

MAC-Layer Performance Enhancement Using Control Packet Buffering in Optical Burst-Switched Networks

Ahmed I. Abd El-Rahman, *Student Member, IEEE*, Sherif I. Rabia, and Hossam M. H. Shalaby, *Senior Member, IEEE*

Abstract—One of the major hurdles in optical burst-switched (OBS) networks is the contention problem. In this paper, we tackle this problem by proposing a new contention resolution technique based on control packet buffering. In particular, a rigorous mathematical model is developed to analyze the performance of an OBS network core node employing this technique along with just-in-time one-way reservation protocol. The key idea that makes this proposal novel is that this buffering is implemented in the electronic domain, thus avoiding the complexity that exist in the optical domain solutions. A detailed evaluation of the system performance using both the burst loss probability and the steady-state throughput is presented. Our results reveal that a considerable improvement in the performance is achieved when adopting the proposed technique.

Index Terms—Just-in-time (JIT), optical burst-switched (OBS) networks, queuing theory.

I. INTRODUCTION

THE evolution of communication technology is strongly correlated with the demand for higher bit rates. This justifies the endeavors exerted by researchers to optimize the utilization of the vast bandwidth available on the optical fiber. However, the unavoidable need for electronic devices in optical networks to perform switching operations, which implies the conversion of the signal from the optical domain to the electronic domain and back to the optical domain (OEO conversion), imposes restrictions on the maximum achievable bit rate. In principle, the OEO conversion limits the overall transmission speed of the optical fiber system. Thus, many research work addressed this problem and many suggestions aimed to overcome the OEO hurdle and build an all optical network (AON). On the way to an AON, and especially due to lack of advanced optical devices that can effectively replace their peer electronic devices, optical burst switching has gained a great potential as it represents

a good compromise between optical circuit-switching and optical packet-switching [1]. In particular, optical burst-switching (OBS) aims at maintaining the fiber high bandwidth along the burst path by creating an all optical path that bypasses the OEO conversion. The salient feature in OBS networks is that the packets heading the same destination and belonging to the same class are aggregated by means of appropriate aggregation algorithms to form a *data burst* (DB) and the switching operation is then to be performed on a *per burst* basis [2]–[4]. This is constructed as follows: the burst header, referred to as *the control packet* (CP), which is carrying the information required for the routing operations, travels on a separate out-of-band control channel prior to the burst transmission from the ingress node to the egress node through the core nodes. Its mission is to reserve, on each core node along the way from source to destination, appropriate resources for its ensuing burst [5]. This allows the coming optical burst to pass smoothly through the whole way without being optically buffered nor being passed through the bottleneck of an OEO conversion. At the egress node the burst is disassembled back to packets. This scenario implies the following.

- Since OBS is designed to be employed mainly in long haul optical networking, one-way reservation protocols like “just-enough-time” (JET) [1] and “just-in-time” (JIT) [6], [7] are the most suitable to reach an ultra-low-latency burst transport. Indeed, the delay of two way reservation protocol would degrade the service drastically.
- The burst must wait at the ingress node for a predetermined time, called *offset time*, to account for the CP processing time. This way, the burst will arrive at the core node only when the switch fabric is configured to bypass it.
- In the core nodes, the control packets contend for available resources, i.e., wavelength (WDM) or code (OCDMA). Consequently, failing CPs and their ensuing bursts will be blocked, which, in turn, results in the loss a large number of packets, as one burst may extend from one packet to a whole session.

Many efforts have been exerted by researchers to present mathematical models which analyze the performance of OBS networks. In [8] Shalaby proposed a simplified mathematical model to study the performance of an OBS core node assuming Bernoulli distribution for arrivals per time slot, which proved to be a good assumption until a certain traffic load when compared to the simulation results that assumed Poisson distribution for arrivals. In [9] Morsy *et al.* proposed an enhanced mathematical model for the performance evaluation of OBS core nodes in order to relax some of the constraints given in [8]. In addition, researchers addressed the contention problem in many

Manuscript received June 12, 2011; revised November 16, 2011 and December 22, 2011; accepted January 19, 2012. Date of publication February 06, 2012; date of current version April 04, 2012.

A. I. Abd El-Rahman is with the Department of Electrical Engineering, Alexandria University, Alexandria 21544, Egypt (e-mail: ahmedibrahim766@hotmail.com).

S. I. Rabia is with the Department of Engineering Mathematics and Physics, Alexandria University, Alexandria 21544, Egypt (e-mail: sherif.rabia@alexu.edu.eg).

H. M. H. Shalaby is currently with the Department of Electronics and Communication Engineering, Egypt-Japan University of Science and Technology (E-JUST), New Borg EL-Arab City, Alexandria 21934, Egypt, on leave from the Electrical Engineering Department, Alexandria University, Alexandria 21544, Egypt (e-mail: shalaby@ieee.org).

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.2012.2187043

occasions. In [10], Akar *et al.* elaborated on a WDM system and suggested using wavelength conversion for contention resolution. In [11], Sowailam *et al.* proposed a new system that employs the code domain instead of the wavelength domain. In fact, they adopted spectral amplitude coding optical CDMA (SAC-OCDMA) techniques. They have shown that this SAC-OCDMA system outperforms the traditional WDM system, as it can handle more users, but it suffers from complexity. In both systems, the improvement in the system performance is correlated with the number of converters.

The aim of this paper is to add a new feature, namely control packet buffering, to the MAC layer of the OBS network as a new contention resolution technique. This feature does not depend on the medium access technique and might be regarded as a new modification to the JIT one-way reservation protocol. Therefore, it can be easily implemented either above SAC-OCDMA or WDM based optical layer. The key idea of this feature is that the CP that fails in reserving its required resource will not be dropped immediately, rather electronically buffered for some threshold time X which is determined at the ingress node according to each burst duration. Meanwhile, the required resource may be released and consequently immediately reserved for the new burst. Otherwise, the CP will be dropped, and the ensuing DB will be lost. This way, the probability of being dropped, namely *the per node burst loss probability*, is decreased. Hence, this feature can be employed to replace resource converters or to enhance the system performance in presence of partial conversion, which would lead to a great reduction in the system complexity. This suggestion requires some modifications in the burst offset time, in order to avoid the burst arrival while the core nodes are still not ready to bypass it.

In addition to our proposal of CP buffering, two more novel aspects in this paper comprise the following: we develop a new mathematical model, which makes use of queuing theory impatience concept, in order to measure the performance of our proposed technique. Moreover, we present a rigorous mathematical notation to make it easy to generalize the model.

This paper is organized as follows. The system description is presented in Section II. Section III is devoted to both the mathematical model and the performance analysis. In Section IV, we present some numerical results for the derived performance measures. Finally, our conclusions are given in Section V and some future work is presented in Section VI.

