Adaptive Path Selection in OBS Networks

Li Yang and George N. Rouskas, *Senior Member, IEEE*

***Abstract*—In this paper, the authors investigate the concept of adaptive path selection in optical burst-switched networks and its potential to reducing the overall burst drop probability. Specifically, the authors assume that each source maintains a (short) list of alternate paths to each destination and uses information regarding the recent congestion status of the network links to rank the paths; it then transmits bursts along the least congested path. The authors present a suite of path selection strategies, each utilizing a different type of information regarding the link congestion status, and evaluate them using simulation. The results demonstrate that, in general, adaptive path selection outperforms shortest path routing, and, depending on the path strategy involved, the network topology, and the traffic pattern, this improvement can be significant. A new framework for the development of hybrid (or meta) path selection strategies, which make routing decisions based on a weighted combination of the decisions taken by several independent path selection strategies, has been presented. This paper presents two instances of such hybrid strategies, i.e., 1) one that assigns static weights and 2) one that dynamically adjusts the weights based on feedback from the network; it has been shown that these strategies can further improve the overall burst drop probability in the network.**

***Index Terms*—Optical burst switching, path switching.**

I. INTRODUCTION

OPTICAL burst switching (OBS) is a promising switching paradigm that aspires to provide a flexible infrastructure for carrying future Internet traffic in an effective yet practical manner. The transmission of each burst is preceded by the transmission of a setup message (also referred to as burst-header control message), whose purpose is to reserve switching resources along the path for the upcoming data burst. An OBS source node does not wait for confirmation that an end-to-end connection has been set up; instead, it starts transmitting a data burst after a delay (referred to as "offset"), following the transmission of the setup message. For a recent overview of the breadth and depth of current OBS research, the reader is referred to [4].

One of the most important issues in OBS networks is that of burst loss due to congestion caused by transient or permanent overload. Therefore, appropriate mechanisms must be in place in an OBS network to manage the increased demand for resources during a period of congestion. Such mechanisms can be implemented either inside the network (i.e., at OBS switches) or at the source nodes where bursts originate. At the OBS switches, contention-resolution mechanisms can be employed to alleviate the effects of congestion. Contention-resolution schemes can be based on one of four orthogonal approaches or a combination thereof: buffering, wavelength conversion, burst segmentation, or deflection [16], [21]. All these approaches require additional hardware or software components at each OBS switch, increasing their cost significantly; furthermore, practical implementations of some of these components require technology that may be several years from maturity. For instance, optical buffering solutions based on fiber delay lines are not currently cost-effective or scalable. Similarly, although wavelength conversion has been shown to be quite effective in alleviating output port contention [21], wavelength converters are expensive and complex devices, and this state of affairs is expected to continue in the foreseeable future.

Burst segmentation [16] refers to the process of discarding parts of a burst that overlap with another burst at the output port of an OBS switch. Burst segmentation has been proposed to improve the data loss due to contention, as well as a mechanism for providing differentiated quality of service (QoS) in an OBS network. However, segmentation also faces technological challenges, such as the ability to optically detect segments in a burst, to accurately truncate a burst in a way that the remaining data can be recovered at the receiver, and to signal downstream nodes of the reduced burst length. Deflection routing [3] is another mechanism that can be used to reduce the burst loss due to output port contention. In this approach, each switch maintains several paths to a destination, with one path designated as primary (default). When the primary path of an incoming burst is not available, the switch deflects the burst to one of the secondary paths. A deflection routing protocol for OBS networks was proposed in [17], whereas [8] and [22] analyzed the performance of deflection routing. However, deflection routing in OBS networks has several disadvantages. A practical implementation would require intermediate switches that deflect a burst to somehow increase its offset, an operation that is impossible without the use of buffers (alternatively, each burst must have an offset large enough to account for all possible deflections in its path, severely degrading the performance of the network). When deflection decisions are made at each switch without coordination with the rest of the network (a typical approach given the limited amount of time between the setup message and the data burst), there is great potential for routing loops that can have disastrous effects in an optical network [9]. Finally, deflection routing is by nature suboptimal since it only considers the congestion of the current switch, not the state of the links further along the path; in addition, it may cause undesirable vibration effects, as explained in [22].

Manuscript received February 4, 2006; revised May 4, 2006. This work was supported in part by MCNC-RDI as part of the Jumpstart project and by the NSF under Grant ANI-0322107.

L. Yang was with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695-7534 USA. She is now with Scalable Networks, Los Angeles, CA 90045 USA.

G. N. Rouskas is with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695-7534 USA (e-mail: rouskas@ncsu.edu).

Digital Object Identifier 10.1109/JLT.2006.878087

It is also possible to employ a different set of mechanisms at the edge of the network (i.e., at ingress nodes or the burst sources) to reduce the level of burst contention inside the network. These approaches typically require only minimal support from the signaling protocol (e.g., feedback regarding the status of burst transmissions) and can be used when it is desirable to simplify the architecture of core OBS switches to improve their scalability and contain the overall network cost. Alternatively, edge and core nodes may each implement their own mechanisms and coordinate to further reduce burst loss.

One edge node strategy that has the potential to improve burst contention significantly, especially when wavelength conversion is unavailable or sparse, is wavelength assignment. A priority wavelength selection algorithm was presented in [18], and a comprehensive study of wavelength selection strategies for OBS networks can be found in [12]. An algorithm for buffering bursts electronically at edge nodes and scheduling them to prevent burst overlap at links inside the network was proposed in [10]. A traffic engineering approach to select paths for source routing to balance the traffic load across the network links was investigated in [13]. Finally, a dynamic scheme for selecting routes at the burst sources was proposed in [15]. Each source maintains a (short) list of alternate paths to each destination and uses link congestion information to rank each path. The source uses the least congested path to transmit its bursts. We note that a similar technique, referred to as "end-to-end path switching," was proposed and evaluated recently for selecting one among a set of Internet paths [11]; the main finding was that path switching can result in improvement in packet loss.

In this paper, we undertake a comprehensive study of adaptive (end-to-end) path selection in OBS networks. Our objective is to develop a methodology for sources to dynamically switch traffic between a predetermined set of paths in a way that minimizes the overall burst drop probability in the OBS network. The remainder of the paper is organized as follows: In Section II, we discuss our assumptions regarding the OBS network we consider in our study. In Section III, we describe pure path switching strategies, each utilizing partial information about the network state, to select one of a set of available paths to route bursts. In Section IV, we develop a framework for combining several path switching strategies into hybrid (or meta) strategies that base their routing decisions on the decisions of multiple individual methods. In Section V, we present simulation results to demonstrate the effectiveness and benefits of adaptive path selection, and we conclude the paper in Section VI.

