Evaluation of preemption probabilities in OBS networks with burst segmentation

Eduardo Magaña, Daniel Morató, Mikel Izal and Javier Aracil
Universidad Pública de Navarra, Dept. de Automática y Computación
Campus Arrosadia, 31006 Pamplona, Spain
Tel. +34 948 169853
Email: {eduardo.magana, daniel.morato, mikel.izal, javier.aracil}@unavarra.es

Abstract

Preemption techniques have been recently proposed for service differentiation in Optical Burst Switching (OBS) networks. According to [1], an incoming burst *with the same priority that the burst in service* will preempt the wavelength if the residual length of the burst in service is smaller than the incoming burst transmission time. For a network scenario with no wavelength conversion, the preemption probability is evaluated assuming Exponential, Gaussian and Pareto-distributed burst sizes. Knowledge of the preemption dynamics at an OBS switch is a fundamental issue in performance evaluation, since the downstream switches will surely be affected. An analytical upper bound is provided, that shows that the preemption probability depends on the burst size distribution, which in turn depends on the burst assembly technique used at the network edges. On the other hand, not only truncated bursts result from preemption, as reported in other studies, but also the burst size distribution for preempting bursts is shifted to larger values.

Index Terms

Optical Burst Switching, priority schemes, service differentiation, preemption, performance evaluation.

Evaluation of preemption probabilities in OBS networks with burst segmentation

I. INTRODUCTION

Optical burst switching is a transfer mode that provides intermediate granularity between circuit switching and packet switching [2]. Several packets are encapsulated in an optical burst, which is preceded by a Burst Control Packet (BCP). Such BCP is in charge of setting up resources for the incoming burst beforehand.

A most interesting feature of OBS is the flexibility in resource allocation. In fact, two different priority schemes have been proposed. The first priority scheme is implemented by adequately tuning the offset time between the BCP and the actual burst [3]. Broadly speaking, bursts with higher offset times are given more chances to reserve a wavelength. As an alternative, preemption techniques can also be adopted to provide differentiated quality of service. For such techniques, high priority bursts may preempt the wavelength from low priority bursts [1]. This paper focuses on preemption for differentiated quality of service.

On the other hand, the use of *burst segmentation* provides significant throughput advantages in combination with preemption. The fact that a burst is composed by several packets makes it possible to drop *part* of a burst, so that the remaining packets may continue transmission in subsequent hops. With burst segmentation, either the head of the incoming burst or the tail of the burst in service can be dropped. Thus, in case of partial overlapping of two contending bursts, there is no need to drop the entire burst. It has been shown that the burst segmentation technique provides significant throughput benefits and allows for a higher flexibility in quality of service allocation, by placing packets either towards the burst tail or head [1].

In this paper, the case of two contending bursts *with the same priority* is considered, for an optical switch with burst segmentation capabilities. More specifically, the algorithm proposed in [1, pp. 1201,Table I] is considered, namely for bursts *within the same priority class* in a bufferless switch with no wavelength conversion:

- If the residual length of the burst in service is larger than the incoming burst length then the burst in service wins the contention. The incoming burst is dropped (either entirely or partially -head-).
- If the residual length of the burst in service is smaller than the incoming burst length then the incoming burst wins the contention. The burst in service is segmented and the tail is dropped.

Such policies correspond to the *Drop Policy* (DP) and *Segment and Drop Policy* (SDP) in [1].

In this paper, the preemption probability, or probability that the incoming burst wins the contention, is evaluated, within the same priority class. This probability is relevant for OBS network engineering for a twofold reason. First, since the incoming burst and the burst in service contend for the same resources, it is likely that they both follow the same route. Thus, due to tail dropping upon preemption, packet misordering may occur. Second, optical networks are limited by the so-called "electronic bottleneck". If preemption occurs, the optical switch must drop the tail of the burst in service and then switch the contending burst to the corresponding wavelength. This implies a processing cost not only in the optical domain but also in the electronic domain. Actually, additional signalling must be created to re-schedule bursts in the downstream nodes. Another control packet called "trailer" [1] is sent as soon as preemption happens in order to update scheduling information for the rest of OBS switches. Since this implies a processing cost, the likelihood of preemption becomes a relevant issue in OBS network performance.

A. Assumptions

In what follows, the following assumptions will be made. First, it will be assumed that the switching time is negligible in comparison to the average burst length. This is also assumed in other papers, see for instance [4], [5]. Actually, switching time has been strongly reduced due to the phenomenal advances in optical technology, specially in Semiconductor Optical Amplifiers (SOAs). For example, a switching time of 200 ps to 2 ns is reported in [6]. On the other hand, one may argue that the size-based preemption scheme proposed in [1], and adopted in this paper, may favor those users that create larger bursts, since the preemption scheme is basically size-based. We believe that this is not the case, since the burstifier is *part of the optical network and not user equipment*. Thus, burst assembly is not within the user responsibility and unfair behavior is not possible.

In order to perform the analysis, both the burst arrival process and burst size distribution must be determined. Concerning the burst arrival process, it has been shown [7] that the burst arrival process can be assumed to be Poisson, despite the possible long-range dependence of incoming traffic. Concerning the burst size distribution, not only it depends on the burst assembly algorithm which is used at the burstifiers but it also depends on burst truncation at the upstream nodes. As the bursts traverse OBS switches, the burst size distribution is likely to change. However, a burst size distribution must be selected for the analysis. While some of our results are valid *for any burst size distribution with finite first moment*, other results require explicit knowledge of the burst size distribution. In that case, for analytical tractability, it will be assumed that the incoming burst size distribution is "fresh", namely, the size

distribution is determined by the burstifier only and it is not biased by truncation or preemption.

