A Virtual One-Way Signaling Protocol With Aggressive Resource Reservation for Improving Burst Transmission Delay

Kyriakos G. Vlachos, *Member, IEEE*, and Demetris Monoyios

***Abstract*—In this paper, we propose a new scheme for the on demand use of capacity in OBS networks, combining a two-way reservation protocol and an assembly scheme process that incorporates an aggressive and forward resource reservation mechanism. The key idea is to tune the assembly timer to be equal to the time associated with the establishment of the end-to-end connection (round-trip-time) and, thus, synchronize the resource reservation with the assembly process. In this way, upon the arrival of the first packet in the queue, reservation of resources may start simultaneously based on an aggressive prediction of the burst length.**

***Index Terms*—Least mean square filter, one-way signaling, optical burst switching (OBS).**

I. Introduction

OPTICAL burst switching (OBS) has been introduced to couple the merits of packet and circuit switching [1]. One of the key developments of OBS was the introduction of one-way reservation schemes for the "on-demand" use of capacity. In one-way reservation schemes (also called "Tell-and-Go"), a setup packet is sent in advance to precede the arrival of a burst of packets by a time offset. This allows for minimizing the pretransmission delay. A number of one-way reservation schemes have been proposed for OBS, including the just-enough-time (JET) [1], horizon [2], and just-in-time (JIT), [3]. One-way schemes are very promising, when applied to a network operating at light load, but may result in a high burst loss ratio when load increases and there is limited or no buffering in the core. Various studies have been carried out to estimate the burst loss ratio when wavelength converters and/or FDL-based optical buffers are employed. In addition, QoS provision schemes were proposed to assure a constant loss ratio. Such schemes include the offset-time-based scheme [4] that provides an extra time offset to isolate different classes of traffic, the composite-burst assembly scheme that mixes traffic classes during burst assembly and provides QoS via prioritized burst segmentation [5], the "*preemptive wavelength reservation mechanism*," where each class is associated with a predefined usage limit, [6] and the "*early dropping mechanism*" that probabilistically drops bursts of a lower priority class in order to guarantee the loss probability of higher priority classes of traffic [7]. However, all these schemes require optical buffers or additional scheduling/processing that makes their deployment difficult. On the other hand, two-way reservation protocols guarantee loss-less operation in a buffer less OBS network. However, two-way signaling induces a large delay, associated with the establishment of an end-to-end connection. A detailed evaluation of two-way OBS in terms of packet delay and buffer requirements in the edge notes can be found in [8].

In this paper, we propose a new scheme that differentiates from the abovementioned ones and which truly emulates one-way reservation. It relies on a two-way reservation protocol and a burst length prediction mechanism. The key idea is to tune the assembly timer to be equal to the time associated with the establishment of the end-to-end connection to synchronize the resource reservation with the assembly process. In this way, upon the arrival of the first packet in the queue, reservation of resources may start simultaneously based on a prediction of the burst length. In our study, we have used an N-order normalized LMS (Least Mean Square) linear predictive filter (LPF) that provides adequate accuracy and has been previously used in OBS [9], [10], while we have further extended its mechanism with an aggressive correction parameter to accommodate bursty traffic. The overall scheme is a *virtual* one-way reservation protocol, in the sense that the burst is transmitted immediately after the assembly timer expires. The advantage of the scheme is that latency is reduced to the minimum possible, burst transmission is guaranteed to be lossless in the core, while data losses may only occur at the edge and only when prediction underestimates burst size.

The rest of the paper is organized as follows. Section II presents the network concept of the proposed virtual one-way scheme, while Section III presents in detail the burst length prediction mechanism and provides a latency comparison with conventional one-way and two-way OBS protocols. Section IV presents evaluation results from static experiments for the determination of the filter convergence speed and accuracy, while Section V presents performance evaluation results over a large scale network with emphasis in achieving packet loss-free operation.

Manuscript received June 02, 2008; revised November 29, 2008. First published April 24, 2009; current version published July 15, 2009. This paper was supported by the BONE-project ("Building the Future Optical Network in Europe"), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

The authors are with the Computer Engineering and Informatics Department and Research Academic Computer Technology Institute, University of Patras, GR26500, Rio, Greece (e-mail: kvlachos@ceid.upatras.gr; dmonogio@ceid.upatras.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2009.2017728

II. Network Concept

OBS networks have been widely associated with the one-way signaling protocols. However, burst losses increase fast with the increase of network load and it is difficult or quite impossible to guarantee a certain level of QoS to end-users. In addition, assuming that each OBS edge router services concurrently thousands active TCP connections, QoS support become an unrivaled task that requires cross layer (transport, network and physical layer) processing. In the proposed framework, a two-way reservation protocol is used in combination with a timer-based assembly scheme, where the timer has been tuned to be equal to the round-trip time delay. Fig. 1 illustrates graphically the conventional OBS approach, while Fig. 2 the proposed one. In particular, Fig. 1(a) shows the usual case of an one-way protocol, where a burst is transmitted after the expiration of the assembly timer. In that case, a setup packet is first transmitted to reserve resources, while burst data follows with a time offset (denoted as TO). Fig. 1(b) shows the case of two-way signaling. Aggregation of packets takes place for a time equal to the assembly time, while burst transmission starts only after an end-to-end wavelength channel has been setup [8]. This includes the time that a control packet has to traverse from source to destination to reserve resources, including the packet processing time and twice the propagation time. It is usually denoted as round trip time ($T_{\rm RTT}$). Over long transmission distances, the round-trip propagation delay may be comparable to, or even larger than, the duration of a burst. For example, the round trip time delay, for a distance of 500 km, correspond to 5 ms in addition to the packet processing time. Assuming a 10-Gb/s channel capacity and that the packet traverses three nodes with 10-ms processing, the total round trip time corresponds to the transmission of 350-Mb of data. In general, the maximum deterministic latency or upper bound on the maximum transmission time that packets experience between entering the core network at the source and leaving the destination routers is