II. OBS NETWORK UNDER STUDY

An OBS network is composed of users, optical switches, and fibers. Users are devices (e.g., high-speed electronic routers or multiplexers) that generate optical bursts. An optical switch consists of two components, namely 1) an optical cross-connect (OXC), which can optically forward a burst from an input to an output port without optical–electrical–optical (OEO) conversion, and 2) a signaling engine, which processes signaling messages and controls the OXC switching fabric. Optical fiber links interconnect the network of switches and connect each user to one or more edge switches. A burst generated by a user travels past a series of fibers and switches in the OBS network and terminates at another user.

We will use $G = (V, E)$ to denote an OBS network. V is the set of switches, $N = |V|$, and $E = \{\ell_1, \ell_2, \ldots, \ell_M\}$ is the set of unidirectional fiber links, $M = |E|$. Each link in the network can carry burst traffic on any wavelength from a fixed set of W wavelengths, $\{\lambda_1, \lambda_2, \ldots, \lambda_W\}$. We assume that each OBS switch in the network has full wavelength conversion capabilities, which are used in the case of wavelength contention. The network does not use any other contention-resolution mechanism. Specifically, OBS switches do not employ any buffering, either electronic or optical, in the data path, and they do not utilize deflection routing or burst segmentation. Therefore, if a burst requires an output port at a time when all wavelengths of that port are busy transmitting other bursts, then the burst is dropped.

The OBS network employs source routing, in that the ingress switch (source) determines the path of a burst entering the network. The path over which the burst must travel is carried by the setup message that precedes the transmission of the data burst. We assume the existence of a routing algorithm that is capable of computing a set of k alternate paths for each source–destination pair; the number k of such alternate paths is relatively small, i.e., $k = 2{-}4$. Each source node maintains the list of paths for each possible destination and is responsible for selecting the path over which a given burst will travel. Once the source has made a routing decision for a burst, the path is recorded in the setup message, and it cannot be modified by downstream nodes (i.e., no deflection is allowed).

All source nodes use the same "path switching strategy" to make routing decisions on a per-burst basis. A path switching strategy is characterized by the metric used to rank the paths to a certain destination node. In general, the metric is designed to reflect the likelihood that a burst transmitted on a particular path will experience resource contention and be dropped before it reaches its destination. Whenever a new burst is ready, the source node selects the "best" path according to the metric used and injects the burst into the network.

The rank of each path maintained at a source node is updated dynamically based on information regarding the state of the network collected by the node. We assume that the control plane of the OBS network provides support for the collection and dissemination of information required by the path switching strategies. For instance, this information may be part of the feedback the source receives from the signaling protocol regarding the success or failure of each burst transmission; the Jumpstart just-in-time (JIT) signaling protocol was designed to provide such feedback [2]. Alternatively, the OBS switches may collect information and statistics regarding the (long-term) congestion status of their links and use a link-state protocol to disseminate this information to the rest of the network. Since signaling and state dissemination protocols are required for a variety of network functions, the additional overhead due to the path switching strategies we propose in this paper is expected to be only moderate.

As network dynamics change due to shifts in traffic conditions, congestion levels, and the actions of burst sources, information about these changes will be propagated to the edge of the network using the mechanisms discussed above. In turn, the path rankings at the source nodes may be updated to reflect the new state of the network. As a result of adaptive path selection, different bursts between a certain source–destination pair may take different paths through the network. However, we assume that path switching takes place at large timescales relative to the burst transmission (and generation) times. In particular, the rate of path switching depends on network constants, such as the network diameter and the parameters of the state dissemination protocol (e.g., update interval), whose values are in the hundreds of milliseconds (or even seconds). Consequently, at steady state, a (relatively large) number of successive bursts will use the same path before path switching takes place.

III. Pure Path Switching Strategies

A path switching strategy uses information about the current state of the OBS network to select one of a small number of routing paths for transmitting burst traffic between a source–destination pair. There are several different pieces of information that could be used to describe the congestion level in the network (for instance, link utilization and end-to-end path burst drop rate), and there are several ways in which this information can be combined into a metric to rank paths. It is unknown which types of information or what metrics perform best for path switching in terms of burst drop probability. In this section, we present a suite of "pure" path switching strategies that use a single-path selection method.

A. Weighted Bottleneck Link Utilization Strategy (WBLU)

The WBLU strategy ranks paths using information on link utilization. The motivation behind this strategy is to reduce or prevent contention by using paths with less utilized links.

Consider a link ℓ of the OBS network, let $\text{Succ}(\ell, t)$ denote the set of bursts that have successfully traversed link ℓ until time t, and let T_i denote the length of burst i. The utilization $U(\ell, t)$ of link ℓ at time t is defined as

$$U(\ell, t) = \frac{\sum_{i \in \text{Succ}(\ell, t)} T_i}{Wt} \tag{1}$$

where W is the number of wavelengths; at time $t = 0$, we assume that the utilization $U(\ell, 0) = 0$ for all links ℓ.

Consider now a source–destination pair (s, d), and let $\{\pi_z, z = 1, \ldots, m\}$ be the set of m candidate paths for transmitting bursts from node s to node d. Let $\{\ell_k, k = 1, \ldots, |\pi_z|\}$ be the set of links comprising path π_z, which has length (in number of hops) $|\pi_z|$. At time t, the WBLU strategy routes bursts from s to d along the path $\pi_{z^\star(t)}$ whose index $z^\star(t)$ is obtained using the following metric:

$$z^\star(t) = \arg \max_{1 \le z \le m} \frac{1 - \max_{1 \le k \le |\pi_z|} U(\ell_k, t)}{|\pi_z|}. \tag{2}$$

The numerator in the above expression is the available capacity of the bottleneck link in a given path π_z. Therefore, the WBLU strategy routes bursts along the path with the highest ratio of available bottleneck link capacity to path length. By taking the number of hops into account as in (2), we ensure that if the bottleneck link utilization is similar for two paths, then the shortest path is selected for routing; the longer path is preferred only if the utilization of its bottleneck link is significantly lower than that of the shorter one. We note that a similar metric for ranking paths was used in [5] as part of a routing and wavelength assignment algorithm for wavelength-routed networks.

We note that it is possible to either reset the utilization values periodically or not. The latter approach is easier to implement, but it is not adaptive to load fluctuation. Resetting the values periodically provides better performance under dynamically changing traffic since they more accurately reflect the recent state of the network. In this case, the length τ of the update period is an important parameter that must be carefully selected to balance a set of conflicting requirements. A small update interval period may cause path oscillations and lead to unstable network behavior, whereas with a very long interval, this strategy may fail to react to changing demands in a timely manner. In general, the length of the update interval will be a function of the diameter of the network and the specific algorithm used to disseminate the link congestion information. For a discussion of the factors that need to be considered in selecting the update interval, the reader is referred to [15].

Recall that we have made the assumption that shifts in traffic demands take place at longer timescales than the operation of the path switching strategy. Therefore, the update interval is taken to be longer than the timescales we consider in this paper, during which no changes in traffic patterns are considered. Selecting an appropriate value for the update interval under changing demands is outside the scope of this paper.

