Forward Resource Reservation for QoS Provisioning in OBS Systems

Jingxuan Liu and Nirwan Ansari
Advanced Network Laboratory
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, NJ 07102-1982, USA.

Abstract– This paper addresses the issue of providing QoS services for Optical Burst Switching (OBS) systems. We propose a Linear Predictive Filter (LPF)-based Forward Resource Reservation method to reduce the burst delay at edge routers. An aggressive reservation method is proposed to increase the successful forward reservation probability and to improve the delay reduction performance. We also discuss a QoS strategy that achieves burst delay differentiation for different classes of traffic by extending the FRR scheme. We analyze the latency reduction improvement gained by our FRR scheme, and evaluate the bandwidth cost of the FRR-based QoS strategy. Our scheme yields significant delay reduction for time-critical traffic, while maintaining the bandwidth overhead within limits.

Index Terms–*QoS, latency reduction, OBS, resource reservation, linear predictive filter.*

I. INTRODUCTION

Optical Burst Switching (OBS) provides a feasible solution to IP-Over-WDM systems, which support multiple types of traffic, such as audio and data, each with different QoS requirements. It becomes increasingly important to design an OBS system that guarantees QoS provisioning for different classes of traffic. Whereas more bandwidth is being provided in Gigabit optical networks, one must address the latency issue in the next generation network.

The basic idea underlying an OBS system is the separation of the transmission and switching of a control header and its coordinating data burst. A data burst is assembled from IP packets at the network ingress, at the timescale of hundreds of microseconds. A control header, also called a burst header packet (BHP)[1], is transmitted in an earlier time window. It reserves resources and sets up a switching path at least before its data payload enters a switching node in the core network, enabling a data burst to be transmitted transparently throughout the core network. The delay on a data burst thus mainly consists of three components: burst assembly delay at the edge routers, path-setup delay caused by the control headers, and propagation delay in the core network.

There have been numerous proposals in the literature focusing on the latency reduction issue in OBS systems. For example, a typical OBS system features one-way reservation that lowers the round-trip delay for signaling transmission. Wei *et al.* [2] proposed a just-in-time (JIT) protocol to reduce burst delay due to lightpath-setup. Xiong *et al.* [1] discussed the optimal switching architectures of the core routers to process control headers. All these strategies are focused on reducing the latency in the core network.

We observe that the bandwidth at the core network (OC192 and upper) is much higher than that in the edge network (OC3-OC48). The time for burst assembly therefore has a significant impact on the end-to-end burst delay. This is especially the case for applications with strict delay constraints (e.g., Internet telephony and videoconferencing). Hence, reducing burst delay at the edge routers will be greatly beneficial to latency reduction and QoS provisioning.

For this reason, we propose in this paper an innovative scheme, called Forward Resource Reservation (FRR), for latency reduction at the edges of an OBS system. The FRR scheme is further extended to achieve controllable QoS differentiation on burst delay for different classes of traffic. Theoretical analysis and simulation results show that our scheme substantially reduces the burst delay caused by the burst assembly at edge nodes, while maintaining the bandwidth wastage of the system within limits.

The rest of this paper is organized as follows: Section II describes the network model, the basic FRR approach, and the FRR-based QoS strategy. Section III analyzes the performance and presents the simulation results. We conclude in Section IV.

II. FRR SCHEME AND QOS SUPPORT IN OBS SYSTEMS

This section describes the system architecture in which the FRR-based QoS strategy applies. We explain our proposed FRR scheme and one of its important features, namely the aggressive resource reservation. Then we present the QoS strategy that can be incorporated in the FRR scheme.

A. System Model

As stated above, the essence of an OBS system is the decoupling of the BHP and the data payload. Fig. 1 highlights the architecture of an OBS network under investigation.

In our scenario, we employ the burst assembly mechanism described in [3], where incoming IP packets of the same destination and attributes, e.g., QoS requirements, are aggregated at edge nodes. When a predefined threshold is reached (e.g., a timer expires), a new burst is generated and is ready to be sent into the core network.

We assume that the lightpath is set up and reserved for a burst according to the RFD (reserve-a-fixed-duration) approach, e.g., the just-enough-time (JET) protocol [4]. In this scenario, a BHP has the knowledge of its payload, including the burst length. This scenario enables a BHP to reserve resources for a proper duration that corresponds to the burst length, and thus delivers efficient bandwidth utilization.

