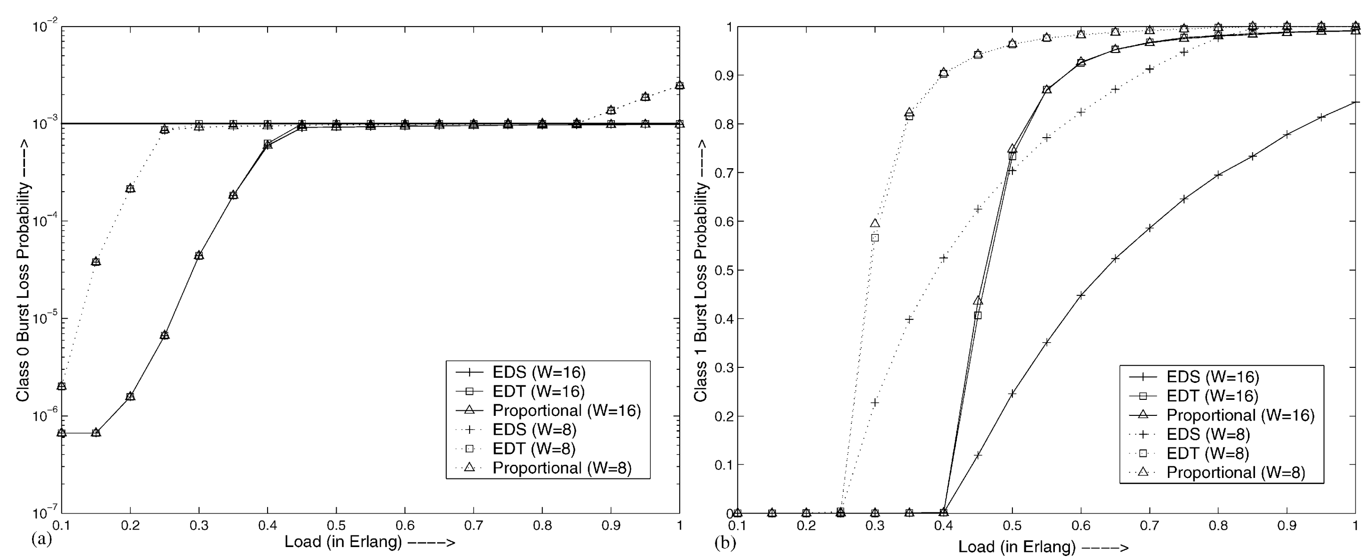


Fig. 6. (a) Class 0 and (b) Class 1 loss probability versus load for the EDS, EDT, and proportional schemes.

and the arrival rate of bursts labeled $L1$ is

$$\lambda_{L1} = P_{C_1}^{\mathrm{ED}} \lambda_{C_1}. \tag{12}$$

Based on the decomposition property of a Poisson process, the arrival processes of bursts labeled $L0$ and bursts labeled $L1$ are also Poisson.

We model each outgoing link as a continuous time Markov chain with the state defined as $X = \{x_0, x_1\}$, where x_0 is the number of wavelengths that are busy serving bursts labeled $L0$ and x_1 is the number of wavelengths that are busy serving bursts labeled $L1$. The state transition diagram for the Markov chain is shown in Fig. 5 and the state transition rates are as follows:

$$\begin{aligned} q_{(x_0,x_1)(x_0,x_1+1)} &= \begin{cases} \lambda_{L1}, & x_0 + x_1 < W \ \& \ x_1 < W_1 \\ 0, & \text{otherwise} \end{cases} \\ q_{(x_0,x_1)(x_0+1,x_1)} &= \begin{cases} \lambda_{L0}, & x_0 + x_1 < W \\ 0, & \text{otherwise} \end{cases} \\ q_{(x_0,x_1)(x_0,x_1-1)} &= \begin{cases} x_1\,\mu, & x_1 > 0 \\ 0, & \text{otherwise} \end{cases} \\ q_{(x_0,x_1)(x_0-1,x_1)} &= \begin{cases} x_0\,\mu, & x_0 > 0 \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{13}$$

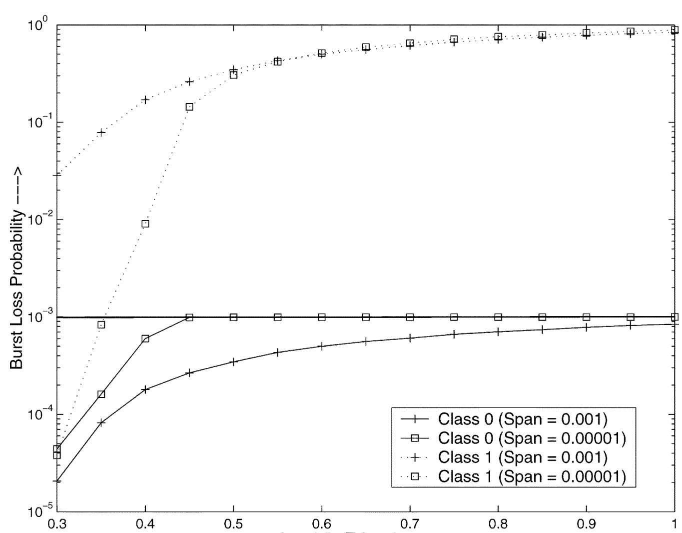


Fig. 7. Burst loss probability versus load with different values of span in EDS with $P_{C_0}^{\mathrm{MAX}} = 10^{-3}$.