II. SYSTEM DESCRIPTION

A. JIT One-Way Reservation Protocol

The JIT one-way reservation protocol is one of the main protocols suggested to be used in optical burst switched networks. As explained in [7] and [9], the protocol is in general based on two main features.

- *Immediate channel reservation*: After CP processing, the core node immediately reserves the required resource, if available, and a channel busy period is declared although the burst has not arrived yet.
- *Explicit channel release*: The resource is maintained busy till the core node receives an explicit release signal. This takes some time after the burst switching process.

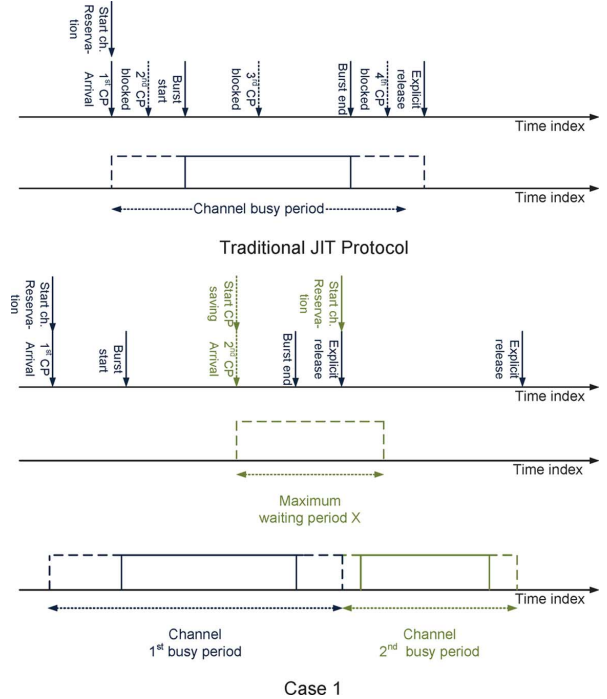


Fig. 1. Timing diagram explaining the traditional JIT operation and the modified operation in presence of buffering (Case 1).

This process is illustrated in the top of Fig. 1. Observing the figure, one can easily notice that the main drawback of this scheme is that the resource busy period is longer than the actual burst duration, which represents a waste of resource and increases the burst loss probability.

In fact, JET is more efficient than JIT, but its strategy is more sophisticated. For example, the control packet in the case of JET must bear information about the burst arrival time and the burst length so that the core node can estimate the start and the end of the required period and perform a delayed reservation with implicit channel release. Hence, the reservation algorithm is complex and adding a new feature like CP buffering would increase this complexity further. On the other hand, the main target of this work is to propose a simple solution to the contention problem that reduces the need for resource converters, which represent a main source of complexity. Furthermore, we present this novel feature as a prelude to further investigations. So, in order to get more insight into the problem under study, we decided to present it in the simplest form possible. Consequently, the choice of the JIT, thanks to its simplicity, seemed more suitable.

One of the main advantages of the layering concept in data networks is that the protocols of each layer can be designed independently. This way, a wide range of interoperability can be achieved. Following the same concept, we propose the CP buffering as a simple MAC layer performance enhancing technique. As stated earlier, it is a slight modification to the MAC layer JIT and does not require any modification in the upper layer protocols. Regardless of the upper and lower layer protocols, (Fig. 3) illustrates the network modified MAC layer.

B. Proposed Control Packet Buffering With JIT

In [5] and [9] a detailed description of the network ingress, egress and core nodes is presented. In this paper, we are simply

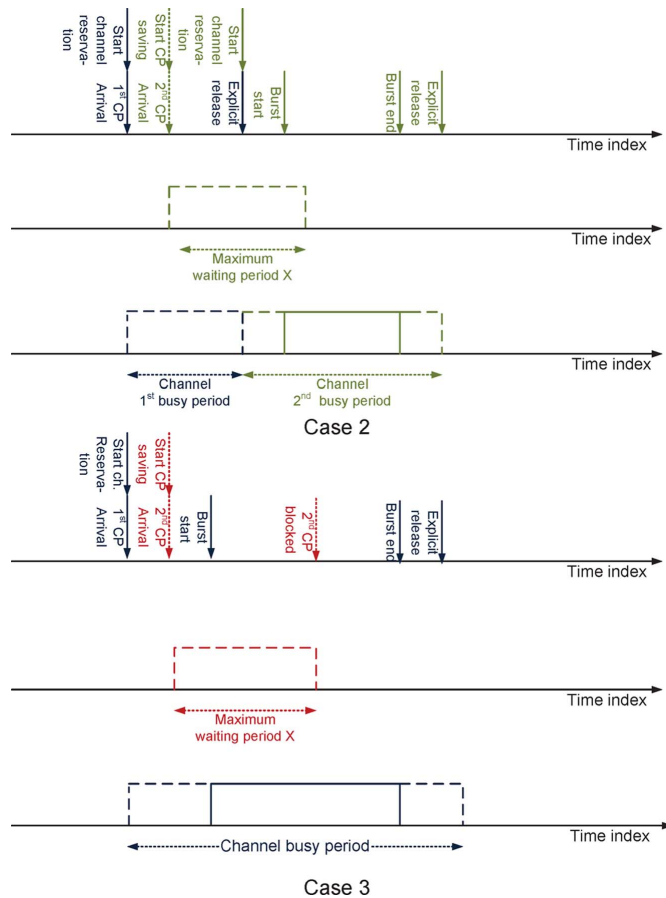


Fig. 2. Timing diagram explaining the JIT modified operation in presence of buffering (Case 2) and the blocking situation (Case 3).

interested in explaining the CP buffering feature. Thus, we adopt the case of no resource conversion. According to our proposal, new functions must be added to the ingress and core nodes as follows:

In addition to its main job, the ingress node assigns to each CP prior to its transmission a threshold time that is directly proportional to its burst length. Furthermore, it increases the offset time of the burst at the ingress node by $c \times$ the assigned waiting time, where c is the expected number of congested hops on the expected way of each burst. This can be easily calculated at the ingress node based on the congestion statistics (this process is incorporated in the offset time generator Fig. 3). Here, it should be noticed that the increment in the offset time is not constant for all bursts, as the assigned waiting time and the parameter c differ from one burst to another. This variable offset time is necessary to help resolving the contention problem.