$$T_{\rm BAT} + 2 \cdot T_{\rm propag.} + h \cdot \delta + \frac{L_{\rm burst}}{C} \quad (1)$$

$$T_{\rm BAT} + h \cdot \delta + \frac{L_{\rm burst}}{C} \quad (2)$$

for two-way and one-way reservation protocols respectively. In (1) and (2), $T_{\rm BAT}$ is the assembly time, h is the number of hops, δ the packet processing time (that includes the optical cross-connect, OXC, reconfiguration time), $T_{\rm propag}$ is the end-to-end propagation time and $L_{\rm burst}/C$ is the burst duration time.

Fig. 2 displays the timing constrains of the proposed scheme. The edge router, that in any case maintains a different queue per destination, assigns to each queue, an assembly timer equal to the RTT of that source-destination pair. Upon the arrival of the first packet in the queue [see Fig. 2(a)], a prediction mechanism estimates the size of the queue, at $\rm T_{RTT}$ time later and immediately transmits a setup packet to reserve resources according to that prediction. Packets for that specific destination continue to arrive and are being stored in the same assembly queue. Upon the return of the acknowledge message [see Fig. 2(b)] burst transmission starts immediately without the need of a control packet to precede. Thus, the time offset that is usually incorporated in between transmission and signaling is not needed in our case, and, therefore, latency is further reduced. In particular, the de facto latency bound for the first packet entering the queue is $T_{\rm RTT}$, while only $L_{\rm burst}/C$ for the last packet. The process repeats with a new prediction, upon the arrival of a new (first) packet in the queue [burst #2 in Fig. 2(b)].

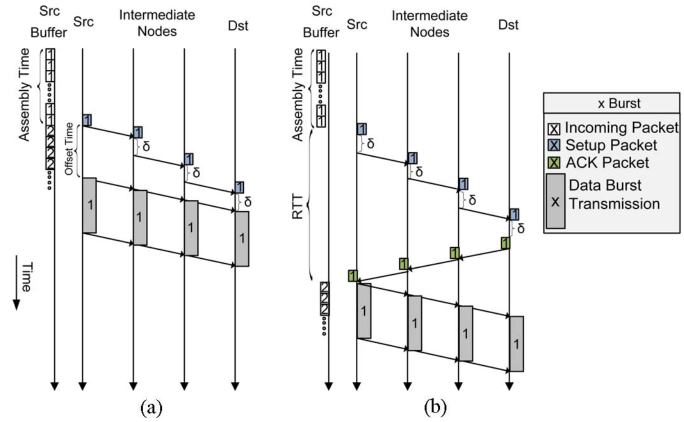


Fig. 1. Timing considerations of the conventional OBS approach with (a) one-way and (b) two-way reservation. δ is the processing time of the setup packet, including the OXC re-configuration time.

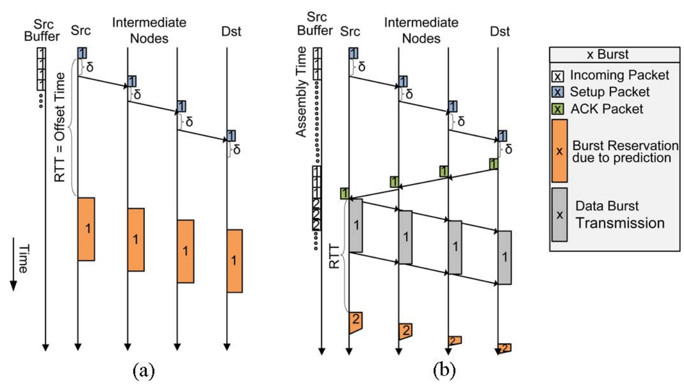


Fig. 2. Timing considerations of the proposed scheme during the (a) reservation phase and (b) the burst transmission phase and reservation of the next burst.

The framework that we propose benefits from the parallel execution of the signaling messages and the assembly process, and for this a setup message requires *a priori* knowledge of the burst length. There is no doubt that to make a prediction algorithm practical for an OBS system, it should not only deliver good estimation performance, but also be simple and fast so the calculations can be done on-line. In general the actual burst size can be different from the predicted one and absolutely depends on the prediction mechanism. When the prediction algorithm underestimates burst length, then the extra packets have to be dropped or delayed for the next assembly cycle. Alternatively, if the prediction algorithm overestimates burst length, then the burst can be transmitted with a percentage of bandwidth being wasted.

III. Burst Length Prediction Mechanism and Aggressive Resource Reservation

The choice of a prediction method is a tradeoff between the prediction interval, prediction error and computational cost. The prediction interval refers to how far into the future can the traffic be predicted with confidence. In our case, the lower bound of this interval is defined by T_{RTT}. For video and network data traffic, linear prediction methods have been considered in the literature as a simple and effective alternative, [9], [10]. Within this framework, we have considered an N-order Normalized LMS (least mean square) linear predictive filter (LPF) that can provide a high accuracy while its time complexity for the coefficient calculation is $O(N)$.