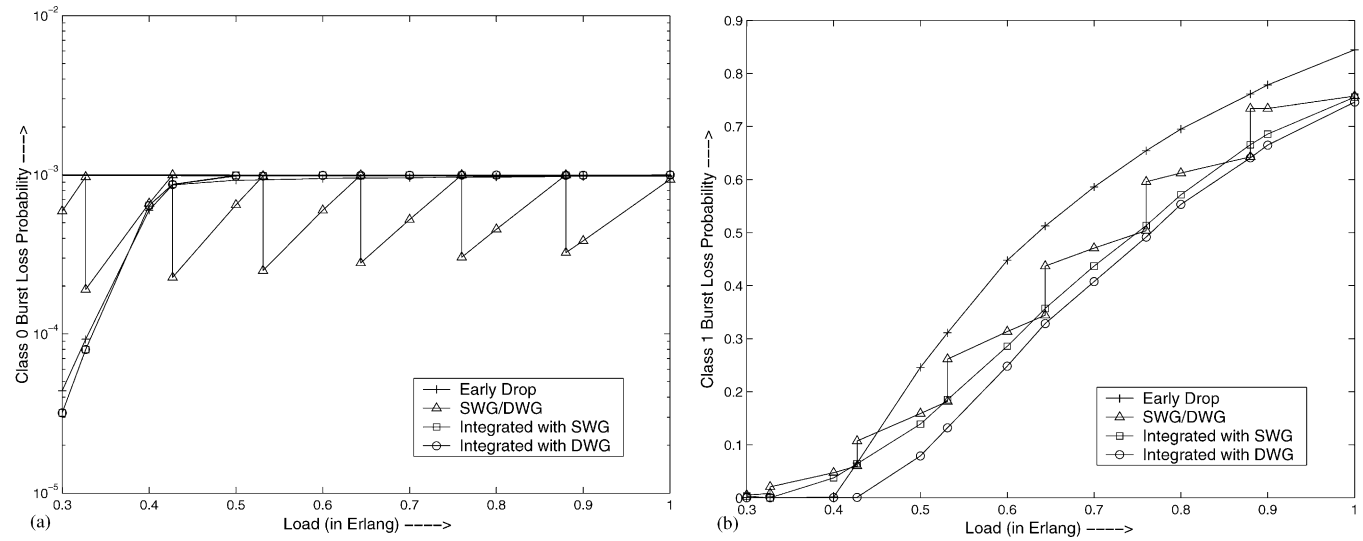


Fig. 8. (a) Class 0 and (b) Class 1 loss probability versus load for the proposed absolute QoS schemes.

From the Markov chain, we can solve the steady-state probabilities $p(x_0, x_1)$. The probability that a burst labeled $L0$ is dropped by the DWG scheduler is equal to the probability that all of the wavelengths are busy. Thus, the loss probability for bursts labeled $L0$ is given by

$$P_{L0} = \sum_{i=0}^{W_{C_1}} p(W-i, i). \tag{14}$$

A burst labeled $L1$ is dropped either when all wavelengths are occupied, or when the number of bursts labeled $L1$ currently scheduled is W_{C1}, but all of the wavelengths are not fully occupied. Therefore, the loss probability for bursts labeled $L1$ is as follows:

$$P_{L1} = P_{L0} + \sum_{i=0}^{W-W_{C_1}-1} p(i, W_{C_1}). \tag{15}$$

Hence, the loss probabilities of Class 0 and Class 1 bursts are given as

$$P_{C_0} = P_{L0}, \tag{16}$$
$$P_{C_1} = P_{L1}\, P_{C_1}^{\mathrm{ED}} + P_{C_0}\,(1 - P_{C_1}^{\mathrm{ED}}) \tag{17}$$

where $P_{C_1}^{\mathrm{ED}}$ is calculated from P_{C_0} using (5).

We observe from the set of equations, that if we know any one of P_{C_0}, $P_{C_1}^{\mathrm{ED}}$, or λ_{L0}, we can obtain the other two. Therefore, we start with an initial value of λ_{L0} equal to λ_{C_0}, since all Class 0 bursts are assigned Label $L0$. We increment the value of λ_{L0} in discrete steps by some small value γ until it reaches the total arrival rate, λ. For each value of λ_{L0}, we obtain the corresponding value of P_{C_0} from (16). Then, we compute $P_{C_1}^{\mathrm{ED}}$ from (5). We can also obtain $P_{C_1}^{\prime\mathrm{ED}}$ from (12), where $P_{C_1}^{\prime\mathrm{ED}} = \lambda_{L1}/\lambda_{C_1}$. We choose the value of λ_{L0} for which $\mid P_{C_1}^{\mathrm{ED}} - P_{C_1}^{\prime\mathrm{ED}} \mid$ is minimum. If this minimum is less than some small value ε, we stop; otherwise, we reduce the value of γ by some amount and repeat the process. Thus, the correct values of P_{C_0}, $P_{C_1}^{\mathrm{ED}}$, and λ_{L0} are obtained.