On the other hand, the bank of resource converters in each core node is removed and a small size buffer at each core node is electronically implemented. As mentioned earlier, this buffer saves a control packet that finds the appropriate resource busy upon arrival. The saving duration is limited to the threshold time predetermined in the ingress node. During this duration the contended resource might be released, and it would be therefore possible for this CP to reserve it immediately for its burst. Otherwise, i.e., the time is elapsed and the resource is still grasped, the CP will be dropped and the coming burst will be ignored. This process comprises three cases that are portrayed in both Figs. 1

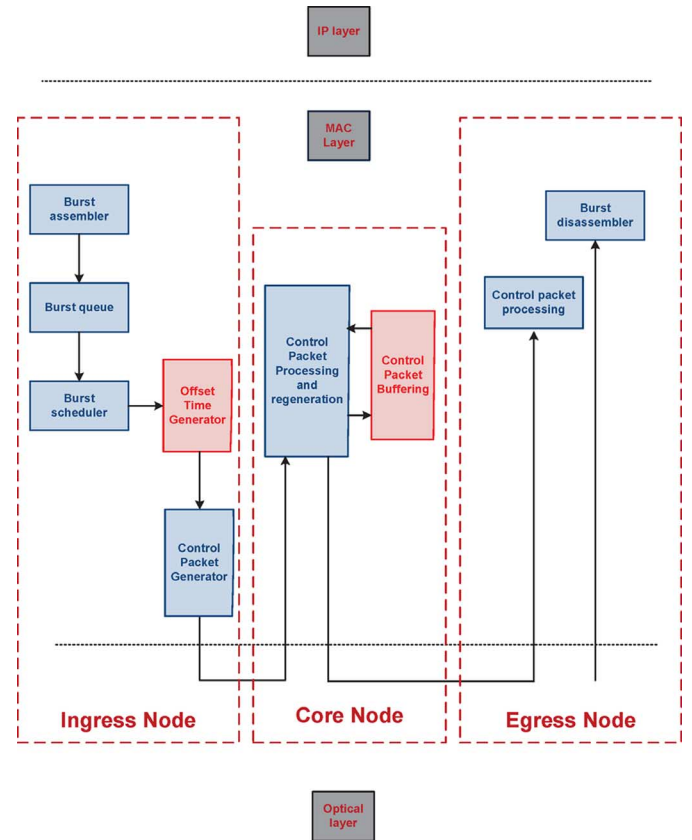


Fig. 3. OBS MAC layer with the proposed CP buffering feature.

and 2. In the first two cases the channel is released while the CP is still saved. Consequently, the CP is extracted and the channel is immediately reserved for its ensuing burst. The third case represents the CP case of failure, in which the waiting period elapses while the channel is still busy (The whole CP buffering process takes place in conjunction with the CP processing and regeneration Fig. 3). The reason for portraying cases 1 and 2 is to show that the channel might be released in two different ways.

- 1) After the normal switching of a scheduled traffic.
- 2) When a scheduled burst does not arrive, which is the case when CP fails in one core node on the way to the destination, then a backward release signal must be sent to free resources already reserved in previous core nodes.

The required ultra-low-latency for burst transport necessitates the use of one-way reservation protocols. Consequently, the burst could be lost due to contention. In this case, the ingress node would detect the burst loss event when it receives a negative acknowledgment from core nodes where the collision had occurred. Then, the ingress node would start a burst retransmission process [12], [13]. In fact, the burst retransmission might be viable, as no contention resolution technique can achieve zero burst loss probability. However, retransmission schemes would increase the traffic load resulting in higher burst loss probability. In addition, owing to the substantial average burst retransmission delay, reducing the number of bursts to be retransmitted by undermining the contention chances remains a main design issue. Furthermore, it should be pointed out that, for long haul communications, i.e., for large dimension networks, the main defect in the retransmission process is attributed to the high

propagation delay experienced by the packet bearing the negative acknowledgment (NACK). That is, the average burst retransmission delay is proportional to the network dimension. Clearly, the proposed CP saving process will delay the burst delivery. Nevertheless, we will show later that this delay is much less than that introduced by the burst retransmission. Moreover, our proposal will reduce the burst blocking probability, thus the MAC layer delivery success will be more likely and the number of bursts to be retransmitted will be lowered. Hence, this delay can be justified.

The purpose behind having limited buffering time is that uncontrolled waiting time might cause intolerable delays. In addition, it might be longer than expected and the DB might arrive before reserving the appropriate resources. This way the buffering will not only be useless but it will also cause a waste of other resources already reserved in precedent nodes. Furthermore, the proportionality between the threshold time and the burst length implies that the burst loss probability will follow the burst length. In other words, it will be less likely to block bursts comprising larger number of packets.

Finally, we may summarize the exact difference of CP buffering feature compared to standard offset-time-based QoS provisioning: The offset time based QoS provisioning is essentially concerned with classifying bursts according to their priority and assigning different extra offset times to different classes so that higher priority classes have privilege over other classes mainly in the burst loss probability. The purpose behind this technique is to achieve higher reliability for mission critical and real time applications by providing lower blocking probability, lower time jitter, etc. On the other hand, our proposal focuses on fairly improving the system burst loss probability by allowing the blocked CP to be saved in the core node buffer for a predetermined time, as meanwhile the contended resource might be released. Moreover, since longer bursts carry larger amount of information, judicious waiting time (patience) assignment implies making the waiting time (patience) proportional to the burst length. Consequently, the CP buffering feature, as suggested in our paper, does not isolate traffic classes. However, the flexibility of the proposed feature and mathematical model make it possible to investigate the introduction of QoS issues with JIT protocol.

III. PERFORMANCE ANALYSIS OF THE SYSTEM MAC LAYER

As discussed in [11], a successful burst that travels from ingress node to egress node implies in principle both a CP success in the MAC layer and a burst transmission success in the optical layer. Thus, in order to find the overall burst loss probability, a top down approach should be adopted. That is, the performance analysis should be carried out first in MAC layer then in the optical layer.

To model the system performance, a continuous time Markovian analysis is conducted to illustrate the performance of a single core node. Then, a recursive formula of the *per node blocking probability*, i.e., *per node burst loss probability*, will be found. First, recall the model's main assumptions.

- Poisson arrivals with average arrival rate λ .
- Exponential service time with average service rate $\mu = (1/\tau)$, where τ is the average burst duration.

- The offered load is thus given by

$$\rho = \lambda/\mu$$

but since each port can carry up to N different users, as there are N distinct resources and assuming that the load is uniformly distributed over N , then the effective offered load is found as

$$\rho_{\text{eff}} = \frac{\lambda}{N\mu}.$$

As stated earlier, the CP buffering feature can be easily implemented either above SAC-OCDMA or WDM based optical layer. Thus, in order to confirm the system flexibility, we assign the general symbol N and define it as the number of available distinct resources per link, which would refer to either the number of available wavelengths or codes per link.

When neither a buffer nor resource conversion are applied, each resource can be safely modeled as an $M/M/1/1$ loss system. From the Erlang "b" table, we get

$$P_{\text{Node-Loss}} = \frac{\rho_{\text{eff}}}{1 + \rho_{\text{eff}}}. \quad (1)$$

While if we introduce a buffer to each resource with m storage capacity and unlimited queue waiting time, each resource, assuming no resource conversion, can be modeled as an $M/M/1/m + 1$ loss system. Hence, the long run $m + 1$ state probability, which represents the loss probability in this case, is given by [14]

$$\pi_{m+1} = \frac{1 - \rho_{\text{eff}}}{1 - \rho_{\text{eff}}^{m+2}} \times \rho_{\text{eff}}^{m+1}. \quad (2)$$

In fact, the last equation does not apply to the case suggested in this paper because the queue waiting time is restricted to some value assigned at the ingress node, which is necessary to avoid waste of resources. Thus the analysis should be carried out in a different way, in which the concept of "impatience" is introduced.