B. Weighted Link Congestion (WLC) Strategy

The objective of the WLC strategy is to route bursts along the path that is most likely to lead to a successful transmission. To this end, the source uses information on link congestion along each path to infer the burst drop rate of the path. This strategy assumes the existence of a link-state protocol that disseminates information on link congestion.

Let $N_{\text{succ}}(\ell, t)$ [respectively, $N_{\text{drop}}(\ell, t)$] denote the number of bursts that have been successfully transmitted along (respectively, dropped at) link ℓ up to time t. We define the congestion level $c(\ell)$ of link ℓ at time t as the fraction of bursts that have been dropped at the link, i.e.,

$$c(\ell, t) = \frac{N_{\text{drop}}(\ell, t)}{N_{\text{drop}}(\ell, t) + N_{\text{succ}}(\ell, t)}. \tag{3}$$

We assume that at time $t = 0$, $c(\ell, 0) = 0 \forall \ell$.

Let π_z be a candidate path for routing bursts between a source–destination pair (s, d), consisting of links $\ell_1, \ldots, \ell_{|\pi_z|}$. Assuming that link drop probabilities are independent, at time t, the probability that a burst will be dropped along this path can be calculated as

$$b(\pi_z, t) = 1 - \prod_{1 \le i \le |\pi_z|} (1 - c(\ell_i, t)) \tag{4}$$

The WLC strategy routes bursts from s to d along the path $\pi_{z^\star(t)}$ whose index $z^\star(t)$ is obtained using the following metric:

$$z^\star(t) = \arg \max_{1 \le z \le m} \frac{1 - b(\pi_z, t)}{|\pi_z|} \tag{5}$$

As in (2), this metric takes the number of hops of each path into account to ensure that longer paths are preferred over shorter ones only when they offer a substantial improvement in drop probability.

C. End-to-End Path Priority-Based (EPP) Strategy

The EPP strategy is similar in spirit to WLC in that it also attempts to route bursts along paths with low drop probability. However, rather than relying on information on individual link congestion levels to infer the burst drop probability, EPP requires the source to directly measure this probability from feedback messages it receives from the network regarding the status of each burst transmission.

Consider the source–destination pair (s, d), and let π_z be one of the m candidate paths for this pair as before. Let $N_z(t)$ denote the total number of bursts that have been transmitted from s to d on path π_z up to time t. The EPP strategy assigns a priority $\text{prio}(\pi_z, t)$ to path π_z at time t, which is updated each time a new burst is transmitted on this path and is recursively defined as

$$\text{prio}(\pi_z, t) = \begin{cases} 1.0, & t = 0 \\ \frac{\text{prio}(\pi_z, t-1) \times N_z(t-1) + 1}{N_z(t-1)+1}, & \text{burst success at } t \\ \frac{\text{prio}(\pi_z, t-1) \times N_z(t-1)}{N_z(t-1)+1}, & \text{burst failure at } t. \end{cases} \tag{6}$$

$N_z(t)$ is also updated as $N_z(t) = N_z(t-1) + 1$ each time a new burst is transmitted on path π_z, with $N(0) = 0$. In the above expressions, the time index t refers to the time the source receives feedback from the network regarding the outcome (success or failure) of the most recent burst transmission along path π_z; similarly, index $t - 1$ refers to the time feedback was received regarding the immediately previous burst transmission over the same path. The priority of a path remains unchanged in the interval $[t - 1, t)$. There are several ways for the source to receive feedback regarding a burst transmission. Without loss of generality, we assume that the feedback is provided by the signaling protocol, as is the case with the Jumpstart JIT signaling protocol [2].

Note that the numerator of the two fractions in the right-hand side of (6) is the total number of bursts successfully transmitted along this path up to time t. Therefore, the priority of a path in (6) is simply the probability of a successful bursts transmission along this path; hence, the range of path priorities is the interval (0,1). Therefore, at any given instant, the priority of a path is a measure of the likelihood that a burst transmission along this path will be successful.

At time t, the EPP strategy routes bursts from s to d along the path $\pi_{z^\star(t)}$ whose index $z^\star(t)$ is obtained using the following metric:

$$z^\star = \begin{cases} z, & \text{prio}(\pi_z, t) - \text{prio}(\pi_x, t) > \Delta \forall x \neq z \\ \arg\max_{1 \le z \le m} \frac{\text{prio}(\pi_z, t)}{|\pi_z|}, & \text{otherwise} \end{cases} \tag{7}$$

In other words, if there exists some path whose priority at time t is higher by the priority of all paths by an amount at least equal to a threshold Δ, then this path is selected for routing bursts. The threshold Δ reflects the degree of confidence in the selection of a given path for routing paths. If we are sufficiently confident that a path is better than others in terms of burst drop probability, then the selection is based solely on path priorities. Otherwise, we discount the priority of each path by its length, and we select a path based on the discounted priorities.

In our experiments, we have found that, as long as the traffic pattern does not change over time, the path priorities initially oscillate but eventually converge to a certain value. Consequently, at steady state, one of the candidate paths is always used for routing bursts between a given source–destination pair. Although we do not have a proof of convergence, we have observed such convergence over all the experiments we have conducted. This convergence procedure of the path priority values can be thought of as a "dynamic optimization" process: The path priority values for each source–destination pair keep affecting each other until a local minimum is reached. Numerical results to be presented later demonstrate that the paths selected in this manner perform better than shortest path (SP) routing in terms of burst drop probability.

IV. HYBRID PATH SWITCHING STRATEGIES

Each of the pure path switching strategies we described in the previous section uses some information regarding the network state to select one of a set of candidate paths for transmitting bursts between a source–destination pair. At the time a new burst is ready for transmission, the source uses a strategy to make a routing decision. We will say that a decision is "correct" if the outcome is a successful burst transmission, and "wrong" otherwise. In general, a strategy will be correct (i.e., make correct decisions) only some fraction of the time. Furthermore, at a given time and set of circumstances, different strategies may result in different decisions. Each pure path switching strategy uses only one piece of information in reaching a decision, and this information provides only a limited "view" of the network state.

In this section, we focus on hybrid strategies that, at each burst transmission instant, combine the decisions of several pure strategies into an overall decision in the hope of improving the accuracy of the path selection process and improve the overall burst drop probability. In general, a hybrid strategy emulates a set of pure strategies that run independently of each other "on the side." Each time a burst is ready to transmit, the decision of each pure strategy is computed, and the hybrid strategy uses a set of rules for selecting one of the decisions.

The motivation for this approach is to combine the different partial views of the network state in a way that improves the performance. The next section presents a general framework for combining a set of pure strategies, followed by several instantiations of this framework into hybrid strategies.