B. Loss Model for Path Clustering

Given the arrival rates of Class 0 and Class 1 traffic, as well as the clustering, we calculate the arrival rates of every priority traffic, namely, λ_{P_0}, λ_{P_1}, and λ_{P_2}. Let $P_{P_1}^{\mathrm{ED}}$ be the mean of $p_{P_1}^{\mathrm{ED}}$, and $P_{P_2}^{\mathrm{ED}}$ be the mean of $p_{P_2}^{\mathrm{ED}}$ at steady state. We then compute the arrival rate for each type of traffic with a given label according to Table III. The arrival rates of traffic labeled $L0$, $L1$, $L2$, $L3$ are as follows:

$$\lambda_{L0} = \lambda_{P_0} + (1 - P_{P_1}^{\mathrm{ED}})\lambda_{P_1} + (1 - P_{P_1}^{\mathrm{ED}})(1 - P_{P_2}^{\mathrm{ED}})\lambda_{P_2} \tag{18}$$
$$\lambda_{L1} = P_{P_1}^{\mathrm{ED}}\lambda_{P_1} + P_{P_1}^{\mathrm{ED}}(1 - P_{P_2}^{\mathrm{ED}})\lambda_{P_2} \tag{19}$$
$$\lambda_{L2} = (1 - P_{P_1}^{\mathrm{ED}})P_{P_2}^{\mathrm{ED}}\lambda_{P_2} \tag{20}$$
$$\lambda_{L3} = P_{P_1}^{\mathrm{ED}}P_{P_2}^{\mathrm{ED}}\lambda_{P_2}. \tag{21}$$

We again model each link as a continuous time Markov chain with the state defined as $X = \{x_0, x_1, x_2, x_3\}$, where x_0, x_1, x_2, and x_3 are the number of wavelengths that are busy serving bursts labeled $L0$, $L1$, $L2$, and $L3$, respectively. The state transition rates are as follows:

$$
\begin{aligned}
q_{(x_0,x_1,x_2,x_3)(x_0-1,x_1,x_2,x_3)} &= \begin{cases} x_0\,\mu, & x_0 > 0 \\ 0, & \text{otherwise} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1-1,x_2,x_3)} &= \begin{cases} x_1\,\mu, & x_1 > 0 \\ 0, & \text{otherwise} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1,x_2-1,x_3)} &= \begin{cases} x_2\,\mu, & x_2 > 0 \\ 0, & \text{otherwise} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1,x_2,x_3-1)} &= \begin{cases} x_3\,\mu, & x_3 > 0 \\ 0, & \text{otherwise} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0+1,x_1,x_2,x_3)} &= \begin{cases} \lambda_{L0}, & x_0 < W \\ 0 & \text{otherwise,} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1+1,x_2,x_3)} &= \begin{cases} \lambda_{L1}, & x_1 < W_{P1} \\ 0 & \text{otherwise,} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1,x_2+1,x_3)} &= \begin{cases} \lambda_{L2}, & x_2 < (W - W_{P1}) \\ 0, & \text{otherwise} \end{cases}\\
q_{(x_0,x_1,x_2,x_3)(x_0,x_1,x_2,x_3+1)} &= \begin{cases} \lambda_{L3}, & x_3 < W_{P2} \\ 0, & \text{otherwise.} \end{cases}
\end{aligned} \tag{22}
$$

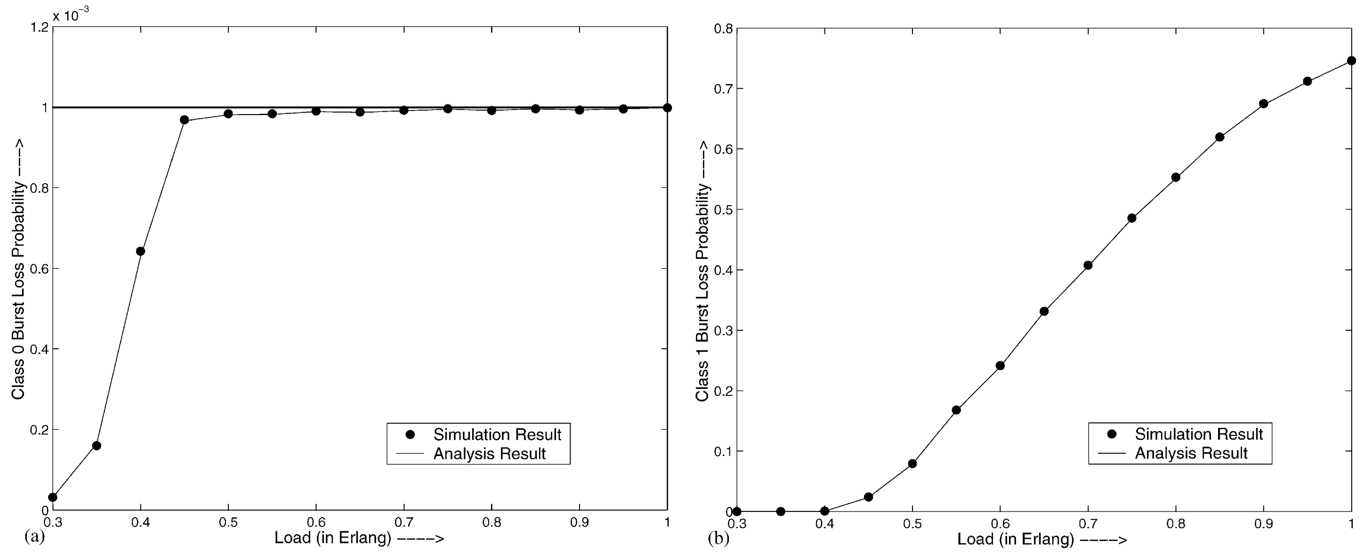


Fig. 9. (a) Class 0 and (b) Class 1 loss probability versus load using analytical and simulation results for the integrated EDS and DWG scheme.