A. Impatience Model "Reneging"

From the perspective of an observer located at the core node, the time that an arbitrary CP would wait in the buffer is random. Although the approach we will develop can be used for generalized models, for simplicity we have assumed an exponential service time with parameter μ . For the same reason, let us assume that each CP is allowed to wait in each core node buffer for an exponentially distributed random time with parameter ν , which is related to the average burst length, i.e., $\nu = c\mu$. In queuing theory such restriction is known as a sort of "queues with impatience" labeled "reneging" [15], in which an arrival joins the queue for some time and may abandon the queue if it runs out of patience before reaching the server. Before expressing this argument mathematically, we define the following random variables, events and probability density functions PDFs.

- " N_i ": N_i is a random variable, that represents the number of CPs that will be eventually served out of i queued CPs.
- " X ", " $f_X(x)$ ": Let X denote the patience of a tagged control packet CP_{tag} and let $f_X(x)$ be its exponential PDF with parameter ν .

- “ T ”, “ $f_{T|N_i}(t|n)$ ”: T represents the time, that CP_{tag} should spend to penetrate through the queue and reach the server, if its patience is infinite. In literature, T is sometimes called *the potential waiting time* [16]. Thus, assuming that CP_{tag} exists currently in the i th position of the queue, then $f_{T|N_{i-1}}(t|n)$ will be the conditional PDF of CP_{tag} potential waiting time given that $N_{i-1} = n$. Based on the memoryless property of the exponentially distributed service [14], the reader would expect that $f_{T|N_{i-1}}(t|n)$ is Erlangian with parameters μ and $n + 1$.
- “ R_i ”: Let R_i^1, R_i^0 refer to the *eventual* reneging event of CP_{tag} currently located in position i and its complement, respectively [Fig. 4(b)]. For simplicity, the superscript will be omitted from R_i^1 .

Consequently, R_i will occur if $T > X$. In the general case, i.e., when CP_{tag} occupies an arbitrary position i , its potential waiting time T depends on N_{i-1} and, recalling that $f_X(x) = \nu e^{-\nu x}$, the conditional probability $P(R_i|N_{i-1} = n)$ can be obtained according to the total probability theorem

$$P(R_i|N_{i-1} = n) = \int_0^\infty \int_x^\infty f_{T|N_{i-1}}(t|n) dt \times f_X(x) dx, \quad 0 \leq n \leq i - 1; \quad 2 \leq i \leq m. \quad (3)$$

Fortunately, the process governing our restricted waiting time queue is still the well known birth-death process. However, the departure rate must be modified, since the arrival that joins the queue is now allowed to leave the system in two mutually exclusive ways, either served with rate μ or gave up before reaching the server with rate ν . Hence, the system notation is changed to $M/M/1/1+m$ to outline that the system has “1” server and can queue up to “ m ” CP’s with exponentially distributed *patience*.

In addition, based on Fig. 4(a), the system state probability can be easily derived to get

$$\pi_k = \pi_0 \prod_{j=1}^k \left(\frac{\lambda/N}{\mu + (j-1)\nu} \right) \quad (4)$$

where π_0 is given by

$$\pi_0 = \left[1 + \sum_{k=1}^{m+1} \prod_{j=1}^k \left(\frac{\lambda/N}{\mu + (j-1)\nu} \right) \right]^{-1}. \quad (5)$$

Our next target is to calculate the per node burst loss probability. First, let us explicitly define the two cases in which a burst will be lost.

- 1) When a CP finds the system full upon arrival, i.e., its required resource is reserved and the buffer assigned to this resource is full. Thus assuming that the buffer size is m , and based on the PASTA property of the Poisson process [15], this probability is the same as π_{m+1} .
- 2) When a CP joins the queue, but *reneges*. As defined earlier, provided that this CP joined the i th position of the queue, this is the event R_i . In order to find $P(R_i)$, it is assumed that the CPs are served in a *first in first out* (FIFO) manner and then the movement of this CP is tracked from its *initial* position to its *departure* position.

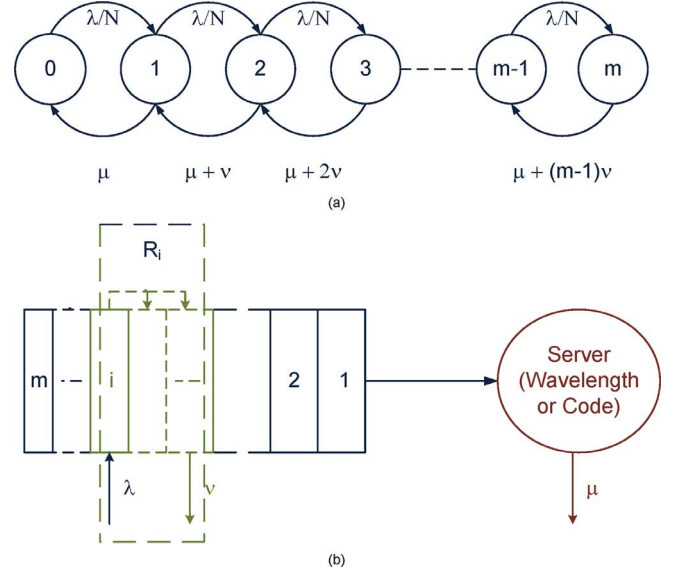


Fig. 4. (a) State transition rate diagram for one-server loss system with an m size buffer with reneging. (b) Reneging event, R_i .

Briefly a general form of $P_{Node-Loss}$ can be expressed as

$$P_{Node-Loss} = \pi_{m+1} + \sum_{i=1}^m \pi_i \times P(R_i) \quad (6)$$

where $P(R_i)$ comprises two cases.

- 1) $P(R_1)$, which can be simply obtained as follows. Since in this situation $f_T(t) = \mu e^{-\mu t}$, then

$$P(R_1) = \int_0^\infty \int_x^\infty \mu e^{-\mu t} dt \times \nu e^{-\nu x} dx = \frac{\nu}{\mu + \nu}. \quad (7)$$

- 2) $P(R_i)$, for $i > 1$, is obtained by conditioning on N_{i-1} and employing the total probability theorem

$$P(R_i) = \sum_{n=0}^{i-1} P(R_i|N_{i-1} = n)P(N_{i-1} = n). \quad (8)$$

Observing (8) we deduce that $P(R_i|N_{i-1} = n)$ is given by (3). Consequently it remains to extract a formula for $P(N_{i-1} = n)$, which is a nontrivial task. Let us start by adopting a step by step approach that will enable us to accomplish this task and develop a recursive formula for $P(R_i)$.