Problem definition: Let $L(t)$ be a random process that gives the number of bytes in the interval $[0, T_{\mathrm{RTT}}]$), estimate the number of aggregated bytes in the next assembly cycle, denoted as $\check{L}(t + T_{\mathrm{RTT}})$.

Let $L_d(k)$ be the length (in the time scale) of the k^{th} burst that corresponds to the k^{th} assembly cycle after $t = k \cdot T_{\mathrm{RTT}}$ time. The length of the next incoming burst is then predicted according to those of the previous N bursts by

$$\check{L}_d(K+1) = \sum_{i=1}^{N} [h(i) \cdot L_d(k-i+1)] \tag{3}$$

where $h(i)$, $i \in \{1, \ldots N\}$ are the coefficients of the N-order LPF. We update the predictive filter coefficients by an efficient algorithm [11], where the coefficients for the $(k+1)^{th}$ prediction are defined as

$$\mathrm{h}^{k+1} = \mathrm{h}^{k} + \mu \cdot e(k) \cdot \mathrm{L_d}^{k} / \|\mathrm{L_d}^{k}\|^2$$

where h is the coefficient vector, μ is an adjustable parameter of the filter, $e(k)$ the residual between the actual and the predicted length of the k^{th} data burst and $(\mathrm{L_d})^k$ the vector of $L_d(j)$, $(j \in \{(k-1) \cdot N + 1, \ldots, k \cdot N\})$. In [10], it has been verified that the LMS-based method provides a low prediction error without knowing the autocorrelation of the input traffic stream in advance and, thus, can be used as an on-line algorithm for each assembly cycle as in the proposed scheme. It must be noted here that only the number of bytes matters and not the particular distribution of packet arrivals and packets sizes within the prediction interval.

A. *Aggressive Burst Length Prediction*

The prediction error, expressed by $e(k)$ can be positive or negative, in the sense that the predicted burst length can be an under- or over-estimate the actual data aggregated. In the first case, part of the data aggregated has to be dropped (or transferred to the next assembly cycle), while in the second case bursts can be transmitted but with a fraction of the reserved capacity being wasted. In general, the prediction error of an LMS-based LPF is less 5% especially for smooth traffic of constant bit rate. However, LPFs cannot accommodate bursty traffic that exhibit fast changes in the packet arrival rate. Thus, it is of key importance to compensate the under-estimation errors, since in that case either the extra packets have to be dropped or to be transmitted in the next assembly cycle, thus increasing their queuing delay. In addition, the prediction mechanism must also be capable of identifying if the error in prediction is the result of a traffic increase or falls within the usual error of the LPF.

For this purpose, we have further extended the filter mechanism to accommodate traffic violations and achieve faster overestimated predictions. We have defined a correction parameter, d, that is added to the prediction value as $L_r(k+1) = \tilde{L}_p(k+1) + d$, where L_r is the reserved length and L_p is the predicted length. The correction parameter d is determined by a general function based on simple remarks, that continuous error signs means that there exist a traffic increase (or decrease), [12], while the magnitude of increase is a function of the previous burst lengths and errors. The estimated adjustment quantity $d(k)$ is added to the LMS prediction value, so that the new predictor could follow the variation of traffic trend more quickly. The general function considered here is as follows:

$$d(k) = \mathrm{sign}(k)^n \times f\left((L_d)^k\right) \times \tilde{e}(k) \tag{4}$$

where $\mathrm{sign}(k)$ is the sign continuity function and is decided based on the sign of the prediction error, $e(k)$ and $\tilde{e}(k)$ is the percentage expression of $e(k)$. For example, if at several continuous moments, $e(k)$ exhibits a negative (or positive) sign, this is probably an indication of a persistent traffic decrease (or increase) in the variation trend. For such cases, the normal LMS-based LPF presents a delay in accommodating the new traffic trend. An exponential weight to the importance of the sign continuity function is given by rising $\mathrm{sign}(k)$ to n. Function $f(\mathbf{L_d})^k$ is a function of the previous N burst lengths, $L_d(j)$, $(j \in \{(\mathbf{k-1}) \cdot N + \mathbf{1}, \ldots, k \cdot N\})$. We have experimented with different statistical functions and we have selected the standard deviation of the average burst sizes denoted here as $\sigma_{\text{mean burst}}$. When combined with the % change in the error value, and sign continuity function, it provides adequate confidence for the next prediction, compensating rapidly, fast increases in the packet arrival rate. Inevitably, it constitutes a tradeoff between large jumps in the prediction (that decrease convergence delay) at the expense, however, of accuracy. It must be noted here that function $f(\cdot)$ should not be relevant to $e(k)$, primarily because statistical properties of $e(k)$ in the past N samples do not reveal any further insight on the data gathered at the network edge and which can be important in OBS networks. For example, the normalization of $e(k)$ may result in a fast tracking performance for small traffic increases, but, however, instant large error values may result to very large errors in the accuracy of the next predictions.