From the Markov chain, we can solve the steady-state probabilities $p(x_0, x_1, x_2, x_3)$. The probability that a burst labeled $L0$ is dropped by the DWG scheduler is equal to the probability that all of the wavelengths are busy. Thus, the loss probability for bursts labeled $L0$ is given by

$$P_{L0} = \sum_{i=0}^{W_{P_1}} \sum_{j=0}^{W-W_{P_1}} \sum_{k=0}^{W_{P_2}} p(W-(i+j+k), i, j, k). \quad (23)$$

A burst labeled $L1$ is dropped either when all wavelengths are occupied, or when the number of wavelengths occupied by burst labeled $L1$ is W_{P_1}, but all of the wavelengths are not occupied. Therefore, the loss probability for bursts labeled $L1$ is as follows:

$$P_{L1} = P_{L0} + \sum_{i=0}^{W} \sum_{j=0}^{W-W_{P_1}} \sum_{k=0}^{W_{P_2}} p(i, W_{P_1}, j, k) \quad (24)$$

where $i+j+k+W_{P_1} < W$. Similarly

$$P_{L2} = P_{L0} + \sum_{i=0}^{W} \sum_{j=0}^{W_{P_1}} \sum_{k=0}^{W_{P_2}} p(i, j, W-W_{P_1}-i-k, k) \quad (25)$$

where $j + (W - W_{P_1}) < W$, and

$$P_{L3} = P_{L0} + \sum_{i=0}^{W} \sum_{j=0}^{W_{P_1}} \sum_{k=0}^{W-W_{P_1}} p(i, j, k, W_{P_2}) \quad (26)$$

where $i+j+k+W_{P_2} < W$.

Hence, the loss probabilities of Priority 0, Priority 1, and Priority 2 bursts are given as

$$P_{P_0} = P_{L0} \quad (27)$$
$$P_{P_1} = P_{L0}(1 - P_{P_1}^{\mathrm{ED}}) + P_{L1}P_{P_1}^{\mathrm{ED}} \quad (28)$$

and

$$P_{P_2} = P_{L0}(1 - P_{P_1}^{\mathrm{ED}})(1 - P_{P_2}^{\mathrm{ED}}) + P_{L1}P_{P_1}^{\mathrm{ED}}(1 - P_{P_2}^{\mathrm{ED}}) + P_{L2}(1 - P_{P_1}^{\mathrm{ED}})P_{P_2}^{\mathrm{ED}} + P_{L3}P_{P_1}^{\mathrm{ED}}P_{P_2}^{\mathrm{ED}} \quad (29)$$

where $P_{P_1}^{\mathrm{ED}}$ and $P_{P_2}^{\mathrm{ED}}$ are calculated from P_{P_0} and P_{P_1} respectively, using (5).

In the Markov chain for the two-priority model shown in Fig. 5, the total number of states is less than $(W+1)(W_{C_1}+1)$. The total number of states in the Markov chain for the three-priority model is less than $(W+1)(W_{P_1}+1)(W_{P_2}+1)(W-W_{P_1}+1)$. Hence, the analytical models are scalable in terms of the number of states as the number of wavelengths increases. However, as the number of priorities increases, the number of states increases rapidly. Finding an approximate analytical model for an arbitrary number of priorities is an open problem.

VIII. Numerical Results

Simulations are developed to evaluate the performance of the proposed schemes and to verify the analytical model.

A. Nonpath Clustering Results

In this section, we compare the scheduling schemes presented in Section III, IV, and V at the node-level. We also verify the analytical model of the integrated EDS and DWG scheme via simulation.

The simulation generates two classes of traffic, with Class 0 (high) and Class 1 (low) bursts. Burst arrivals follow a Poisson process with rate λ, and burst lengths are exponentially distributed with an average burst length of 100 μs. Unless we stated out, otherwise, the number of wavelengths on each link is 16. The transmission rate on a wavelength is 10 Gb/s. We assume that the core node has full wavelength conversion capability and has no buffering. The absolute loss requirement for Class 0 traffic is $P_{C_0}^{\mathrm{MAX}} = 10^{-3}$. Class 1 traffic receives best-effort service. 30% of the traffic is Class 0 and 70% of the traffic is Class 1. We set the span of the EDS scheme, $\delta_{C_0} = 0.1 \cdot P_{C_0}^{\mathrm{MAX}}$.

We first compare the performance of the early dropping schemes and the proportional scheme. We also compare the performance of these schemes under different number of wavelengths on a link. We adjust the proportionality factor in the proportional scheme, such that the average loss probability of