B. Burst Loss Probability for Simple $M/M/1/1+1$ and $M/M/1/1+2$ Models

Consider the first step, in which the system has only one buffer $M/M/1/1+1$. Applying (6), we get

$$P_{Node-Loss} = \pi_2 + \pi_1 \times P(R_1). \quad (9)$$

Hence, using (4), the blocking probability can be directly written as follows:

$$P_{Node-Loss} = \rho_{eff} \left(\frac{\lambda/N}{\mu + \nu} \right) \pi_0 + \rho_{eff} \frac{\nu}{\mu + \nu} \pi_0. \quad (10)$$

The second step is to consider the $M/M/1/1 + 2$ queue, in which the system can save up to two CPs. Applying (6) and (8), the $P_{\text{Node-Loss}}$ can be obtained as follows:

$$\begin{aligned} P_{\text{Node-Loss}} &= \pi_3 + \pi_2 \times [P(R_2|N_1 = 0)P(N_1 = 0) \\ &\quad + P(R_2|N_1 = 1)P(N_1 = 1)] \\ &\quad + \pi_1 \times P(R_1). \end{aligned} \quad (11)$$

In the last equation we have $P(R_2|N_1 = 0) = P(R_1)$ and $P(N_1 = 1) = P(R_1^0)$, while $P(R_2|N_1 = 1)$ can be calculated using (3). Thus, we may write $P_{\text{Node-Loss}}$ as

$$\begin{aligned} P_{\text{Node-Loss}} &= \rho_{\text{eff}} \frac{(\lambda/N)^2}{(\mu + \nu)(\mu + 2\nu)} \pi_0 \\ &\quad + \rho_{\text{eff}} \left(\frac{\lambda/N}{\nu + \mu} \right) \pi_0 \left[\frac{\mu\nu}{(\mu + \nu)^2} + \frac{\mu^2\nu}{(\mu + \nu)^3} \right] \\ &\quad + \rho_{\text{eff}} \left(\frac{\mu}{\mu + \nu} \right) \pi_0. \end{aligned} \quad (12)$$

C. Burst Loss Probability for the General Model $M/M/1/1 + m$

Now, it is required to find $P(N_{i-1} = n)$ in order to develop a formula that applies to the general system $M/M/1/1 + m$, for $m > 0$. Indeed, to avoid notation ambiguity and to simplify the general model formulation process, we consider the following notation.

- Define the binary sequence $K_i = (k_i(1), k_i(2), k_i(3), \dots, k_i(i-1))$, $2 \leq i \leq m$.
- Let $\Psi_i(n)$ be the set of all possible sequences K_i counting n zeros and $i - n - 1$ ones. That is, the cardinality of $\Psi_i(n) = \binom{i-1}{n}$, $0 \leq n \leq i-1$.
- Define the function $\phi_j(K_i)$ which gives the number of zeros in the subsequence $(k_i(1), k_i(2), \dots, k_i(j))$, for $1 \leq j \leq i-1$.

First it is customary to emphasize the following important remarks.

- *Remark 1:* Obviously, R_2 depends on R_1 . Hence as stated earlier in case of larger buffers, the event R_i depends on the events R_j and R_j^0 for all $j < i$.
- *Remark 2:* Note that $P(R_i|N_{i-1} = n)$ does not depend on the positions of the reneging items, but only on their numbers. On the other hand, in order to find $P(N_{i-1} = n)$, all combinations of mutually exclusive reneging events of n CPs in different positions must be taken into consideration.

The second remark gives rise to an important question: What is the joint probability that an arbitrary number n of queued CPs will not give up regardless of their position? The answer can be obtained by employing the notation explained earlier as follows:

$$\begin{aligned} P(N_{i-1} = n) &= \sum_{K_i \in \Psi_i(n)} P \left(R_1^{k_i(1)}, R_2^{k_i(2)}, R_3^{k_i(3)}, \dots, R_{i-1}^{k_i(i-1)} \right). \end{aligned} \quad (13)$$

Using the multiplication rule, we get

$$\begin{aligned} P(N_{i-1} = n) &= \sum_{K_i \in \Psi_i(n)} P \left(R_1^{k_i(1)} \right) \prod_{j=2}^{i-1} \\ &\quad P \left(R_j^{k_i(j)} | R_1^{k_i(1)}, \dots, R_{j-1}^{k_i(j-1)} \right) \\ &= \sum_{K_i \in \Psi_i(n)} P \left(R_1^{k_i(1)} \right) \prod_{j=2}^{i-1} \\ &\quad P \left(R_j^{k_i(j)} | N_{j-1} = \phi_{j-1}(K_i) \right). \end{aligned} \quad (14)$$

We can now take advantage of (3) to find $P(N_{i-1} = n)$.

Inspired by the notation explained above, we have reached a simple algorithm that employs (3), (6), (8), and (14) to obtain $P_{\text{Node-Loss}}$ for a general system with a queue of arbitrary length. To illustrate this idea, let us apply this result to extend the $M/M/1/1 + 2$ queue to an $M/M/1/1 + 3$ queue. Applying (6) the burst loss probability can be directly written as follows:

$$P_{\text{Node-Loss}} = \pi_4 + \pi_3 \times P(R_3) + \pi_2 \times P(R_2) + \pi_1 \times P(R_1). \quad (15)$$

Clearly, the third and fourth terms of (15) are known and it remains to calculate $P(R_3)$. Thus, using (8)

$$P(R_3) = \sum_{n=0}^2 P(R_3|N_2 = n)P(N_2 = n). \quad (16)$$

Next, applying the suggested notation to obtain $P(N_2 = n)$ for $n = 0, 1, 2$, we get the following cases.

- 1) $N_2 = 0$, $\Psi_3(0) = \{(1, 1)\}$, and using (13) and (14) we get

$$P(N_2 = 0) = P(R_1, R_2) = P(R_1) \times P(R_2|N_1 = 0).$$

- 2) $N_2 = 1$, $\Psi_3(1) = \{(0, 1), (1, 0)\}$, and using (13) and (14) we get

$$\begin{aligned} P(N_2 = 1) &= P(R_1^0, R_2) + P(R_1, R_2^0) \\ &= P(R_1^0) \times P(R_2|N_1 = 1) \\ &\quad + P(R_1) \times P(R_2^0|N_1 = 0). \end{aligned}$$