B. *Latency Comparison With One-Way and Two-Way Signaling Protocols*

In usual one-way signaling protocols, the upper bound of packet latency occurs, when all packets arrive simultaneously at the beginning of the assembly cycle, and, thus, it is $T_{\mathrm{BAT}} + (h \cdot \delta)$, excluding for simplicity but without loss of generality the burst transmission. The corresponding average packet waiting time can be easily derived to be: $T_{\mathrm{BAT}}/2 + (h \cdot \delta)$. Similarly, the average and maximum packet delays when using a two-way reservation protocol can be derived from (1). In the proposed virtual one-way reservation scheme with aggressive traffic prediction, the corresponding delays are, by $(h \cdot \delta)$ less, in the sense that the time offset is no needed, but increased by T_{RTT}, in the sense that the under-estimated data left over sent in the next

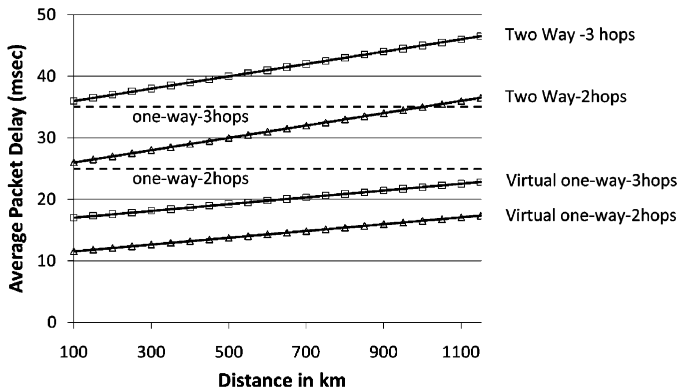


Fig. 3. Average packet delay versus distance for conventional one-way and two way protocols, as well as for the proposed virtual one-way scheme.

cycle. Thus, the upper bound of latency is $2T_{\mathrm{BAT}} = 2T_{\mathrm{RTT}}$ while the average packet waiting time: $T_{\mathrm{RTT}}/2+|\tilde{e}(k)|\cdot T_{\mathrm{RTT}}$, if $e(k) > 0$.

Fig. 3 displays the average packet delay versus the transmission distance for the case a burst is transmitted over two or three hops. In this analysis, we have assumed an assembly timer of 10 ms, (used only in the case of one-way protocols and two-way protocols without prediction), a 5% prediction error (used only in the case of the proposed virtual one-way scheme), 10 ms time for OXC re-configuration, which is a realistic value with current available MEMS switch technology [13] and a 12.5 μs packet processing time as derived in [14] using JITPAC controllers. The dashed lines in Fig. 3, denote the delay of a usual one-way protocol that is constant and independent of the transmission distance. It can be seen that when using a two way reservation protocol (with or without traffic prediction), average packet delay increases with distance and with number of hops. This is as expected, since round trip time, (the return of the ACK message) determines when the burst is transmitted. When comparing, one-way schemes with the proposed scheme, then the latter induces a significantly lower average packet delay. This is due to the fact that the time offset employed in one-way schemes is significant large, while in the proposed scheme this is compensated during assembly time. For example, the average packet delay, for the case of one-way schemes and two hops is: $T_{\mathrm{BAT}}/2 + (h \cdot T_{\mathrm{process}}) \cong 25$ ms, while in the proposed scheme: $T_{\mathrm{RTT}}/2 + |\tilde{e}(k)| \cdot T_{\mathrm{RTT}} = T_{\mathrm{RTT}}(1/2 + |\tilde{e}(k)|) = [20 \text{ ms} + 2L/(2 \cdot 10^8)] \cdot 0,55$, which depends on the path length and becomes comparable with one-way scheme when $T_{\mathrm{BAT}} = 10$ ms and L $= 2550$ km or $T_{\mathrm{BAT}} = 1$ ms and L $= 1800$ km. It is, therefore, clear that the proposed virtual one-way signaling scheme induces a significantly lower average packet delay, and this is because assembly time has been tuned to be equal to round-trip-time.

IV. Prediction Mechanism Evaluation

In order to evaluate the proposed scheme, we have carried out static experiments with a single edge router and Poisson packet arrivals of constant as well as varying mean rates with and without the aggressive correction parameter. Tables I–III summarize our findings. Table I shows the mean and variance of the prediction error for constant mean rates, namely for 200, 400, 600, and 800 kpackets/s. In general, the prediction error was found to be ~1.5% that translates to +/−15-KB error per MB transmitted. It can be seen that the LMS-based LPF performs very good for constant arrival rates and even better for large arrival rates (i.e. 800 kpacket/s), primarily because of the higher number of samples that lends the prediction algorithm a higher accuracy.

TABLE I
Performance of LMS-Based LPF for Constant Packet Arrival Rates ($\mu = 0.1, N = 8$)

λ (kpackets/sec)	AVE	VAR
100	2.01	5.21
200	1.72	4.32
400	1.48	3.22
600	1.32	2.13
800	1.21	1.49

TABLE II
Performance of LMS-Based LPF for Varying Packet Arrival Rates ($\mu = 0.1, N = 8$)