A. General Framework

The principles underlying the hybrid path switching strategies are based on ideas from the domain of machine learning [6], [7]. Specifically, it has been shown [7] that the "ensemble" decision reached by a set of voters is more accurate than the decision of any individual voter, provided that each voter reaches a decision in a manner that is largely independent of other voters. Consider, for instance, three voters, v_1, v_2, and v_3, who are called upon to make a binary decision. If the three voters use identical (or very similar) methods to reach a decision, then, whenever voter v_1 is wrong, voters v_2 and v_3 are likely to be wrong as well. However, if the errors made by the different voters are not correlated, then, whenever v_1 is wrong, it is possible that voters v_2 and v_3 are correct, so that a majority vote may reach the correct decision. Let us assume for simplicity that each voter has the same error rate $b < 1/2$ and that errors are independent. In this case, the probability that the majority vote will be wrong is equal to the area under the binomial distribution where more than one-half of the voters are wrong; this area diminishes quickly as the number of voters increases.

In the context of path switching in an OBS network, a pure path switching strategy corresponds to a voter, and the selection of a path corresponds to a (routing) decision. A strategy is correct if transmitting the burst over the path selected by the strategy is successful, and it is wrong if the burst is dropped along the path before it reaches its destination. We can think of the overall burst drop probability of a strategy as its "error rate," i.e., the fraction of time the method is incorrect in successfully selecting a path for a burst. Obviously, the drop probability "overestimates" the real error rate of the strategy, since the fact that a burst is dropped along a given path does not necessarily imply that the burst would have been successful had another path been chosen. Each of the pure strategies we have discussed exploits a different piece of information regarding the network state (e.g., bottleneck link congestion or utilization, or path priority) to reach a decision. Although these pieces of information partially overlap in the sense that they are based on observations of the state of network links, they are not identical and provide separate views of the network. Therefore, we expect that making routing decisions by considering several different views simultaneously will lead to better performance in terms of burst drop probability.

In the remainder of this section, we consider a single source–destination pair (s, d). The source node s maintains $m > 1$ candidate paths, $\pi_1, \ldots, \pi_m$, for routing bursts to destination d. Therefore, for ease of presentation and to avoid repetition, we will drop any references to the source–destination pair (s, d). We also emphasize that our observations and hybrid path switching strategies apply similarly to all other source–destination pairs.

To formalize our approach, let us assume that there are n pure path switching strategies available, $S_1, S_2, \ldots, S_n$. A strategy S_i takes as input some information regarding the network state and produces a probability distribution $p_i^{(z)}$ over the indexes of the candidate paths; we shall discuss shortly how this probability distribution is obtained. The probability $p_i^{(z)}$, $z = 1, \ldots, m$, represents the degree of confidence that strategy S_i has in selecting candidate path π_z for routing the burst traffic. Obviously, we have $p_i^{(1)} + p_i^{(2)} + \cdots + p_i^{(m)} = 1$.

A hybrid strategy H assigns a probability distribution q_i over the n pure path switching strategies $S_1, \ldots, S_n$. The probability q_i represents the degree of confidence of the hybrid strategy H that strategy S_i is correct in its selection of a path. Again, we have that $q_1 + q_2 + \cdots + q_n = 1$. Then, the expected confidence of the hybrid strategy in selecting candidate path π_z is

$$E_z = \sum_{i=1}^{n} q_i p_i^{(z)}, \quad z = 1, \ldots, m. \tag{8}$$

Therefore, the decision of the hybrid strategy H is to route bursts along the path $\pi_z^\star$ with the maximum expected confidence, i.e., the one whose index $z^\star$ is given by

$$z^\star = \arg \max_{1 \le z \le m} E_z. \tag{9}$$

A hybrid strategy is characterized by the set of pure path switching strategies it utilizes and the probability distribution q_i it assigns over these strategies. In the following sections, we introduce three hybrid strategies, each of increasing sophistication.

B. Majority Binary Voting (MBV) Strategy

MBV is the simplest hybrid strategy. Let us assume that there are n pure strategies, $S_i, \ldots, S_n$, with n odd. Each strategy S_i makes a binary decision for each of the m paths: whether to select it for routing bursts or not. Formally, the probability distribution $p_i^{(z)}$ returned by each strategy S_i is as follows:

$$p_i^{(z)} = \begin{cases} 1, & S_i \text{ selects path } \pi_z \\ 0, & \text{otherwise} \end{cases} \qquad i = 1, \ldots, n, z = 1, \ldots, m. \tag{10}$$

The path selected by the hybrid MBV strategy is the one with the most votes. We note that this strategy assumes a uniform distribution q_i over the set of strategies $\{S_i\}$.

C. Weighted Nonbinary Voting (WNV) Strategy

The MBV strategy restricts the pure path switching strategies to vote for a single path, i.e., the one ranked the highest based on the metric used by the respective strategy. Nonbinary voting allows each pure strategy S_i to assign a degree of confidence to each candidate path π_z through a probability distribution $p_i^{(z)}$. One straightforward approach to obtaining the probability distribution is to normalize the values $v_i^{(z)}$ (e.g., priority and congestion level) assigned to the various paths by strategy S_i, i.e.,

$$p_i^{(z)} = \frac{v_i^{(z)}}{\sum_{l=1,\ldots,m} v_i^{(l)}}, \quad i = 1, \ldots, n, \quad z = 1, \ldots, m. \tag{11}$$

The WNV strategy further assigns a probability distribution q_i over the set of pure strategies $\{S_i\}$ and reaches a decision using (8) and (9). The main motivation for using a nonuniform distribution q_i is the fact that, as we shall demonstrate shortly, each pure strategy results in a different burst drop probability; furthermore, the relative performance of the various pure strategies depends on system parameters such as the network topology, traffic load, and pattern. Since each strategy has a different error rate, their contribution to the overall decision of the hybrid strategy should be weighted accordingly. In general, as we shall see, the performance of the hybrid strategy depends strongly on the choice of weights, with the best performance achieved when the weights reflect the relative error rate of the pure strategies.

D. Dynamic WNV (DWNV) Strategy

Under the WNV strategy, the probability distribution q_i over the set of pure strategies $\{S_i\}$ remains fixed at all times. One problem with such an approach is the difficulty in appropriately selecting the weights (degrees of confidence) q_i since an inappropriate choice has the potential to result in poor performance. Instead, it would be desirable to have a method for dynamically adjusting the probability distribution q_i in real time in a way that will minimize the overall burst drop probability; in this case, the probability distribution q_i would also converge to the optimal one. We now present a DWNV strategy to achieve this objective.