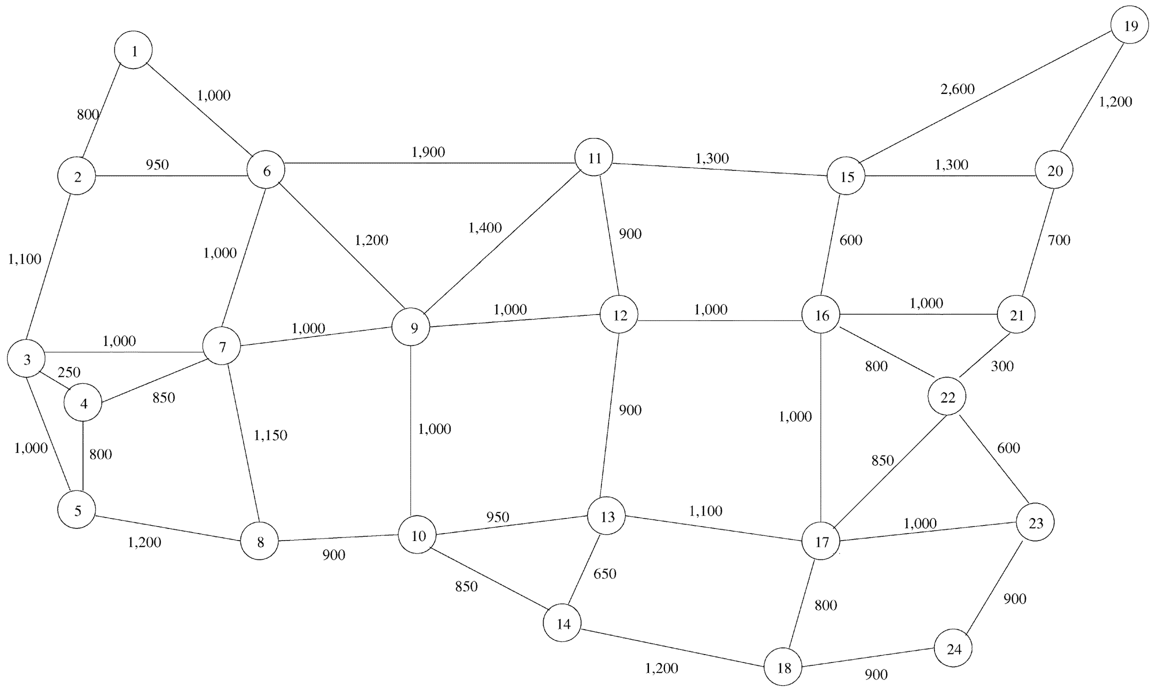


Fig. 10. 24-node mesh network (43 bidirectional links, average hop-distance $= 2.992$, average nodal degree $= 3.583$).

Class 0 traffic in the proportional scheme is mapped to the average loss probability of Class 0 traffic in EDS. Fig. 6(a) and (b) plots Class 0 and Class 1 burst loss probabilities versus load. We observe from Fig. 6(a) that when the number of wavelengths on a link is 16, the loss probabilities of Class 0 bursts in the EDS and proportional schemes are lower than the loss guarantee; on the other hand, EDT exceeds the guaranteed loss occasionally. For the case in which there are eight wavelengths on a link, the loss probabilities of Class 0 in all schemes exceeds the loss guarantee when the loads are higher than 0.8 Erlang. Hence, the early dropping schemes need wavelength provisioning for Class 0 traffic in order to meet the loss guarantee. In Fig. 6(b), at low loads, we see that the early dropping schemes are not triggered, since the loss probabilities are lower than their trigger points. The early dropping schemes are triggered earlier when the number of wavelengths on a link is 8 than when the number of wavelengths on a link is 16. We also see that once the loss probabilities reach the trigger points, EDS outperforms the other two schemes.

For the EDS scheme, we also study the performance of EDS with different span values. Fig. 7 plots the burst loss probability versus load with different values of span in EDS. We observe that at high loads, the total loss probability decreases with wider spans, since there is a wider range over which the loss probability of Class 0 bursts can be corrected. At low loads, the total loss probability increases with longer span, since the EDS scheme is triggered earlier.

We next compare the performance of EDS, SWG, DWG, and the integrated schemes. Fig. 8(a) and (b) plot Class 0 and Class 1 burst loss probabilities versus load for the absolute QoS schemes, respectively. In the absence of EDS, the loss performances of SWG and DWG are identical. We observe from Fig. 8(a) that all of the schemes can support the loss requirement of Class 0 traffic. As the load increases, the number of wavelengths provisioned for Class 0 traffic increases and the number of wavelengths provisioned for Class 1 traffic decreases accordingly. In the SWG and DWG schemes, if one more wavelength is provisioned to be dedicated to Class 0 traffic, the loss probability of Class 0 traffic will be reduced. For example, at loads between 0.53 and 0.64, the number of wavelengths provisioned for Class 0 is 10, and at loads between 0.64 and 0.76, the number of wavelengths provisioned for Class 0 is 11. Hence, there is a drop in the loss probability at a load of 0.64 in the SWG and DWG schemes. In Fig. 8(b), we observe that the loss probability experienced by Class 1 traffic in the integrated EDS and DWG scheme is the lowest, since Class 1 bursts have the flexibility of being assigned on any available wavelength, while in the integrated EDS and SWG scheme, Class 1 bursts must be scheduled on a particular set of wavelengths.

Fig. 9(a) and (b) compare the analytical results with the simulation results for the integrated EDS and DWG scheme. In the analytical model, we set $\varepsilon = 10^{-2}$. We see that the simulation results match very closely with the analytical results.

B. Path Clustering Results

We adopt a 24-node mesh network, shown in Fig. 10, for the network-level simulation in which the maximum hop-distance is 6 and the number of wavelengths on each link is 16. Traffic is uniformly distributed among the source-destination pairs. Fixed shortest-path routing is used. Since the maximum hop-distance in the 24-node network is 6, let $C1 = \{1,2,3,4,5,6\}$ represent the nonpath clustering case, and let $C2 = \{1\}\{2,3,4,5,6\}$, $C3 = \{1,2\}\{3,4,5,6\}$, $C4 = \{1,2,3\}\{4,5,6\}$, $C5 = \{1,2,3,4\}\{5,6\}$, and $C6 = \{1,2,3,4,5\}\{6\}$ be the five possible two-cluster combinations. The end-to-end loss guarantee for Class 0 traffic is $P_{C_0}^{NET} = 10^{-2}$, which results in $P_{C_0}^{\mathrm{MAX}} = 1.67 \times 10^{-3}$ for the nonpath clustering case. We study the loss performance of