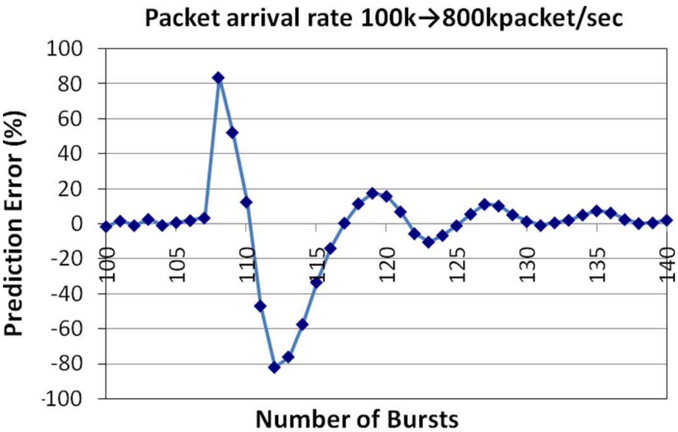
λ (kpackets/sec)	Time (sec)	Bursts	AVE	VAR
100→200	1.16	29	12.9	408.5
100→400	1.24	31	17.2	776.7
100→600	1.28	32	17.5	860.8
100→800	1.42	35	17.7	924.23

TABLE III
Performance of LMS-Based LPF With the Addition of the Aggressive Correction Parameter for Varying Packet Arrival Rates ($\mu = 0.1, N = 8, n = 2$)

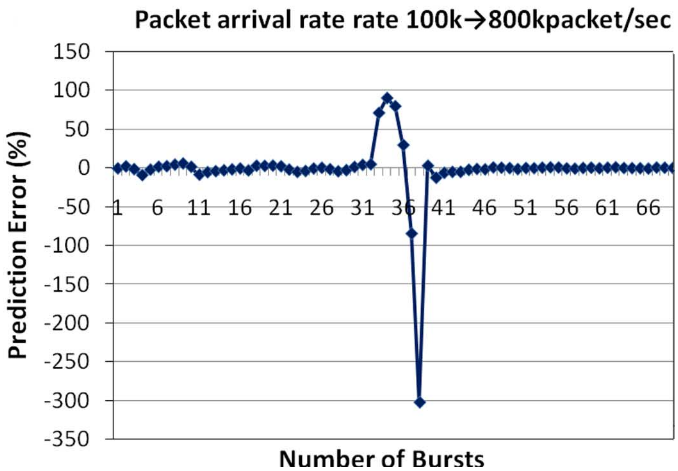
λ (kpackets/sec)	Time (sec)	Bursts	AVE	VAR
100→200	0.36	9	-25	4489
100→400	0.36	9	-49	11949
100→600	0.48	12	-54	14740
100→800	0.40	10	-23	13074

Table II shows the performance of the filter against instant increases in the packet arrival rate (see first column of Table II) without the addition of the correction parameter.

In particular, Table II provides the elapsed time until the filter error, $e(k)$ reaches a steady state with a variance below 10, the number of the bursts transmitted within that period as well the $\tilde{e}(k)$ average and variance only for that period. It is clear that the LPF exhibits a delay in following the traffic increase and which delay increases with the magnitude of increase. For example, the filter mechanism needs 1.42 s to adapt to an increase from 100 kpackets to 800 packets/s, while 1.16 s for an increase from 100 kpackets to 200 packets/s. A major difference is denoted in the mean and variance of the error in those periods. In particular, in the latter case (100-to-200 kpackets/s) the mean error of the filter is only 12.9% with a variance of 408.3 while in the first case (100-to-800 kpackets/s) both values are by far larger. Fig. 4(a) shows the error variation per burst transmitted around the traffic change for an increase of 100 k to 800 kpackets/s in the arrival rate. This behavior is inherent with LMS-based algorithms since they constitute a good compromise of convergence speed and tracking performance. While applying LMS-based LPF for traffic prediction, on one hand, a large change in traffic reduces prediction delay, but brings the problem of convergence



(a)



(b)

Fig. 4. Error variation per burst transmitted for a rate change of 100 k to 800 kpackets/s (a) without and (b) with the aggressive correction parameter.

that leads to an increased prediction error, while on the other hand, a smaller change gives less prediction error but a longer prediction delay.

Table III displays the same performance metrics against the same instant increases in the packet arrival rate, but with the addition of the aggressive parameter. It can be seen, that convergence delay has been significantly reduced down to 0.36 s, and, thus, fewer bursts are transmitted with an error in their predicted length. Therefore, tracking performance has been improved at the tradeoff, however, of accuracy. This can be seen from Table III, where average and variance of the prediction error are increased. This is due to the fact, that the addition of the correction parameter induces jumps in burst length prediction in order to fast compensate errors of the previous predictions. This can be also seen from Fig. 4(b) that displays the error variation per burst transmitted around the traffic change. It can be seen that within a few bursts with a positive error in their predicted length, the filter mechanism performs a jump to negative values (overestimation), having identified the increase in the traffic. It must be noted here, that prediction errors that concern overestimations of the burst size, do not affect traffic but only waste a fraction of the reserved resources. In addition the mechanism does not perform uniformly for all traffic changes. For example for the specific case of 100-to-600 kpacket/s, the filter performs worse and this is due to the fact that the jump performed is not optimum for that specific increase and for that specific parameters $(f(\cdot), n)$ used in $d(k)$. However, this is an inherent tradeoff between greedy overestimations and convergence speed.