We further assume that traffic in the system is partitioned into two classes depending on their QoS requirements: real-time traffic that has a stringent burst delay constraint (denoted as class-*0*), and non-real-time traffic that is delay-tolerant (denoted as class-*1*). The QoS requirement considered in this

0-7803-7632-3/02/$17.00 ©2002 IEEE

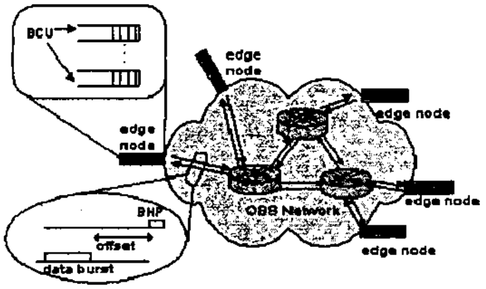

Fig. 1. The system model

paper is the burst delay. The traffic with the lower delay requirement maps to the lower class of service.

A brief summary of the design objectives of our system is:

1. A BHP specifies a reservation duration, which corresponds to the length of its data payload;
2. While preserving the all-optical lightpath advantage for its payload, a BHP should enable the data burst to be transmitted as early as possible, thus minimizing the latency at edge nodes;
3. The system can behave differently for different classes of traffic to achieve service differentiation in terms of the burst delay.

Our FRR scheme meets the first two requirements by an LPF-based method, and is extended to facilitate the QoS capability required by the third one.

B. *Basic Forward Resource Reservation*

To explain the FRR scheme, we first make the following notations to simplify our description ($i = 0,1$):

T_b^i: The time when a new burst of class-*i* traffic begins to assemble at an edge node;

T_h^i: The time when a class-*i* BHP is sent into the core network;

T_d^i: The time when a class-*i* data burst is sent into the core network;

τ_a^i: The duration to assemble a burst of class-*i* traffic;

τ_o^i: The offset between a class-*i* BHP and its data payload. τ_o^i can adopt a pre-existing protocol that is most convenient in the system, e.g., it may be the BHP processing time (end-to-end) in the core network [2], or it may support some QoS capability, such as that described in [4].

In part of this paper, we will discuss the behavior and performance of only the traffic class to which the FRR scheme applies, and thus, for notational simplicity, omit the referencing of the class in this case.

An FRR scheme involves a three-step procedure as follows:

- *Phase 1*: *Prediction.* Before a data burst begins to assemble, the burstification control unit (BCU) predicts the reservation length for the incoming data burst. This estimation is derived from an LPF-based method, as will be discussed in the next subsection.
- *Phase 2*: *Pre-transmission.* As soon as a burst begins to assemble at an edge node, i.e., when the first bit of the first packet in a burst arrives at the burst assembly queue at time T_b, the BCU fills the information necessary for path setup, including the reservation length, into a BHP. The BHP is then sent into the core network at time T_h ($T_h = \max\{T_b,\ T_b + \tau_a - \tau_o\}$).
- *Phase 3*: *Examination.* When the burst assembly finishes, the actual burst length is compared with the reservation length in the pre-transmitted BHP. One of the following cases may occur:
    - i) If the actual burst length is less than or equal to the pre-reserved length, i.e., the BHP has reserved enough bandwidth for the data payload, the BHP pre-transmission is deemed a success. In this case, the data burst is sent into the core network at $T_d = T_h + \tau_o$.
    - ii) If the actual burst duration exceeds the reservation length, the BHP pre-transmission is deemed a failure. The BHP has to be re-transmitted for this burst at a later time of $T_b + \tau_a$ with the actual burst size, and the data payload lags behind by the offset τ_o.

Fig. 2 depicts the principle of the basic FRR scheme when $\tau_a > \tau_o$ and the pre-transmission of a BHP succeeds.

C. *Aggressive Resource Reservation*

The FRR scheme requires *a priori* knowledge of the burst length before it is fully assembled. This is made possible with an *N*-order LPF. Let $L_d(k)$ be the length (in the time scale) of the *k*-th burst, then the length of the next incoming burst is predicted according to the lengths of the previous *N* bursts by:

$$\tilde{L}_d(k+1) = \sum_{i=1}^{N} w(i) \cdot L_d(k-i+1), \qquad (1)$$

where $w(i)$, $i \in \{1,\ldots,N\}$, are the coefficients of the adaptive filter, and are updated by the normalized LMS (Least Mean Square) algorithm [5].