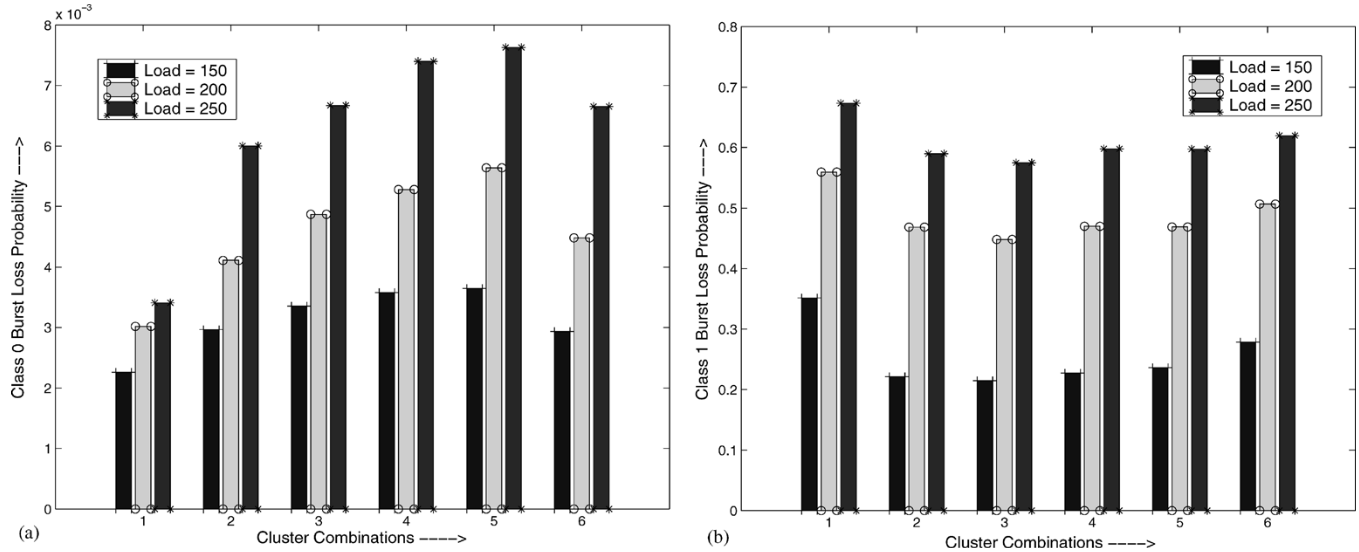


Fig. 11. (a) Class 0 and (b) Class 1 end-to-end loss probability versus different cluster combinations for the 24-node network.

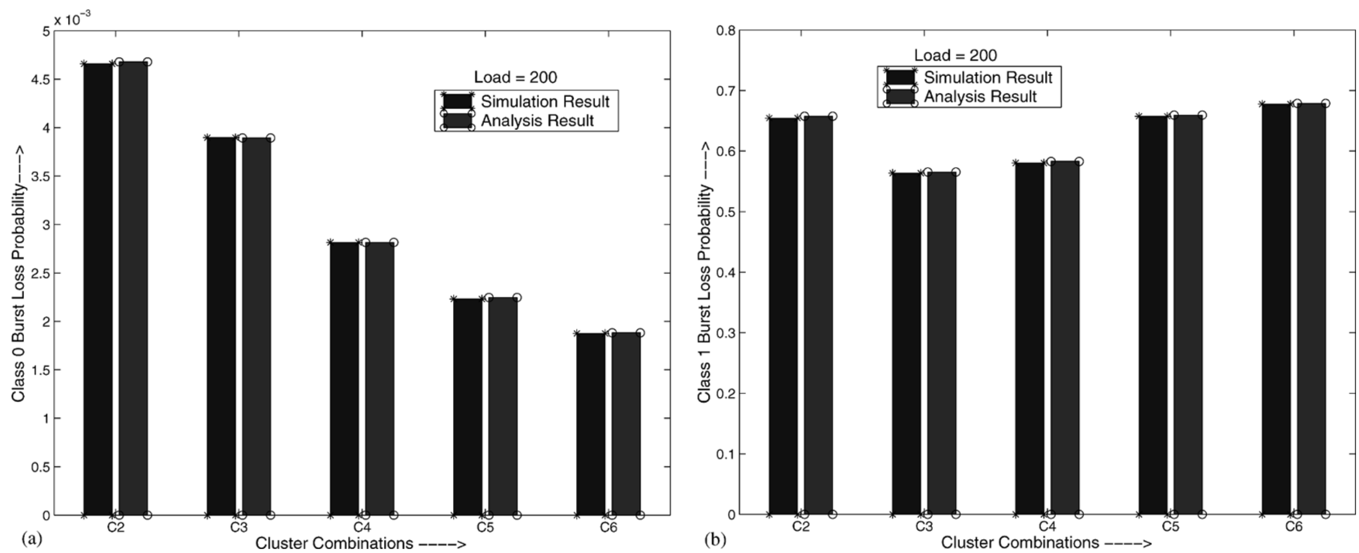


Fig. 12. (a) Class 0 and (b) Class 1 loss probability versus different cluster combinations for the bottleneck link (9, 10) at a network load of 20 Erlang.

these cluster combinations and apply the analytical model to compute the optimal path clustering.

Fig. 11 compares Class 0 and Class 1 end-to-end burst loss probabilities for different cluster combinations over an entire network at loads of 150, 200, and 250 Erlang. We observe in Fig. 11(a), that the end-to-end loss probabilities experienced by Class 0 traffic are below $P_{C_0}^{\mathrm{NET}} = 10^{-2}$. By using clustering, we can see an increase in loss probability experienced by Class 0 traffic, as compared to the nonclustering ($C1$) case. In Fig. 11(b), the end-to-end loss probability experienced by Class 1 traffic reduces considerably with clustering. At a load of 150, 200, and 250 Erlang, we see that cluster combination $C3$ outperforms all other combinations.