Let $q(t) = (q_1(t), \ldots, q_n(t))$ be the probability distribution at time t, and let $B(t, q(t))$ be the burst drop probability of the hybrid strategy at time t when the current distribution is $q(t)$. Our objective is to obtain the distribution $q(t+1)$ at time $t+1$ such that the burst drop probability is minimized (recall that the time indexes refer to the times a burst is ready to be transmitted between the given source–destination pair). In other words, we need to select the distribution $q^\star(t+1)$ such that

$$q^\star(t+1) = \arg \min_{q(t+1)} B\left(t+1, q(t+1)\right). \tag{12}$$

Since it is not possible to solve the above optimization problem directly, we employ a heuristic to dynamically update the q-distribution. We assume that the confidence $c_i(t)$ in the decision of a strategy S_i is reversely proportional to its burst drop probability $b_i(t)$ at time t, i.e.,

$$c_i(t) = \frac{1}{b_i(t) + \epsilon}, \quad i = 1, \ldots, n \tag{13}$$

where ϵ is a smoothing value to avoid division by zero when $b_i = 0$. Based on the confidence c_i of choosing strategy S_i, we compute the new weight q_i as

$$q_i(t+1) = \frac{c_i(t)}{\sum_{l=1,\ldots,n} c_l(t)}, \quad i = 1, \ldots, n. \tag{14}$$

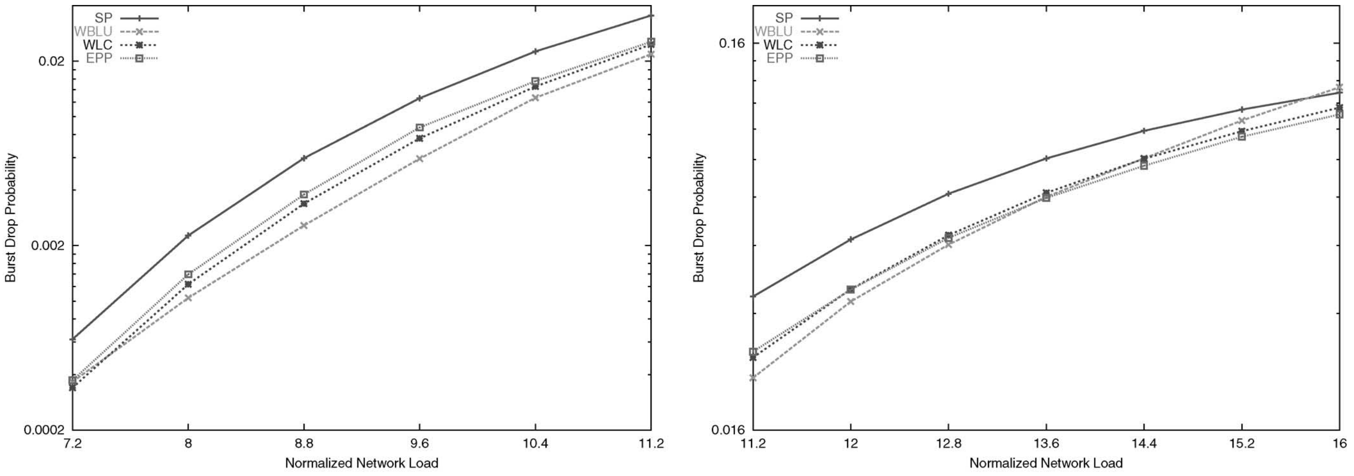
The computation of each of (14) warrants further discussion. The overall burst drop probability $B(t, q(t))$ of the hybrid policy is calculated at the source node using the feedback messages from the network. However, it is not possible for the source node to calculate directly (i.e., based only on feedback from the network) the burst drop probability $b_i(t)$ of each pure strategy S_i as required by (14). To see why a direct calculation of $b_i(t)$ is not possible, consider what happens if the hybrid strategy adopts a decision that is different from the decision of some pure strategy S_i. In this case, the feedback received by the source provides information regarding the decision made by the hybrid policy but no information regarding the decision made by pure strategy S_i; in other words, the source has no way of knowing with certainty whether the burst transmission would have been successful had it used the path selected by S_i instead.

To overcome this difficulty, we use the following approach to compute the burst drop probability $b_i(t)$ for a pure strategy S_i whose decision at time t does not coincide with the decision of the hybrid strategy. Let π be the path chosen by S_i, and let $\text{prio}(\pi, t)$ be the priority of (burst drop probability along) this path; this priority is computed in the course of the operation of the hybrid policy as in (6). Then, we use $\text{prio}(\pi_t)$ to update the drop probability of strategy S_i by making the approximation that the outcome of routing a burst over path π at time t will be failure with probability $1 - \text{prio}(\pi, t)$ and success with probability $\text{prio}(\pi, t)$. Of course, for any pure strategy whose decision coincides with that of the hybrid, the drop probability can be directly updated based on feedback from the network. Therefore, the burst drop probability for any pure strategy S_i, whose decision at time t is to use path π, is updated as follows:

$$b_i(t+1) = \begin{cases} 0, & t = 0 \\ \frac{b_i(t) \times N}{N+1}, & \text{burst success over } \pi \\ \frac{b_i(t) \times N + 1}{N+1}, & \text{burst failure over } \pi \\ \frac{b_i(t) \times N + \left(1 - \text{prio}(\pi, t)\right)}{N+1}, & \pi \text{ was not used.} \end{cases} \tag{15}$$

V. Numerical Results

In this section, we use simulation to investigate the performance benefits of path switching in OBS networks. We use the simulator that was developed as part of the Jumpstart project [14]. The simulator accounts for all the details of the Jumpstart OBS signaling protocol [2] that employs the JIT reservation scheme [19], including all messages required for setting up the path of a burst and feedback messages from the network; the Jumpstart signaling protocol has been implemented in a proof-of-concept testbed on the ATDNet [1]. (We emphasize, however, that the path switching strategies we develop and evaluate in this paper are independent of the specifics of the reservation protocol, and can be deployed alongside either the JET or the Horizon reservation schemes.) We use the method of batch means to estimate the burst drop probability, with



(a) (b)

Fig. 1. Burst drop probability for NSF network, uniform traffic. (a) Low load. (b) High load.

each simulation run lasting until 6×10^5 bursts have been transmitted in the entire network. We have also obtained 95% confidence intervals for all our results; however, they are so narrow that we omit them from the figures we present in this section to improve readability.

We consider two different traffic patterns in our study.

1) Uniform pattern: Each switch generates the same traffic load, and the traffic from a given switch is uniformly distributed to other switches.
2) Distance-dependent pattern: The traffic between a pair of switches is proportional to $|\pi|$ (if $|\pi| = 1$) or $|\pi|/(|\pi| - 1)$ (if $|\pi| > 1$); here, π denotes the shortest path between the pair of switches.

In our simulation experiments, we used a 4 × 4 Torus network based on a regular topology and a 16-node network derived from the 14-node National Science Foundation (NSF) network. All the figures in this section plot the burst drop probability against the "normalized network load" ρ_W, which is obtained by dividing the total load offered to the network by the number W of wavelengths.