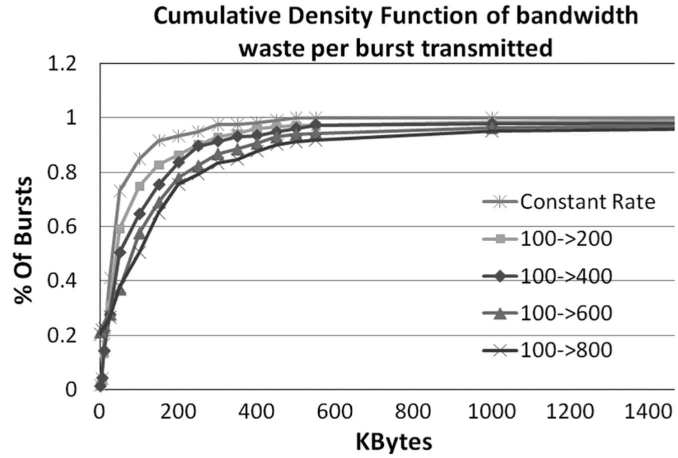


Fig. 5. Cumulative density function of bandwidth waste.

The use of the aggressive correction parameter for making the reservation may result in extra blocking in the core, primarily because of the higher work load offered in the network. To this end, the extra bandwidth reserved has been measured in order to determine the magnitude of the blocking increase. Fig. 5 shows the CDF function of the extra bandwidth reserved for constant (without the aggressive correction function) and nonconstant (with the aggressive function parameter) traffic. It can be seen, that the extra bandwidth reserved is less than 300 KB for the 80% of the bursts transmitted in all cases of traffic increases, while only a very small percentage ($<1\%$), exhibit a bandwidth waste higher than 800 KB. On average, it has been found out that in the case of 100–$>$200 kpacket/s, the average burst size is 2.94 MB, while the average extra bandwidth reserved is 212 KB. Therefore, we may argue that the average % increase of the actual burst size is $\sim 7\%$. This increase in the reservation horizon will slightly affect blocking performance and only in the first burst transmissions, during the traffic increase, some extra burst losses will occur. The yielding % increase of blocking can be derived from the work presented elsewhere [15], [16], having in mind that the proposed scheme emulates an one-way protocol.

V. Packet Loss-Free Operation in Large Networks

We have evaluated the performance of the proposed virtual one-way scheme on the National Science Foundation (NSF) network topology with emphasis in achieving packet loss-free operation. In such a case, packets that should be dropped, due to an underestimation in the burst length or a blocked setup message, are delayed for transmission during next assembly cycles. Especially in the case of a failure in establishing an end-to-end path, then within a very short time ($< T_{\mathrm{RTT}}$) a large amount of data are left over at the ingress node. This will result in increasing both the size of the transmitted bursts as well as the queuing delay of the assembled packets. Therefore, both the buffering and delay requirements should be investigated.

TABLE IV
ROUND TRIP TIME (ms) PER EDGE NODE FOR THE NSF NETWORK TOPOLOGY

Edge Node	0	1	2	3	4	5	6	7
Ave	30.4	36.9	22.1	23.4	23.4	28.6	29.4	32.6
Min	6.0	11.0	8.0	3.0	3.0	3.0	13.0	6.0
Max	49.0	59.0	41.0	46.0	46.0	59.0	51.0	46.0

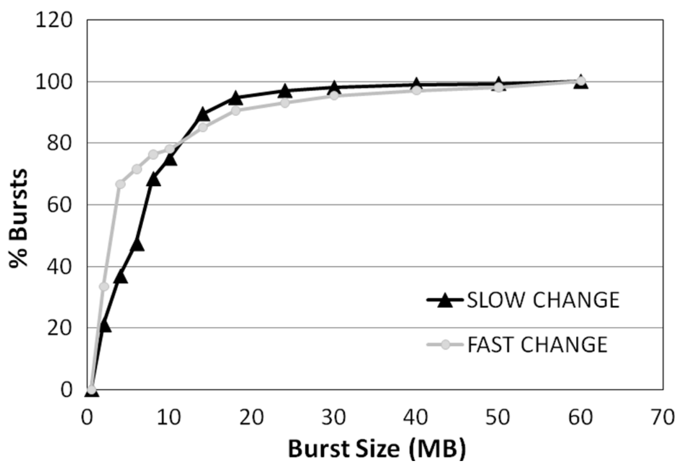


Fig. 6. Burst length density function for fast and slow rate changes.

We used ns-2 simulator and modeled the NSF network consisting of 8 edge nodes, 6 core nodes, one wavelength at 10 Gbps per link and a mean packet arrival rate of 100 kpackets/s, with packet size drawn from an Internet mix size distribution, [17]. The latter values were selected for attaining 1% blocking in the network. A two-way signaling protocol was employed with delayed burst-level reservations, [18], to match the operation of JET reservation scheme. Burst assembly time was set equal to the round-trip-time for each set of source-destination pair. Table IV presents the average, minimum and maximum values of these assembly times per edge node, when Dijkstra algorithm is used to calculate the routing.

We have evaluated the proposed aggressive resource reservation scheme considering two cases of traffic violation, namely fast with an increase of 400 kpacket/s per 200 ms and slow with an increase of 100 kpacket/s per 200 ms. Fig. 6 and Fig. 7 display the distributions of burst sizes and packet delay respectively. From Fig. 6, it is clear that the fast changes in the packet arrival rate increases the yielding burst size. This increase is evidence that the queuing time of the assembled packets increases, since a higher number of packets postpone their transmission for the next assembly cycle, thus forming larger bursts. The average sizes of the transmitted bursts were measured to be 7.8 and 7.3 MB, respectively (see Fig. 6), for the two cases, while only a small percentage of them exhibited a very large size of more than 20 MB. This implies that large buffers are not needed, and, thus, we may argue that the proposed virtual one-way reservation scheme fast compensates traffic violations deterring the accumulation of large data at the network edges.