- 3) $N_2 = 2$, $\Psi_3(2) = \{(0, 0)\}$, and using (13) and (14) we get

$$P(N_2 = 2) = P(R_1^0, R_2^0) = P(R_1^0) \times P(R_2^0|N_1 = 1).$$

Now it is easy to obtain $P(R_3)$ as follows:

$$\begin{aligned} P(R_3) &= P(R_3|N_2 = 0) \times P(R_1) \\ &\quad \times P(R_2|N_1 = 0) + P(R_3|N_2 = 1) \\ &\quad \times [P(R_1^0) \times P(R_2|N_1 = 1) \\ &\quad + P(R_1) \times P(R_2^0|N_1 = 0)] \\ &\quad + P(R_3|N_2 = 2) \times P(R_1^0) \times P(R_2^0|N_1 = 1). \end{aligned} \quad (17)$$

In fact, through the analysis we have found all the terms required to write down $P(R_3)$ except $P(R_3|N_2 = 2)$, which can

be directly calculated from (3). Finally, we are ready to write $P_{\text{Node-Loss}}$

$$\begin{aligned}
 P_{\text{Node-Loss}} &= \pi_0 \rho_{\text{eff}} \frac{(\lambda/N)^3}{(\mu + \nu)(\mu + 2\nu)(\mu + 3\nu)} \\
 &+ \rho_{\text{eff}} \frac{(\lambda/N)^2}{(\mu + \nu)(\mu + 2\nu)} \pi_0 \times P(R_3) \\
 &+ \rho_{\text{eff}} \frac{\lambda/N}{(\nu + \mu)} \pi_0 \left[\frac{\mu\nu}{(\mu + \nu)^2} + \frac{\mu^2\nu}{(\mu + \nu)^3} \right] \\
 &+ \rho_{\text{eff}} \left(\frac{\mu}{\mu + \nu} \right) \pi_0. \quad (18)
 \end{aligned}$$

The developed approach can be employed in more general loss systems $M/G/1/1+m$ but with modifications in the state probability equations and in (3). The reason for these modifications is that in the absence of the memoryless property of the service time, which is the main characteristic of the Markov process, a full track of the elapsed service time will be required in order to determine the system future state. In other words, a two dimensional vector is used to describe the system state. Similarly, the carried approach can also be further moved to analyze the nonexponential *patience* case. In such case, we have to extend the system state description to include the elapsed time of all waiting bursts, i.e., an $m + 1$ dimensional system state must be generated.

Finally, as defined in [11], the overall MAC layer burst loss probability $P_{\text{MAC-Loss}}$ is the probability that a CP fails in reserving a resource at any node along the way from the ingress node to the egress node. Furthermore, assume a uniformly distributed traffic across the whole network, i.e., the offered load is the same for all nodes. Then $P_{\text{MAC-Loss}}$ can be found as follows:

$$P_{\text{MAC-Loss}} = 1 - (1 - P_{\text{Node-Loss}})^H \quad (19)$$

where H is the expected number of core nodes that the CP path can contain. To consider the worst case, which is equivalent to the maximum MAC loss probability, then H must be maximized to represent the longest CP path.

D. Throughput

The second valuable parameter to measure the system performance would be the *steady-state system throughput* β , which is defined to be the number of successful bursts within a time interval equal to the burst duration. Thus

$$\begin{aligned}
 \beta &= (\text{Average arrival/Burst duration}) \\
 &\times \text{Probability of success} \\
 &= \lambda \times \text{Burst duration} \times (1 - P_{\text{Node-Loss}}). \quad (20)
 \end{aligned}$$

IV. NUMERICAL RESULTS

So far we derived a mathematical model that describes the buffering feature proposed in this paper. In order to get more insight into the problem, we employ (1), (3), (6), (8), (14), (19), and (20) to present a numerical evaluation of this model. In this evaluation, we assume a buffer size $m = 5$, an average burst

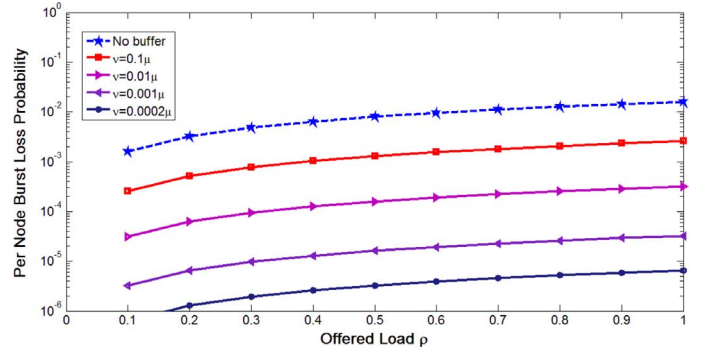


Fig. 5. Semilog plot of the *per node loss probability* versus the offered load ρ with no buffer and with buffer under different values of ν .

length $L_b = 100$ kbits, $H = 10$ hops, and apply this proposal to a WDM system with 62 channels with bit rate of 100 Gbps for each single user.

First, the per node burst loss probability is plotted in Fig. 5 versus the offered load under different values of ν . Needless to say, the per node burst loss probability increases with offered load, since the higher the offered load, the more expected to find resources reserved. It is rather more important to note that the per node burst loss probability in the absence of a buffer is reduced by adding a buffer to the system. In other words, the presence of a buffer improves the system performance. Moreover, inspecting Fig. 5, we find that the burst loss probability curve is improved by reducing the reneging rate, i.e., by increasing the average patience time. This is quite expected, as this means that the CP is allowed to wait longer time in the queue before quitting. Simply stated, it will be more likely for the required resource to be released before the core node discards the buffered CP. In other words, the reneging probability, which is a main contributing factor to the burst loss probability, is inversely proportional to the waiting (patience) time. Accordingly, since the patience is proportional to the burst length, we may also deduce that longer bursts have less blocking probability, given fixed offered load and fixed reneging rate. This aspect can be verified by computing the conditional burst loss probability assuming a specific patience x . For example, in the simple $M/M/1/1+1$ case, (9) is still applicable except for $P(R_1)$ which should be replaced by $P(R_1|X = x)$. The last quantity is obtained using (7) by integrating with respect to t only to get $P(R_1|X = x) = e^{-\mu x}$. Obviously, this result reveals that the loss probability is a decreasing function of the patience and hence a decreasing function of the burst length. Similar results can be obtained for larger buffer sizes. As stated earlier, the purpose behind this preferential treatment to longer bursts is that these bursts represent the aggregation of larger number of packets and their loss leads to more serious service degradation.