A. Pure Path Switching Strategies

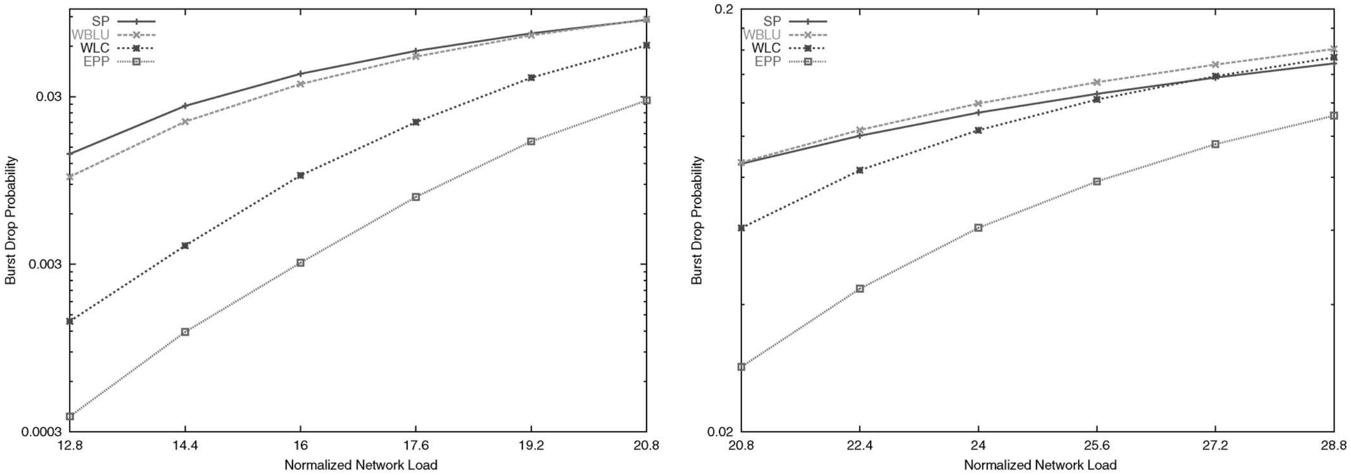
We first investigate the performance improvement that is possible with path switching over SP routing. In our experiments, we assume that each source has to select among $m = 2$ candidate paths to each destination; these are the two shortest link-disjoint paths for the given source–destination pair. We compare four schemes.

1) SP routing: Bursts are routed over the shortest path (in terms of hops) between source and destination, with ties broken arbitrarily.
2) WBLU path switching: Bursts follow the path determined by the WBLU path switching strategy.
3) WLC path switching: Bursts are sent over the path determined by the WLC strategy.
4) EPP path switching: Bursts are routed over the path determined by the EPP strategy.

Fig. 1(a) and (b) plots the burst drop probability of the above four routing schemes for the NSF network with uniform traffic. Fig. 1(a) [respectively, Fig. 1(b)] plots the burst drop probability for low (respectively, high) loads. As we can see, all three path switching strategies perform consistently better than SP routing throughout the load range considered in the figures; the only exception is at very high loads, where the high burst drop probability is due to a saturated network. This result demonstrates the benefits of path switching over SP routing.

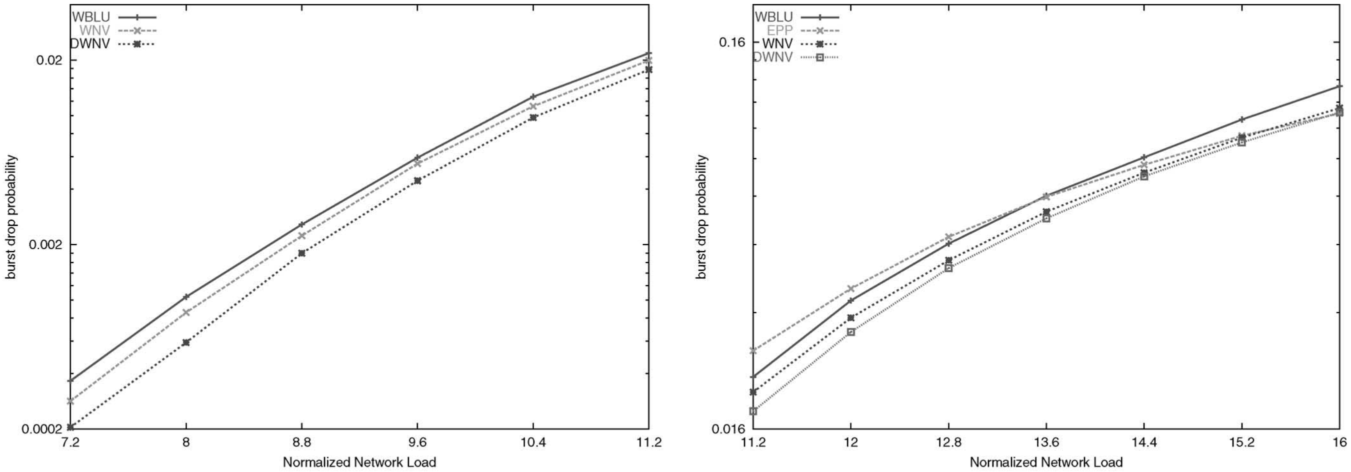
Another important observation from the two figures is that none of the three path switching strategies is a clear winner over the entire range of loads shown. In general, WBLU performs the best at low loads, EPP is the best strategy at high loads, whereas the burst drop probability of WLC is between the values of the other two strategies. Note that at low network loads, most links have low utilization, and avoiding the few highly utilized (bottleneck) links can significantly improve the burst drop probability. Since the WBLU strategy takes account the bottleneck link utilization in determining the burst path, it is not surprising that it performs well at low loads. At high loads, on the other hand, the EPP strategy outperforms the WBLU and WLC strategies. This behavior can be explained by the manner in which the three strategies update their path decisions. Under EPP, path priorities are updated immediately upon the receipt of feedback messages from the network, whereas the WBLU and WLC strategies update their routing decisions periodically (i.e., once they receive the most recent information on link utilization or congestion). The period of update for WBLU and WLC is independent of the network load. With the EPP strategy, however, as the load (i.e., the rate of transmitted bursts) increases, the rate of feedback from the network increases accordingly, providing a more accurate view of the network state and resulting in better routing decisions.

The performance of the four routing methods for the Torus network and uniform traffic is shown in Fig. 2(a) and (b). The WLC and EPP strategies perform consistently better than SP routing, and in fact, the burst drop probability of EPP is significantly lower than that of both WLC and SP across the whole range from low to high loads. The WBLU strategy, on



(a) (b)

Fig. 2. Burst drop probability for Torus network, uniform traffic. (a) Low load. (b) High load.



(a) (b)

Fig. 3. Burst drop probability for NSF network, uniform traffic. (a) Low load. (b) High load.

the other hand, is only slightly better than SP at low loads and slightly worse than SP at high loads. This result can be explained by the fact that WBLU makes a routing decision based only on the utilization of the bottleneck link. In a symmetric topology such as the Torus, the WBLU strategy leads to routing oscillations, which tend to hurt the overall performance. We have observed that the oscillations persist throughout the simulation and that they become worse as the offered load increases. In the asymmetric NSF network, on the other hand, we have observed that the routing decision of WBLU oscillates at first, but it later settles down to a fixed path. The only exception is at very high loads when the bottleneck links are saturated, in which case WBLU keeps oscillating among the candidate paths; this is reflected in Fig. 1(b) for a load of 16, when WBLU performs worse than SP routing.