Fig. 7(a) and (b) displays the corresponding distribution of packet delays with and without the aggressive burst length prediction mechanism, while for reference, we have also added Fig. 7(c) that shows the corresponding distribution, excluding the extra packets from the previous assembly cycles. Thus,

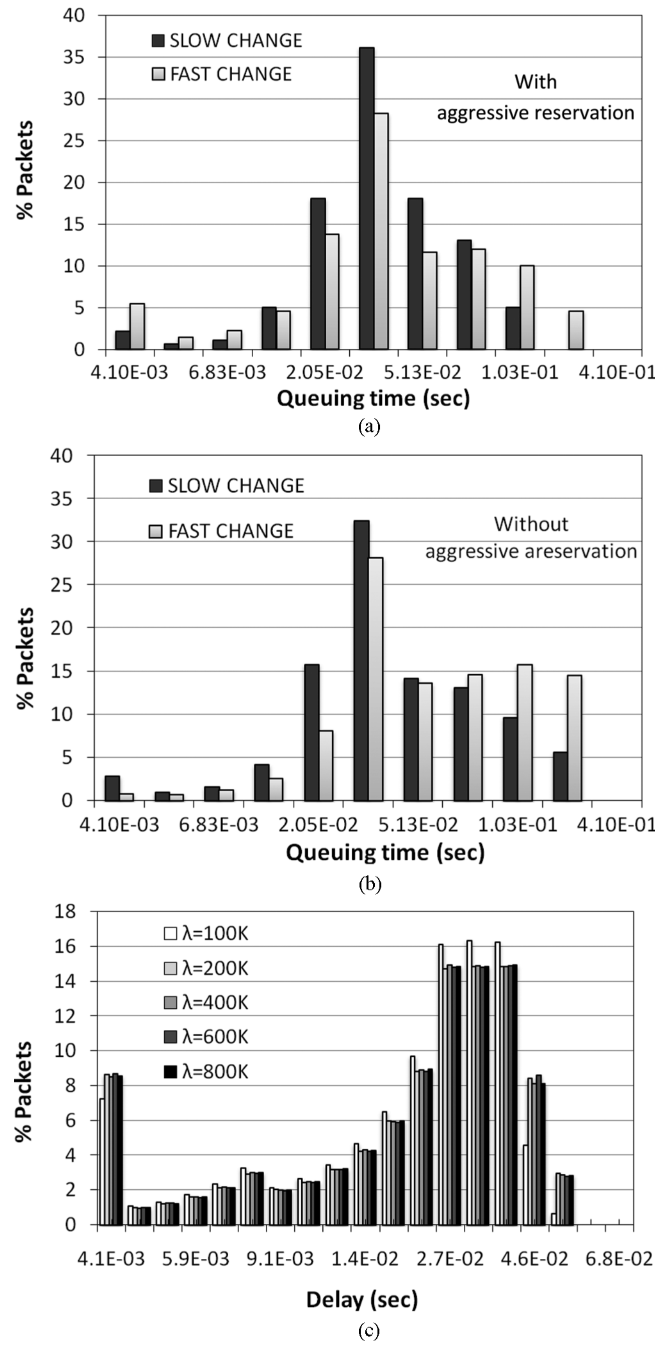


Fig. 7. Packet queuing time distribution (a) with and (b) without the addition of the correction parameter, $d(k)$. (c) Corresponding queuing delays for constant arrival rates.

Fig. 7(c) displays the initial packet delay distribution, that is toggled to that of Fig. 7(b), when employing only the LMS-based prediction filter and to that of Fig. 7(a), when adding the aggressive correction parameter $d(k)$. It must be noted here that the results of Fig. 7 concern a certain source-destination pair with $T_{\mathrm{RTT}} = 51.3$ ms. Thus, packets that exhibit a higher than T_{RTT} delay, are transmitted in the next assembly cycles. From Fig. 7, it is clear that when employing the aggressive mechanism, the percentage of packets with a delay higher than T_{RTT} is reduced. In particular, in the case of a slow change in the arrival rate, 29% of the packets exhibit a delay higher than

T_{RTT}, when employing only the LMS-based prediction filter [black columns of Fig. 7(b)], that is then reduced to 18%, when employing the aggressive parameter as well [black columns of Fig. 7(a)]. Similarly, in the case of fast changes in the packet arrival rate, 45% of the packets exhibited a delay higher than T_{RTT}, from which 14% a delay even higher than $2T_{\mathrm{RTT}}$ (these packets depart during the third assembly cycle). These are then reduced down to 26% and 4%, respectively, when adding the aggressive correction parameter.

To this end, we may argue that the proposed aggressive resource reservation mechanism is capable of compensating rapidly bursty increases in the traffic, assuring an average packet delay close to round trip time and a worst case delay twice the round trip time.

VI. Conclusion

In this paper, we have presented a novel virtual one-way reservation scheme for the on demand use of capacity in OBS networks. The scheme relies on a two-way reservation protocol and a burst assembly scheme that incorporates a burst length prediction mechanism. The burstification delay is enforced to be equal to the round-trip-time, so that the two-way reservation of resources to start immediately, upon the arrival of the first packet in the queue, for the estimated duration of the burst. We have evaluated the proposed scheme for constant as well as varying packet arrival rates and assessed its convergence speed and accuracy. We have further evaluated its performance on a large scale network with emphasis in achieving packet loss-free operation. It was shown, that the proposed virtual one-way protocol can guarantee zero data losses, with an average packet delay close to round trip time and a worst case delay twice the round trip time.