A control header makes an advance resource reservation according to the predicted value. The forward reservation length, denoted as $L_r(k+1)$, if optimal, should be equal to the actual burst length. Due to the imperfection of a predictor, however, an estimated length may turn out to be smaller or larger than the actual burst duration. Suppose the reservation length is set to be equal to the predicted length, a smaller prediction of burst length ($e(k+1) = (L_d(k+1) - \tilde{L}_d(k+1)) > 0$) will result in an insufficient reservation of path holding time for the data burst. This requires the BHP to be re-transmitted after the burst assembly finishes, thus degrading the FRR latency reduction performance.

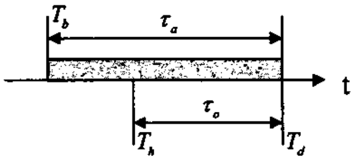

Fig. 2. FRR discipline

This problem is compensated by an aggressive reservation method. Instead of making $L_r(k+1) = \tilde{L}_d(k+1)$, we define the reservation length as $L_r(k+1) = \tilde{L}_d(k+1) + \delta$, where δ is a small margin of correction. δ may be any of the multiple of the sample Root Mean Square (RMS) of the LPF, defined as $\delta = m \cdot \sqrt{\sum_{i=1}^{N} e^2(k-i+1) \Big/ N}$, where m is a real value and is determined according to the tradeoff consideration between the bandwidth cost and the successful BHP pre-transmission probability. Fig. 3 presents the principle of our LPF-based aggressive resource reservation.

D. FRR-Based QoS Provisioning

An intrinsic feature of the FRR scheme, namely the advance transmission of a BHP with an estimated data burst length, facilitates the parallel between the resource reservation and the burst assembly, thereby reducing the burst delay at the ingress nodes. As discussed above, real-time traffic has a higher class and a more stringent constraint for burst delay. To reduce the latency of a class-*0* traffic, and to achieve a flexible QoS differentiation for different classes of applications, we extend the FRR scheme for QoS provisioning in an OBS network, and will refer to it as the FRR-based QoS provisioning.

We present the discipline of our QoS strategy by illustrating the behaviors of BHPs belonging to different traffic classes. For simplicity, we assume the two traffic classes have the same burst assembly time and offset time, i.e., $\tau_a^0 = \tau_a^1$, $\tau_o^0 = \tau_o^1$, and denote them as τ_a and τ_o, respectively.

For a burst of class-*1* traffic (i.e., non-real-time traffic), a simple resource reservation is executed, where a BHP is generated and is sent into the core network when the burst is fully assembled. The BHP carries the actual burst length (Fig. 4(a)).

For a burst of class-*0* traffic, however, an FRR-based process is triggered. A BHP is launched into the core network prior to the burst assembly completion by time τ_p (Fig. 4(b)). The delay of the time-critical traffic at the ingress node is thus decreased ($T_d^0 < T_d^1$). The advanced period τ_p is a system parameter and can be determined from a user or a system perspective. The user could specify the τ_p as a QoS constraint.

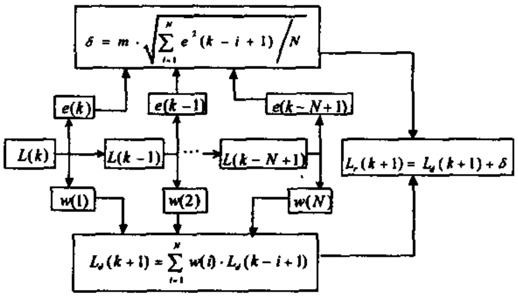

Fig. 3. The prediction and aggressive reservation

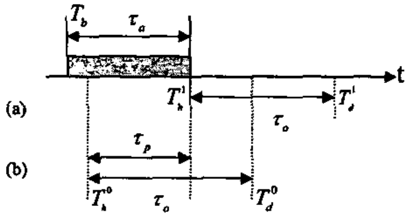

Fig. 4. An example of FRR-based QoS strategy $(\tau_o > \tau_a)$.
(a) class-*1* burst; (b) class-*0* burst.

Alternatively, the network operator can adapt the τ_p as a matter of policy, varying with the differentiation degree requirement between classes. In Fig. 4, $\tau_a > \tau_p$.