We apply the path clustering loss model to determine the optimal clustering. Since the traffic is uniformly distributed among all source-destination pairs, we analyze the bottleneck link (9, 10), and we compare the analytical results with the simulation results on link (9, 10) at a network load of 200 Erlang. Fig. 12(a) and (b) show the simulation results and the analytical results for the bottleneck link. We see that the analytical results closely match the simulation results. We also observe that cluster combination $C3$ has the least burst loss probability of Class 1 traffic at a load of 200 Erlang on link (9, 10), which is confirmed by the network-wide loss result in Fig. 11(b). Hence, using the offline calculation, we can find which of the many possible clustering combinations performs the best for a given topology at a given operating load range.

Fig. 13(a) and (b) plots Class 0 and Class 1 end-to-end burst loss probabilities versus load under the two best cluster combinations, $C2$ and $C3$, as well as the nonpath clustering case (C1). We observe from Fig. 13(a) that $C1$, $C2$, and $C3$ can satisfy the end-to-end loss requirement of Class 0 traffic. However, $C1$ incurs significantly higher loss probability for Class 1 traffic compared to $C2$ and $C3$, as illustrated in Fig. 13(b). We also observe that, at loads between 100 Erlang and 250 Erlang, cluster combination $C3$ performs the best.

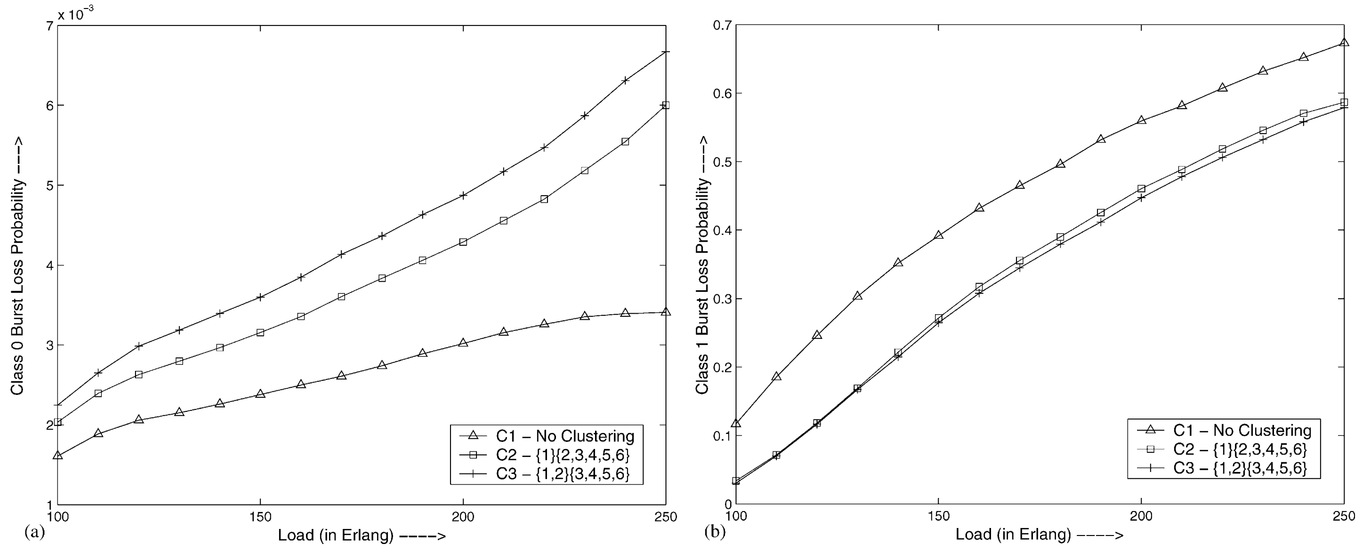


Fig. 13. (a) Class 0 and (b) Class 1 end-to-end loss probability versus load for the 24-node network.

IX. Conclusion

In this paper, we addressed the issue of absolute QoS support in an OBS network. We described two mechanisms, namely, early dropping and wavelength grouping, and integrated these two mechanisms to support loss guarantee in OBS core nodes. We showed that integrated early dropping by span with dynamic wavelength grouping has the best performance. We also proposed a path clustering technique to support absolute QoS over an entire network. We observed that the path clustering technique further reduces the loss probability experienced by the nonguaranteed traffic while satisfying the loss requirement of the guaranteed traffic, thereby improving the network-wide loss performance. We developed analytical models for the integrated early dropping by span and dynamic wavelength grouping scheme, as well as for the path clustering technique. The analytical models were verified via simulations. Using the path clustering analytical model, we computed the optimal clustering for a given load and verified the results with the network-level simulation.

References

[1] C. Qiao and M. Yoo, "Optical burst switching (OBS) – A new paradigm for an optical internet," *J. High Speed Networks*, vol. 8, pp. 69–84, Jan. 1999.

[2] J. S. Turner, "Terabit burst switching," *J. High Speed Networks*, vol. 8, pp. 3–16, Jan. 1999.

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

[4] J. Xu, C. Qiao, J. Li, and G. Xu, "Efficient channel scheduling algorithms in optical burst switching networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. 2003.

[5] V. M. Vokkarane, G. P. V. Thodime, V. B. T. Challagulla, and J. P. Jue, "Channel scheduling algorithms using burst segmentation and FDL's for optical burst-switched networks," in *Proc. IEEE ICC*, Anchorage, AK, May 2003.

[6] R. Ramaswami and K. N. Sivarajan, "Routing and wavelength assignment in all-optical networks," *IEEE/ACM Trans. Networking*, vol. 3, pp. 489–500, Oct. 1995.