References

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

[2] Y. Xiong *et al.*, "Control architecture in optical burst-switched WDM twoemrule="yes" networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, p. 1838, Oct. 2000.

[3] J. Y. Wei and R. I. MacFarland, Jr., "Just-In-Time signaling for WDM optical burst switching twoemrule="yes" networks," *IEEE/OSA J. Light. Technol.*, vol. 18, no. 12, pp. 2019–2037, Dec. 2000.

[4] M. Yoo *et al.*, "QoS performance of optical burst switching in IP over WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 10, p. 2062, Oct. 2000.

[5] V. Vokkarane and J. P. Jue, "Prioritized burst segmentation and composite burst-assembly techniques for QoS support in optical burst-switched networks," *IEEE J. Select. Areas Commnun.*, vol. 21, no. 9, p. 1198, Sep. 2003.

[6] W. Liao and C. Loi, "Providing service differentiation for optical-burst-switched networks," *IEEE J. Light. Technol.*, vol. 22, no. 7, Jul. 2004.

[7] Q. Zhang *et al.*, "Absolute QoS diff. in Optical burst-switched networks," *IEEE Sel. Areas Commun.*, vol. 22, no. 11, p. 1781, Nov. 2004.

[8] M. Dueser and P. Bayvel, "Analysis of a dynamically wavelength-routed optical burst switched network architecture," *IEEE J. Light. Technol.*, vol. 20, no. 4, p. 574, Apr. 2002.

[9] D. Morato, J. Aracil, L. A. Diez, M. Izal, and E. Magana, "On linear prediction of Internet traffic for packet and burst switching networks," in *Proc. ICCCN*, 2001, pp. 138–143.

[10] J. Liu *et al.*, "FRR for latency reduction and QoS provisioning in OBS networks," *IEEE Sel. Areas Commun.*, vol. 21, no. 9, p. 1210, Sep. 2003.

[11] J. R. Treichler *et al., Theory and Design of Adaptive Filters*. New York: Wiley, 1987.

[12] Y. Xinyu, Z. Ming, Z. Rui, and S. Yi, "A novel LMS method for real-time network traffic prediction," *Springer Lecture Notes Comput. Sci.*, vol. 3046, pp. 127–136, 2004.

[13] P. B. Chu, S.-S. Lee, and S. Park, "MEMS: The path to large optical crossconnects," *IEEE Commun. Mag.*, vol. 40, no. 3, pp. 80–87, Mar. 2002.

[14] 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. IEEE GLOBECOM*, San Francisco, CA, Dec. 1–5, 2003, vol. 5, pp. 2777–2781.

[15] K. Dolzer *et al.*, "Evaluation of reservation mechanisms for optical burst switching," *Int. J. Electron. Commun.*, vol. 55, no. 1, 2001.

[16] J. Aracil and M. Izal, "On the influence of self-similarity on optical burst switching traffic," in *Proc. GLOBECOM*, vol. 3, pp. 2308–2312.

[17] Agilent Technologies, Aug. 2001 [Online]. Available: http://advanced.comms.agilent.com/insight/2001-08/Questions/traffic_gen.htm

[18] K. Vlachos *et al.*, "Relaxing delayed reservations: An approach for quality of service differentiation in optical burst switching networks," in *Proc. Broadnets*, Oct. 2006.

Kyriakos G. Vlachos (M'99) received the Dipl.-Ing. and Ph.D. degrees in electrical and computer engineering from the National University of Athens (NTUA), Greece, in 1998 and 2001, respectively.

From 1997 to 2001, he was a senior research associate in the Photonics Communications Research Laboratory (ICCS/NTUA). In April 2001, he joined Bell Laboratories, Lucent Technologies, working on behalf of the Applied Photonics Group. He conducted research on high-speed optical networks and DWDM transmission techniques. During 2003, he joined the National Regulation Authority of Telecommunication and Postal Service of Greece (EETT), where he served as a scientific advisor for various techno-economic issues for the promotion of broadband technologies. Since 2003, he has also been a member of the Computer Engineering Laboratory, Technical University of Delft, and since 2005, a faculty member of Computer Engineering and Informatics Department of the University of Patras, Greece. In 2007, he founded the Photonic Networks and Technology Laboratory (PNET) (http://ru1photonicslab.cti.gr/). His research interests are in the areas of high-speed protocols and technologies for broadband, high-speed networks, optical packet/burst switching, and grid networks. He has participated in various research project funded by the European Commission (IST-STOLAS, IST-PRO3, ESPRIT-DOALL, ephoton/ONe+, IST-PHOSPHROUS, ICT-BONE και ICT-DICONET).

Prof. Vlachos is a member of the Technical Chamber of Greece and he periodically acts as a scientific reviewer for the General Secretariat for Research and Technology of Greece (GSRT), as well as for the European Commission and the Netherlander Organization for Scientific Research, Technology Foundation. He is the (co)author of more than 90 journal and conference publications and holds five patents.

Demetris Monoyios received the diploma of Computer Engineering from the Computer Engineering and Informatics Department (CEID), University of Patras, Greece, in 2007. He is currently pursuing the Ph.D. degree.

His research interests include optical networks, optical burst switching, algorithms, and protocols. He is the (co)author of three publications.