Next, in Fig. 6 the steady-state throughput is portrayed versus the average burst arrivals per burst duration when no buffer is introduced and in presence of buffer with two different values of ν . Observing this figure, we find a normal behavior of the system, in which the system throughput increases rapidly with small values of average burst arrivals, then gradually as the number of arrivals increases. Furthermore, comparing the three curves, we find that this behavior improves when a

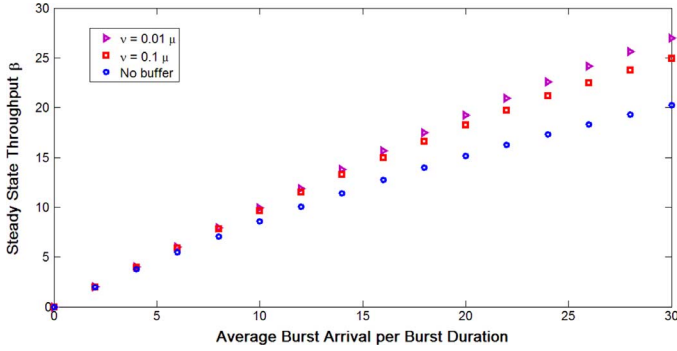


Fig. 6. Plot of the *steady-state throughput* versus the average burst arrival per burst duration with no buffer and with buffer under different values of ν .

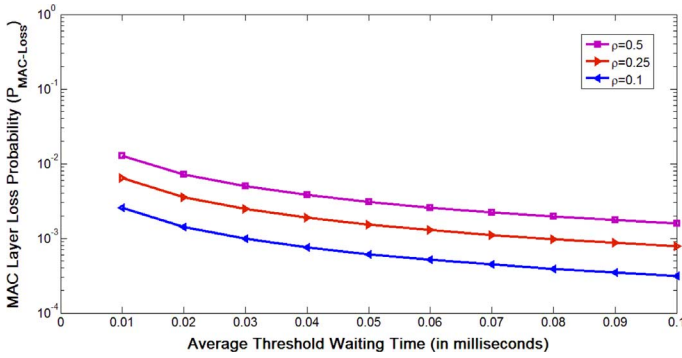


Fig. 7. Plot of the *MAC layer loss probability* versus the average threshold waiting time for different values of offered load.

buffer is added and when the average threshold waiting time is increased. This interesting effect appears and becomes more obvious with the grow in the average arrivals. That is, the proposed feature makes the system capable of handling higher traffic and allowing the control packet to wait longer time in the buffer strengthens this capability.

In Fig. 7 the explicit relationship between the MAC burst loss probability and the average threshold waiting time is illustrated. This figure indicates that the aforementioned improvement in the system behavior will be at the expense of the delay that the burst would experience. Thus, it is more reasonable to use this feature in presence of partial conversion. This way the system performance would be enhanced with small values of ν and limited number of resource converters. Furthermore, consider the following argument that would justify the introduced delay: Let t_1 represent the average CP buffering delay and t_2 refer to the average retransmission delay (assume it is composed of only the NACK propagation delay, which is the time elapsed between the generation of the NACK in the core node where the collision occurred and the start of the burst retransmission at the ingress node, ignoring processing delay). Indeed, if $t_2 \gg t_1$, then the CP buffering delay will be worthy. As stated earlier, t_2 is a function of the network dimensions. Hence t_2 can reach an average of several milliseconds as the average link length in long haul networks extends to several hundreds of kilometers (typically, the propagation delay is 2.5 ms for a 500-km link, and the link may exceed 2000 km yielding $t_2 > 10$ ms) while, as seen in Fig. 7, t_1 is much lower (t_1 of 0.1 ms would reduce the MAC layer burst loss probability from 10^{-1} with no buffer

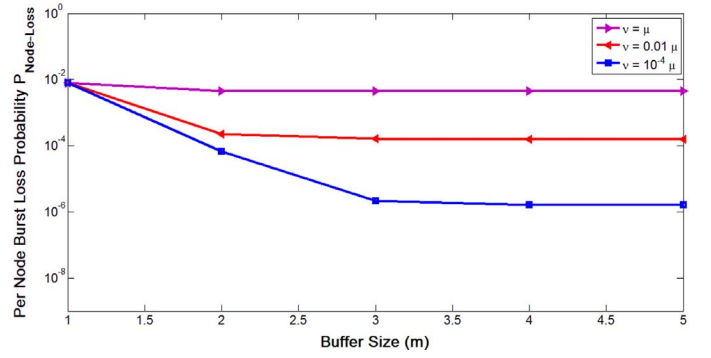


Fig. 8. Semilog plot of the *per node burst loss probability* versus the buffer size for different values of ν .

to $\approx 10^{-3}$). Moreover, using (5) to calculate π_0 , we find that for $\nu = 0.1 \mu$ and in the worst case when the offered load approaches 1, $\pi_0 \gg 0.99$. This means that less than one of every 100 CPs would be saved in the buffer and hence the effect of the waiting time on the traffic of the precedent nodes can be safely neglected.

Finally, a plot of the per node blocking probability versus the buffer size is presented in Fig. 8 for different values of ν . We observe that for a given value of ν , increasing the buffer size reduces the blocking probability until a certain buffer size is reached. This is the maximum value of the buffer size for this value of ν as adding more space beyond this maximum value will be useless. For example, for $\nu = 0.01 \mu$ the blocking probability is rapidly decreased when the buffer size is increased from 0 to 1, then slowly from 1 to 2, and it is almost constant from 2 to 3. Consequently, in this case there is no need to increase the buffer size beyond 3. In order to understand this effect, recall that the queuing time is limited. Thus the chance of an arriving CP that joins the queue to reach the server strongly depends on the length of the queue. This chance decreases as the queue grows up till it vanishes when the queue length reaches a certain value that is related to the average waiting time, i.e., this value depends on ν .

V. CONCLUSION

In this paper we have proposed a new solution to the contention problem in OBS networks by means of control packet buffering. This suggestion can be easily implemented above an OCDMA or WDM optical layer without any extra requirements. Moreover, the buffering time is restricted to a certain value and the offset time is increased accordingly. In addition, a detailed mathematical model of the proposal is explained. The most interesting part in this proposal is that this buffer is implemented in the electronic domain. This way the proposal has damped the system complexity accompanied with optical domain solutions, e.g., code or wavelength converters, fiber delay lines (FDLs), etc. Furthermore, the limitation we put on the queuing time is flexible and can be easily adjusted according to the average burst duration or priority. For example, this feature might be added to the SAC-OCDMA or WDM system proposed in [11] and [10], respectively, in order to enhance the system performance in presence of partial resource conversion. This way the system complexity would be strongly reduced with minor delay. It can also be used to provide a QoS to the system by assigning longer threshold time to bursts belonging to higher priority classes.

VI. FUTURE WORK

The developed mathematical model can be applied in more general cases. Hence, this model can be extended to investigate the cases of generally distributed inter-arrival, service and/or patience time. Moreover, as our results have demonstrated an improvement in system performance when adopting CP to JIT protocol, further research work can be performed to investigate the feasibility of applying the CP buffering feature to the JET protocol with a simplified algorithm. Furthermore, the novelty of the proposal necessitates a rigorous investigation of the proposed feature. Hence, we have focused on studying the CP buffering process in a single core node without identifying a specific node connectivity degree. However, our intention is to extend the work and conduct a real network simulation, in which various network topologies with different node connectivity degrees will be considered.