Although the BHP pre-transmission may fail due to the insufficiency of a pre-reserved bandwidth, and therefore the transmission of the class-*0* burst becomes the same as that of the low-class traffic, we note that the FRR-based QoS scheme does not increase the end-to-end delay of either traffic class.

III. PERFORMANCE ANALYSIS AND SIMULATION RESULTS

In this section we analyze the system performance in terms of latency reduction improvement, bandwidth overhead, and the impact of the aggressive bandwidth reservation on the BHP pre-transmission success probabilities. For simplicity, we do not consider the wavelength assignment time or the reconfiguration time due to the control header update. We assume they are negligible compared with the burst assembly time and the basic offset. We also assume the BHP advanced time is equal to the burst assembly duration ($T_p = T_a$). Table I summarizes the notations we will use in the analysis.

A. Latency Reduction Improvement

We study the burst delay at the network edge with the simple (called NFRR for No Forward Resource Reservation) or the FRR mode of reservation, and the latency improvement by the FRR scheme. In our two-class QoS system scenario, we focus on the class-*0* traffic to which the FRR scheme applies.

The delay of a data burst is defined as the average delay of all the packets composed of this burst. Therefore, the burst delay due to burst assembly is $1/2 \cdot \tau_a$.

1) Burst Delay in an NFRR system

In an NFRR system, the burst delay at an ingress node due to the burst assembly and the basic offset time is:

TABLE I
NOTATIONS

Term	Explanation
D_n	Average burst delay in an NFRR system
D_a	Average burst delay in an FRR system
D_f	Burst delay when the BHP pre-transmission fails
D_s	Burst delay when the BHP pre-transmission succeeds
P_s	The BHP pre-transmission success probability

$$D_n = \frac{1}{2}\cdot\tau_a + \tau_o. \qquad (2)$$

2) Burst Delay in an FRR system

In a system with an FRR scheme, the burst delay at an ingress node differs according to the success or failure of the pre-transmission of a BHP. If fails, the delay is the same as D_n ($D_f = D_n$). Otherwise, the delay is $D_s = 1/2\cdot\tau_a$ when $\tau_c \geq \tau_o$, or $D_s = 1/2\cdot\tau_a + (\tau_o - \tau_a)$ when $\tau_a < \tau_o$. Suppose the forward resource reservation succeeds with a probability of P_s, the average burst delay of a class-*0* burst therefore is:

$$D_a = P_s\cdot D_s + (1-P_s)\cdot D_f$$
$$= \begin{cases} \frac{1}{2}\cdot\tau_a + \tau_o - \tau_o\cdot P_s & \tau_a \geq \tau_o, \\ \frac{1}{2}\cdot\tau_a + \tau_o - \tau_a\cdot P_s & \tau_a < \tau_o. \end{cases} \qquad (3)$$

Now that the burst delay depends on both τ_a and τ_o, we assume $\tau_o = \mu\cdot\tau_a$, where μ is a real value that represents the ratio of τ_o over τ_a. Hence, the latency improvement (η) of the FRR scheme over the NFRR scheme is given by:

$$\eta = 1 - \frac{D_a}{D_n} = \begin{cases} \frac{2\cdot\mu\cdot P_s}{1+2\cdot\mu} & \mu \leq 1, \\ \frac{2\cdot P_s}{1+2\cdot\mu} & \mu > 1. \end{cases} \qquad (4)$$

The systems performance improvement η depends on two parameters: the ratio of τ_o over τ_a (μ) and the probability that a forward reservation succeeds (P_s). Fig. 5 presents the latency reduction percentage versus P_s, when μ varies. It shows that η increases as τ_o approaches τ_a, and reaches its maximum gain when the ratio is 1. Specifically, if the burst length can be predicted precisely such that the pre-transmission of the BHP succeeds with a high probability ($P_s \rightarrow 100\%$), our FRR scheme can reduce the latency for the high-class traffic by 66% when $\tau_o = \tau_a$. This observation can be further exploited when studying the design issues related to the burst assembly time and the offset values.