[7] A. Bononi, G. A. Castanon, and O. K. Tonguz, "Analysis of hot-potato optical networks with wavelength conversion," *J. Lightwave Technol.*, vol. 17, pp. 525–534, Apr. 1999.

[8] I. Chlamtac *et al.*, "CORD: Contention resolution by delay lines," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 1014–1029, June 1996.

[9] C. Gauger, "Dimensioning of FDL buffers for optical burst switching nodes," in *Proc. Optical Network Design and Modeling*, Torino, Italy, 2002.

[10] A. S. Acampora and I. A. Shah, "Multihop lightwave networks: A comparison of store-and-forward and hot-potato routing," *IEEE Trans. Commun.*, vol. 40, pp. 1082–1090, June 1992.

[11] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit, "All-optical packet-switched networks: A study of contention resolution schemes in an irregular mesh network with variable-sized packets," in *Proc. SPIE OptiComm*, Dallas, TX, Oct. 2000, pp. 235–246.

[12] V. M. Vokkarane, J. P. Jue, and S. Sitaraman, "Burst segmentation: An approach for reducing packet loss in optical burst-switched networks," in *Proc. IEEE ICC*, vol. 5, New York, May 2002, pp. 2673–2677.

[13] A. Detti, V. Eramo, and M. Listanti, "Performance evaluation of a new technique for IP support in a WDM optical network: Optical composite burst switching (OCBS)," *J. Lightwave Technol.*, vol. 20, no. 2, pp. 154–165, Feb. 2002.

[14] A. Demers, S. Keshav, and S. Shenker, "Analysis and simulation of a fair queuing algorithm," in *ACM Comput. Commun. Rev.*, 1989, pp. 3–12.

[15] C. Dovrolis and P. Ramanathan, "A case for relative differentiated services and the proportional differentiation model," *IEEE Network*, vol. 13, pp. 26–34, Oct. 1999.

[16] C. Dovrolis, D. Stiliadis, and P. Ramanathan, "Proportional differentiated services: Delay differentiation and packet scheduling," *IEEE/ACM Trans. Networking*, vol. 10, pp. 12–26, Feb. 2002.

[17] C. Dovrolis and P. Ramanathan, "Dynamic class selection: From relative differentiation to absolute QoS," in *Proc. IEEE ICNP 2001*, Nov. 11–14, 2001, pp. 120–128.

[18] Y. Chen, M. Hamdi, D. H. K. Tsang, and C. Qiao, "Proportional differentiation – A scalable QoS approach," *IEEE Commun. Mag.*, vol. 41, pp. 52–58, June 2003.

[19] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-Over-WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.

[20] F. Poppe, K. Laevens, H. Michiel, and S. Molenaar, "Quality-of-service differentiation and fairness in optical burst-switched networks," in *Proc. SPIE OptiComm*, vol. 4874, Boston, MA, July 2002, pp. 118–124.

[21] Y. Chen, M. Hamdi, and D. H. K. Tsang, "Proportional QoS over OBS network," in *Proc. IEEE GLOBECOM*, vol. 3, San Antonio, TX, Nov. 2001, pp. 1510–1514.

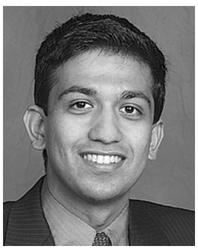
[22] C.-H. Loi, W. Liao, and D.-N. Yang, "Service differentiation in optical burst switched networks," in *Proc. IEEE GLOBECOM*, vol. 3, Taipei, Taiwan, R.O.C., Nov. 2002, pp. 2313–2317.

[23] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: An approach to high bandwidth optical WAN's," *IEEE Trans. Commun.*, vol. 40, pp. 1171–1182, July 1992.

[24] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. Networking*, vol. 1, pp. 397–413, Aug. 1993.

Qiong Zhang (S'03) received the B.S. degree from Hunan University, China, in 1998 and the M.S. degree from the University of Texas, Dallas, in 2000, both in computer science. She is currently working toward the Ph.D. degree in computer science at the University of Texas, Dallas.

Her research interests include quality of service in optical burst-switched networks and secure multicast communication.



Vinod M. Vokkarane (S'02–M'04) received the B.Eng. degree with honors in computer science and engineering from the University of Mysore, India, in 1999, and the M.S. and Ph.D. degrees in computer science from the University of Texas, Dallas, in 2001 and 2004, respectively.

He is an Assistant Professor in the Department of Computer and Information Science at the University of Massachusetts at Dartmouth. His primary areas of research are in optical and wireless networking, with a special focus on the design of architectures and protocols for optical networks.

Dr. Vokkarane is a recipient of the University of Texas at Dallas Best Dissertation Award (2003-2004).

Jason P. Jue (M'99–SM'04) received the B.S. degree in electrical engineering from the University of California, Berkeley, in 1990, the M.S. degree in electrical engineering from the University of California, Los Angeles in 1991, and the Ph.D. degree in computer engineering from the University of California, Davis, in 1999.

He is an Associate Professor in the Department of Computer Science at the University of Texas, Dallas. His research interests include optical networks, network control and management, and network survivability.

Biao Chen (S'96–M'99) received the B.S. degree in computer science from Fudan University, Shanghai, China, in 1988, the M.S. degree in mathematics, and the Ph.D. degree in computer science from the Texas A&M University, College Station, in 1992 and 1996, respectively.

He is an Assistant Professor in the Department of Computer Science, University of Texas, Dallas. His research interests include communications, protocols, network resource management, fault tolerance, and performance analysis.

Dr. Chen received the Best Paper Award at the 12th IEEE International Conference on Distributed Computing Systems (ICDCS). He is a Member of Sigma Xi.