Finally, 100 Gbp/s-switching devices have been demonstrated in literature [17]. In addition, speedy advances in integrated circuits and analog to digital converters promise the realization of practical 100 Gbp/s in the near future. Nevertheless, adverse aspects like device nonlinearity and device high response time can degrade the physical layer performance at such high bit rate. Thus, as a future work, research can be conducted to address these issues and investigate their effects.

REFERENCES

- [1] C. Qiao and M. Yoo, "Optical burst switching (OBS)—A new paradigm for an optical internet," *J. High Speed Netw.*, vol. 8, pp. 69–84, Jan. 1999.
- [2] X. Yu, J. Li, X. Cao, Y. , and C. C. Qiao, "Traffic statistics and performance evaluation in optical burst switched networks," *J. Lightw. Technol.*, vol. 22, no. 12, pp. 2722–2738, Dec. 2004.
- [3] X. Yu, Y. Chen, and C. Qiao, "Study of traffic statistics of assembled burst traffic in optical burst switched networks," in *Proc. Optical Net. Commun. Conf. (OptiComm)*, 2002, pp. 149–159.
- [4] S. Oh and M. Kang, "A burst assembly algorithm in optical burst switching networks," in *Proc. OFC*, 2002, pp. 771–773.
- [5] T. Battestilli and H. Perros, "An introduction to optical burst switching," *IEEE Commun. Mag.*, vol. 41, no. 8, pp. S10–S15, Aug. 2003.
- [6] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "Jump-start: Ajust-in-time signaling architecture for WDM burst-switched networks," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 82–89, Feb. 2002.
- [7] J. Y. Wei and J. 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.
- [8] H. M. H. Shalaby, "A simplified performance analysis of optical burstswitched networks," *J. Lightw. Technol.*, vol. 25, no. 4, pp. 986–995, Apr. 2007.
- [9] M. H. S. Morsy, M. Y. S. Sowailam, and H. M. H. Shalaby, "Enhanced mathematical model for both throughput and blocking probability of optical burst switched networks," *SPIE J. Opt. Eng.*, vol. 48, pp. 1–10, Aug. 2009.
- [10] N. Akar, E. Karasan, and K. Dogan, "Wavelength converter sharing in asynchronous optical packet/burst switching: An exact blocking analysis for Markovian arrivals," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 12, pp. 69–80, Dec. 2006.
- [11] M. Y. S. Sowailam, M. H. S. Morsy, and H. M. H. Shalaby, "Employing code domain for contention resolution in optical burst switched networks with detailed performance analysis," *J. Lightw. Technol.*, vol. 27, no. 23, pp. 5284–5293, Dec. 1, 2009.
- [12] Q. Zhang, V. M. Vokkarane, Y. Wang, and J. Jue, "Evaluation of burst retransmission in optical burst switched networks," in *Proc. Broad-Nets*, 2005, pp. 276–282.

- [13] Y. Hirota, H. Tode, and K. Murakami, "A novel RWA cooperation method considering retransmission in optical burst switched networks," in *Proc. Int. Conf. Photon. Switching*, 2006, pp. 1–3.
- [14] L. Kleinrock, *Queueing Systems*. New York: Wiley, 1975, vol. 1, Theory.
- [15] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory*, 3rd ed. New York: Wiley, 1998.
- [16] O. Garnett, A. Mandelbaum, and M. Reiman, "Designing a call center with impatient customers," *Manufact. Service Oper. Manag.*, vol. 4, pp. 208–227, 2002.
- [17] X. Yu, Y. Chen, and C. Qiao, "Demonstration of a highly scalable 100-gbps OTDM computer interconnect with rapid inter-channel switching capability," in *OFC/IOOC Tech. Dig.*, 1999, pp. 165–167.

Ahmed I. Abd El-Rahman (S'11) was born in Alexandria, Egypt, in 1979. He received the B.S. degree from Alexandria University, Alexandria, Egypt, in 2002, and is currently working toward the M.S. degree from the Department of Electrical Engineering, Alexandria University.

In 2002, joined Fardos, Alexandria, Egypt, for investment and constructions as a Project Assistant Manager. His current research interests include optical burst-switched networks, optical orthogonal frequency division multiplexing, optical CDMA, and wireless systems.

Sherif I. Rabia received the B.Sc. degree in computer science and M.Sc. and Ph.D. degrees in engineering mathematics from Alexandria University, Alexandria, Egypt, in 1992, 1997 and 2001, respectively.

Currently, he is an Associate Professor at the Department of Engineering Mathematics and Physics, Alexandria University. His research papers appeared in the *Alexandria Engineering Journal*, *Journal of the Korean Statistical Society*, *Computer and Mathematics with Applications Journal*, *International Journal of Applied Mathematics and Statistics*, and the *Transactions on Computational Sciences*. His research interest falls in the area of applied probability theory and mainly in queuing theory. This includes analytical, numerical, and algorithmic analysis of queuing models as well as stability and control of such models.

Hossam M. H. Shalaby (S'83–M'91–SM'99) was born in Giza, Egypt, in 1961. He received the B.S. and M.S. degrees from Alexandria University, Alexandria, Egypt, in 1983 and 1986, respectively, and the Ph.D. degree from the University of Maryland at College Park in 1991, all in electrical engineering.

In 1991, he joined the Electrical Engineering Department, Alexandria University, and was promoted to Professor in 2001. Currently he is on leave from Alexandria University, where he is a Professor with the Department of Electronics and Communications Engineering, School of Electronics, Communications, and Computer Engineering, Egypt-Japan University of Science and Technology (E-JUST), New Borg EL-Arab City, Alexandria, Egypt. From December 2000 to 2004, he was an Adjunct Professor with the Faculty of Sciences and Engineering, Department of Electrical and Information Engineering, Laval University, Quebec, QC, Canada. From September 1996 to February 2001, he was on leave from the Alexandria University. From September 1996 to January 1998, he was with the Electrical and Computer Engineering Department, International Islamic University Malaysia, and from February 1998 to February 2001, he was with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include optical communications, optical CDMA, optical burst-switching, OFDM technology, and information theory.

Dr. Shalaby has served as a student branch counselor at Alexandria University, IEEE Alexandria and North Delta Subsection, from 2002 to 2006, and served as a chairman of the student activities committee of IEEE Alexandria Subsection from 1995 to 1996. He received an SRC fellowship from 1987 to 1991 (Systems Research Center, Maryland), State Excellence Award in Engineering Sciences in 2007 (Academy of Scientific Research and Technology, Egypt), Shoman Prize for Young Arab Researchers in 2002 from The Abdul Hameed Shoman Foundation, Amman, Jordan; the State Incentive Award in Engineering Sciences both in 1995 and 2001 from the Academy of Scientific Research and Technology, Egypt; the University Excellence Award in 2009 from Alexandria University; and the University Incentive Award in 1996 from Alexandria University.