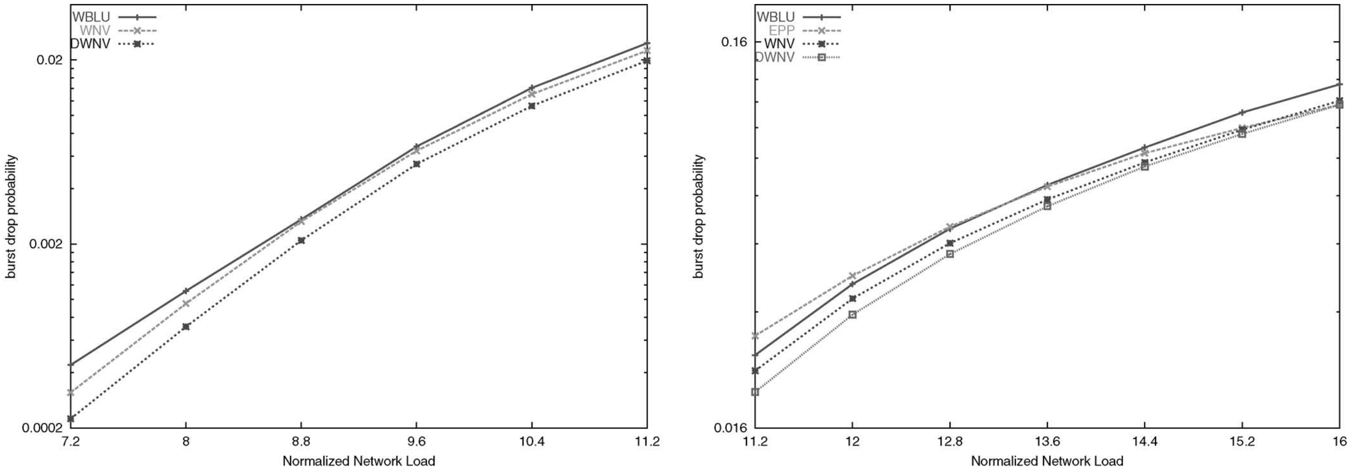
Finally, we note that the relative performance of the four routing methods under the distance-dependent traffic pattern is similar to that under uniform traffic; the relevant results can be found in [20].

Based on the results presented in this section, we conclude that an appropriately selected path switching strategy can significantly reduce the burst drop probability over SP routing. However, the relative performance of a given path switching strategy depends on the network topology and load and is difficult to characterize *a priori*. Therefore, in the next section, we investigate hybrid path switching strategies that combine several independent pure path switching methods to provide consistently good performance.

B. *Hybrid Path Switching Strategies*

We now consider the WNV and DWNV hybrid path switching strategies we introduced in Section IV. Each hybrid strategy utilizes four routing strategies in making its decision: SP routing and the WBLU, WLC, and EPP pure path switching strategies. To characterize the performance of hybrid path switching, in this section, we compare the following three routing schemes.

1) WNV path switching: Bursts are sent along the path with the maximum weighted overall confidence. By definition, WNV assigns static weights (q-distributions) to each of the four pure strategies (voters). We have found



(a) (b)

Fig. 4. Burst drop probability for NSF network, distance-dependent traffic. (a) Low load. (b) High load.

that different weights perform differently for each of the two topologies we consider here. Therefore, after some experimentation, we have used the following weights: For the NSF network, all weights are equal to 1/4 (a uniform distribution), whereas for the Torus network, the weights are 1/8 (SP and WBLU), 1/4 (WLC), and 1/2 (EPP).

2) DWNV path switching: Bursts are sent along the path with the maximum weighted overall confidence. Under this strategy, initially, the weights of all pure strategies are equal, but they are adjusted dynamically during the operation of the network as we explained in Section IV.
3) Best pure strategy: Bursts are sent along the path determined by the pure strategy with the best performance among the four strategies, namely 1) SP, 2) WBLU, 3) WLC, and 4) EPP. Note that, if it turns out that one pure strategy is best across some range of loads while another strategy is best across a different range, we will present both strategies.

Fig. 3(a) and (b) compares DWNV, WNV, and the two best pure path switching strategies for the NSF network with uniform traffic; note that WBLU has the best performance among the pure path switching strategies at low loads, whereas EPP is the best pure strategy at high loads, which is consistent with the behavior we observe in the previous section. We also observe that the hybrid WNV path switching scheme improves the burst drop probability over both pure strategies; in effect, the WNV curve tracks the best of the WBLU or EPP curves. This result confirms our intuition that combining and taking into account several different views of the network state increases the performance. We also note that when the weight of each pure strategy is adjusted dynamically to reflect the real-time network performance, as accomplished by the DWNV hybrid strategy, the burst drop probability is further improved. Our experiments demonstrate that through dynamic adjustments, the weights assigned to each pure strategy by a source–destination pair are tuned to prefer one (or the combination of more than one) pure strategy. Which strategies are preferred depends on the source–destination pair (i.e., path through the network) and the traffic pattern. This tuning procedure can be viewed as a dynamic optimization process that allows the hybrid DWNV strategy to achieve a final set of weights (q-distribution) that is near optimal in the sense of minimizing the burst drop probability.

Similar observations to the above can be made from Fig. 4(a) and (b), which presents results for the NSF network with the distance-dependent traffic pattern. The results for the Torus network are also similar, and we omit them; they can be found in [20].

Overall, we can summarize our results as follows.

- Pure path switching strategies can reduce the burst drop probability in an OBS network as compared with SP routing.
- The performance improvement depends on various parameters, including the congestion information utilized by the path switching strategy, the network topology, the traffic pattern, and the network load; in many cases, the performance improvement over SP routing can be dramatic.
- Hybrid path switching strategies can be used to further improve the network performance. However, if there is a single pure switching strategy that clearly outperforms all others for a given set of parameters within the network's operating region, then a hybrid strategy may not provide any improvement (since it relies on strategies other than the best one). In this case, it is best to simply use the most successful pure strategy instead.
- If it is possible to appropriately select the weights assigned to the pure strategies by a hybrid strategy, then the overall network performance is optimized. Otherwise, a hybrid strategy that dynamically adjusts the weights performs best.

VI. CONCLUDING REMARKS

We have considered the problem of multipath routing in OBS networks with the objective of improving the burst drop probability. We have developed a suite of path switching strategies, each utilizing one type of dynamic information regarding the network state to select one of a set of paths to route a given burst. We also developed a probabilistic framework for hybrid path switching strategies, which make routing decisions by taking into account the decisions of multiple pure path switching strategies. We presented two instances of such hybrid strategies: one static and one dynamic. Experimental results have shown that the pure path switching strategies perform significantly better than SP routing and that hybrid strategies can further improve performance in terms of burst drop probability.