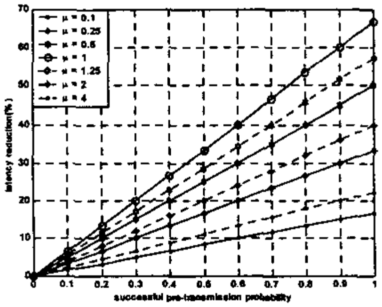

Fig. 5. Latency improvement vs P_s

B BHP Pre-transmission Success Probability

Eq. (4) indicates that the probability that a BHP pre-transmission succeeds (P_s) has an important impact on the latency improvement η. Since P_s depends on, among others, the difference between the pre-reserved duration and the actual burst length ($L_r(k) - L_a(k) = \delta - e(k)$), we study the effect of the correction margin (δ) on P_s.

Conceptually, P_s can be derived from:

$$P_s = P(e(k) < \delta) = \int_{-\infty}^{\delta} f(e(k))de(k), \qquad (5)$$

where $f(e(k))$ is the distribution of the prediction errors ($e(k)$). If we assume $f(e(k))$ could be approximated by a zero-mean Gaussian function with variance equal to σ^2 (further justification of this assumption is omitted due to space constraints), we get $P_s = P(e(k) < \delta) = 1 - Q(\frac{\delta}{\sigma})$, where $Q(\cdot)$ is the Q-function. The theoretical value of P_s, together with the simulation results of P_s under the real IP traffic and the video traffic, is plotted in Fig. 6. We observe that the BHP pre-transmission succeeds at a probability of more than 95%, if $\delta \geq 2\cdot\sigma$, at which point the latency improvement is more than 60% ($\tau_o = \tau_a$), as shown in Fig. 5.

We also conducted a set of simulations tracing the probability density function (PDF) of the number of bursts whose actual lengths differ with the pre-reserved length by a small region of reservation correction. The simulation platform is OPNET. We assume that packets arrive according to a Poisson process. The simulation results are shown in Fig. 7.

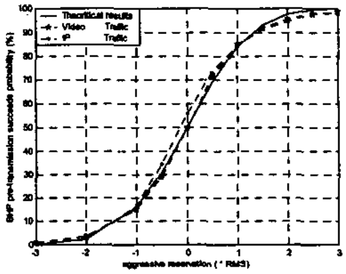

Fig. 6. P_s vs aggressive reservation

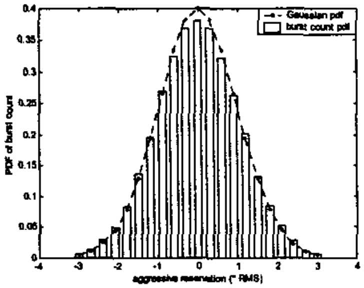

Fig. 7. PDF of burst numbers vs aggressive reservation

For comparison, we also draw the PDF curve of a standard Gaussian distribution function (the dotted line in Fig. 7). It shows that the PDF of the simulation results matches the theoretical curve very well. This also implies that we can achieve controllable successful BHP pre-transmission probabilities as a function of the extra bandwidth reservation.

C *Bandwidth Overhead*

The FRR scheme increases the BHP pre-transmission success probability and improves the latency reduction performance for class-*0* traffic by means of an aggressive reservation. Let γ represent the ratio of the extra reservation length of the aggressive reservation method to the actual burst length. γ can be referred to as the reservation overhead, or the bandwidth overhead. Now consider the bandwidth overhead as a long-term system performance and omit the index of the burst sequence number. This way, an advanced reservation length is simply denoted as L_r, and the estimated burst length and the actual burst length are referred to as L_d and $\tilde{L}_d$, respectively. Let ε and ζ represent the difference between L_d and $\tilde{L}_d$, and that between L_r and L_d, respectively. Then, we have the relationships of $\varepsilon = L_d - \tilde{L}_d$, $\zeta = L_r - L_d$, and $L_r = \tilde{L}_d + \delta$.

By definition, an overhead occurs when $\zeta > 0$, i.e.,

$$\zeta > 0 \Rightarrow L_r - L_d > 0 \Rightarrow \tilde{L}_d + \delta - L_d > 0 \Rightarrow \varepsilon < \delta. \qquad (6)$$