REFERENCES

[1] I. Baldine, M. Cassada, A. Bragg, G. Karmous-Edwards, and D. Stevenson, "Just-in-time optical burst switching implementation in the ATDnet all-optical networking testbed," in *Proc. Globecom*, San Francisco, CA, Dec. 2003, vol. 5, pp. 2777–2781.

[2] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "JumpStart: A just-in-time signaling architecture for WDM burst-switched networks," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 82–89, Feb. 2002.

[3] J. Bannister, F. Borgonovo, L. Fratta, and M. Gerla, "A performance model of deflection routing in multibuffer networks with nonuniform traffic," *IEEE/ACM Trans. Netw.*, vol. 3, no. 5, pp. 509–520, Oct. 1995.

[4] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: A new area in optical networking research," *IEEE Netw.*, vol. 18, no. 3, pp. 16–23, May/Jun. 2004.

[5] X-W. Chu, B. Li, and Z. Zhang, "A dynamic RWA algorithm in a wavelength-routed all-optical network with wavelength converters," in *Proc. IEEE INFOCOM*, 2003, pp. 1795–1802.

[6] T. G. Dietterich, "Machine learning research: Four current directions," *AI Mag.*, vol. 18, no. 4, pp. 97–136, Apr. 1997.

[7] L. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 10, pp. 993–1001, Oct. 1990.

[8] C. Hsu, T. Liu, and N. Huang, "Performance analysis of deflection routing in optical burst-switched networks," in *Proc. Infocom*, 2002, pp. 66–73.

[9] J. Iness, B. Ramamurthy, B. Mukherjee, and K. Bala, "Elimination of all-optical cycles in wavelength-routed optical networks," *J. Lightw. Technol.*, vol. 14, no. 6, pp. 1207–1217, Jun. 1996.

[10] J. Li and C. Qiao, "Schedule burst proactively for optical burst switching networks," *Comput. Netw.*, vol. 44, no. 5, pp. 617–629, Apr. 2004.

[11] S. Tao, K. Xu, Y. Xu, T. Fei, L. Gao, R. Guerin, J. Kurose, D. Towsley, and Z.-L. Zhang, "Exploring the performance benefits of end-to-end path switching," in *Proc. ACM SIGMETRICS*, New York, 2004, pp. 418–419.

[12] J. Teng and G. N. Rouskas, "On wavelength assignment in optical burst switched networks," in *Proc. BROADNETS*, Oct. 2004, pp. 24–33.

[13] ——, "Routing path optimization in optical burst switched networks," in *Proc. 9th ONDM*, Feb. 2005, pp. 1–10.

[14] J. Teng, "A study of optical burst switched networks with the jumpstart just-in-time signaling protocol," Ph.D. dissertation, North Carolina State Univ., Raleigh, NC, Aug. 2004.

[15] G. P. V. Thodime, V. M. Vokkarane, and J. P. Jue, "Dynamic congestion-based load balanced routing in optical burst-switched networks," in *Proc. Globecom*, Dec. 2003, vol. 5, pp. 2628–2632.

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

[17] X. Wang, H. Morikawa, and T. Aoyama, "Burst optical deflection routing protocol for wavelength routing WDM networks," in *Proc. Opticomm*, 2003, pp. 120–129.

[18] ——, "Priority-based wavelength assignment algorithm for burst switched WDM optical networks," *IEICE Trans. Commun.*, vol. E86-B, no. 5, pp. 1508–1514, May 2003.

[19] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *J. Lightw. Technol.*, vol. 18, no. 12, pp. 2019–2037, Dec. 2000.

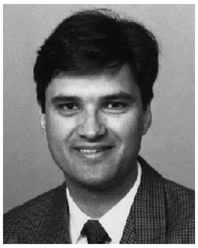
[20] L. Yang, "Congestion control and quality-of-service (QoS) provisions for optical burst switched networks," Ph.D. dissertation, North Carolina State Univ., Raleigh, NC, May 2006.

[21] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *J. Lightw. Technol.*, vol. 21, no. 3, pp. 672–683, Mar. 2003.

[22] A. Zalesky, H. L. Vu, Z. Rosberg, E. M. Wong, and M. Zukerman, "Reduced load Erlang fixed point analysis of optical burst switched networks with deflection routing and wavelength reservation," in *Proc. 1st Int. Workshop Opt. Burst Switching*, Oct. 2003.

Li Yang received the B.S. degree in computer engineering from Xi'an Institute of Posts and Telecoms, Xi'an, China, in 1997, the M.S. degree in computer science from Beijing University of Posts and Telecoms (BUPT), Beijing, China, in 2000, and the Ph.D. degree in computer Science from North Carolina State University, Raleigh, in 2006.

She was a Software Engineer for a year with Alcatel Shanghai Bell-BUPT Research and Development Center, Beijing. She was a Software Intern with Intel Raleigh Design Center. In 2006, she joined Scalable Networks, Los Angeles, CA, as a Research Scientist.



George N. Rouskas (S'92–M'95–SM'01) received the Diploma in computer engineering from the National Technical University of Athens, Athens, Greece, in 1989 and the M.S. and Ph.D. degrees in computer science from the College of Computing, Georgia Institute of Technology, Atlanta, in 1991 and 1994, respectively.

He is a Professor of computer science with North Carolina State University (NCSU), Raleigh. During the 2000–2001 academic year, he spent a sabbatical term with Vitesse Semiconductor, Morrisville, NC, and in May 2000 and December 2002, he was an Invited Professor with the University of Evry, Evry, France. His research interests include network architectures and protocols, optical networks, multicast communication, and performance evaluation.

Dr. Rouskas is a member of the Association for Computing Machinery (ACM) and the Technical Chamber of Greece. He has been on the editorial boards of the IEEE/ACM Transactions on Networking, *Computer Networks*, and *Optical Networks*. He was a Co-Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Protocols and Architectures for Next Generation Optical WDM Networks, published in October 2000. He was the General Co-Chair of the IEEE LANMAN 2005 workshop, Technical Program Co-Chair of the Networking 2004 conference, Program Chair of the IEEE LANMAN 2004 workshop, and Program Co-Chair of the Traffic Grooming workshop 2004. He received the 2004 ALCOA Foundation Engineering Research Achievement Award and the 2003 NCSU Alumni Outstanding Research Award. He was a recipient of the 1997 NSF Faculty Early Career Development (CAREER) Award, a coauthor of a paper that received the Best Paper Award at the 1998 SPIE conference on All-Optical Networking, and the recipient of the 1994 Graduate Research Assistant Award from the College of Computing, Georgia Institute of Technology. He is especially proud of his teaching awards, including his induction in the NCSU Academy of Outstanding Teachers in 2004 and the Outstanding New Teacher Award he received from the Department of Computer Science in 1995.