The average ε in this condition, denote as $\bar{\varepsilon}$, is given by:

$$\bar{\varepsilon} = \int_{-\infty}^{\delta} \varepsilon \cdot f(\varepsilon) d\varepsilon, \qquad (7)$$

where $f(\varepsilon)$ is the distribution function of ε. In the same condition, the average burst length for a given $\tilde{L}_d$ is:

$$\bar{L} = \tilde{L}_d + \bar{\varepsilon}. \qquad (8)$$

The bandwidth overhead when $\zeta > 0$ is thus:

$$\bar{\zeta} = L_r - \bar{L} = \delta - \bar{\varepsilon}. \qquad (9)$$

Since the bandwidth overhead due to the aggressive reservation occurs when the pre-transmission of a BHP succeeds. γ is thus:

$$\gamma = \frac{\bar{\zeta}}{\bar{L}} \cdot P_s = \frac{\delta - \bar{\varepsilon}}{\tilde{L}_d + \bar{\varepsilon}} \cdot P_s. \qquad (10)$$

Provided that $f(\varepsilon)$ is a zero-mean Gaussian function with variance σ^2, we get:

$$\bar{\varepsilon} = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \int_{-\infty}^{\delta} \varepsilon \cdot e^{-\frac{\varepsilon^2}{2 \cdot \sigma^2}} d\varepsilon = -\frac{\sigma}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{\delta^2}{2 \cdot \sigma^2}}. \qquad (11)$$

Since we also have $\delta = m \cdot \sigma$, γ can thus be expressed as a function of m:

$$\gamma = \frac{m \cdot \sigma + \frac{\sigma}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{m^2}{2}}}{\tilde{L}_d - \frac{\sigma}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{m^2}{2}}} \cdot (1 - Q(m)). \qquad (12)$$

In our two-class system scenario, suppose the traffic load distribution of the real-time traffic and the non-real-time traffic is $3:7$, then the bandwidth overhead in the system scale is $0.3 * \gamma$, where γ is the bandwidth overhead of the class-*0* traffic, to which the FRR scheme applies. It can be derived from Eq. (12). Fig. 8 illustrates the system bandwidth overhead as a function of m. Both theoretical curves and simulation results are shown.

Although the aggressive reservation method results in a higher probability of successful BHP pre-transmissions at the cost of a bandwidth overhead, the benefit is more considerable, because bandwidth is no longer a limiting factor in the core network, and latency will be the major challenge to overcome in the future [6]. In addition, our FRR scheme gains significant latency reduction at the cost of very small bandwidth overhead in a system scale, as can be seen from Figs 5, 6, and 8.

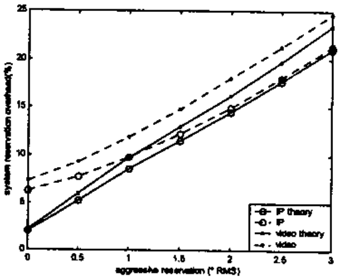

Fig. 8. Bandwidth overhead vs aggressive reservation

IV. CONCLUSIONS

In this paper, a novel FRR scheme has been proposed to reduce the data burst delay at the edge nodes of OBS systems. We have also presented an FRR-based QoS strategy to guarantee the low-delay constraint of the real-time traffic and to offer QoS differentiation in an OBS system. Our FRR-based QoS strategy can dramatically reduce the burst delay due to the burst assembly and the necessary *a priori* all-optical path setup time, while keeping the system bandwidth overhead within limits.

REFERENCES

[1] Y. Xiong, M. Vandenhoute, and H.C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks," *IEEE JSAC*, vol. 18, No. 10, Oct. 2000, pp. 1838-1851.

[2] J.Y. Wei and R.I. Jr. McFarland, "Just-In-Time Signaling for WDM Optical Burst Switching Networks," *IEEE J. of Lightwave Technol.*, vol. 18, No. 12, Dec. 2000, pp. 2091-2037.

[3] A. Ge, F. Callegati, and L. S. Tamil, "On Optical Burst Switching and Self-Similar Traffic," *IEEE Commun. Lett.*, vol. 4, No. 3, March 2000, pp. 98-100.

[4] M. Yoo, and C. Qiao, "A new optical burst switching (OBS) protocol for supporting quality of service," *SPIE Proc. All Optical Commun. Syst.: Architecture, Control Network Issues*, vol. 3531, Nov. 1998, pp. 396-405.

[5] M.L.Honig and D.G. Messerschmitt, *Adaptive Filters Structures, Algorithms, and Applications*, Klumer Academic Publishers, 1990.

[6] D. Farber, "Predicting the unpredictable—The impact of all-optical networking," the Keynote Speech at the IEEE LEOS Summer Topic Workshop on Bandwidth Networking, Monterey, CS, July 1998. http://www.cis.upenn.edu/~farber.