More specifically, timer-based schemes [8] will be assumed for the burstifier. Incoming packets are stored in per-destination queues and a timer is started with the first packet in a queue. Upon timeout, packets are encapsulated in an optical burst and relayed to the first optical switch downstream. For such timer-based schemes, we have shown that the burst size turns out to be (truncated) Gaussian-distributed [7]. Other non-Gaussian burst size distributions that have also been considered in the literature are the exponential distribution [9], the hyperexponential distribution [10], and the Pareto distribution [10]. Three different distributions will be considered in this paper: Exponential, Pareto and Gaussian, thus covering a broad range of burst assembly techniques.

To the best of our knowledge, analytical results for burst segmentation performance evaluation have been restricted to exponential burst sizes [1], [11]. Note that both the Gaussian and Pareto distribution do not fulfill the memoryless property. This complicates matters for the analysis, in contrast to the exponential case.

B. Network scenario

An OBS network is depicted in Figure 1. Incoming IP packets are assembled into bursts by the edge nodes. Such packets are received from the different access networks connected to the edge node. After burst assembly, bursts are relayed to a core node, preceded by the BCP, that is transmitted out-of-band through an independent channel (wavelength). The BCPs are processed electronically in the Switch Control Unit (SCU), which, in turn, is in charge of configuring the Optical Cross Connect (OXC) matrix. To do so, the different burst segmentation policies are taken into account. In case of preemption, the OXC will switch the incoming burst into a particular wavelength and drop the tail of the burst in service.
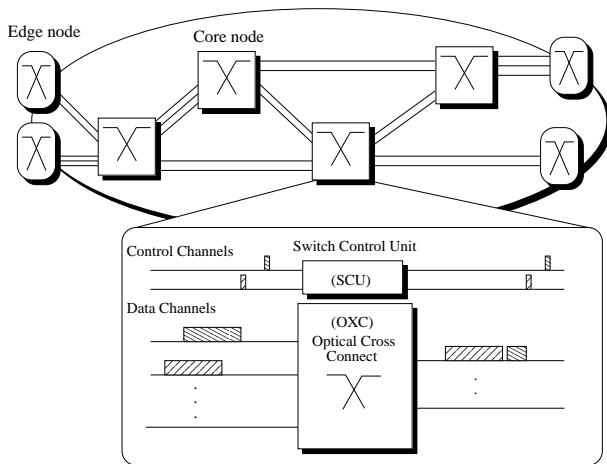

Fig. 1.   Network scenario

It will be assumed that no wavelength conversion capability is available at the switch. Thus, the analysis is restricted to the case of bursts contending for the same wavelength. Let us also assume that the switch is bufferless (no Fiber Delay

Lines -FDLs-) and that the offset between burst and BCP is fixed (which is a usual assumption in the literature, see for instance [1], [11]). Furthermore, wavelength reservation is performed in a *Just Enough Time* (JET) fashion [2]. Finally, due to preemption, the BCP may contain outdated information about the burst size, which becomes shorter. Thus, additional signalling must be employed in order to re-schedule bursts in the downstream nodes. For instance, [1] advocates for the use of trailers. Such signaling issues are out of the scope of this paper.

The rest of the paper is organized as follows: section II presents the analysis and section III is devoted to results and discussion. Finally, section IV presents the ongoing work and conclusions.

II. ANALYSIS

This section will be devoted to the preemption probability analysis. First, a most important burst segmentation property is stated. Let us assume that the wavelength is free and let (t_0, l_0) be the arrival time and the length of the first burst in a busy period. Such burst will be served immediately. Let $(t_i, l_i), i = 1, \ldots, n$ be the arrival times of subsequent bursts that arrive *during* the service time of burst 0, i.e., $t_0 < t_i < t_0 + l_0$ for all $i = 1, \ldots, n$. It must be noted that burst i preempts burst 0 if

$$l_i > l_0 - \sum_{j=1}^{i}(t_j - t_{j-1}) = l_0 - (t_i - t_0) \qquad (1)$$

and $l_k \leq l_0 - \sum_{j=1}^{k}(t_j - t_{j-1}), k = 1, \ldots, i-1$. Figure 2 serves to illustrate the introduced notation.
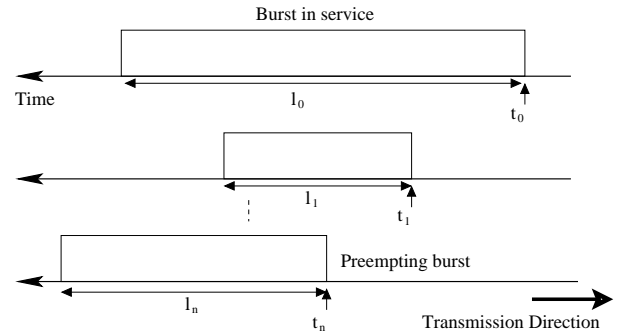

Fig. 2.   Notation

Note that (1) is fulfilled iff the residual length of the burst in service at the time of arrival of burst i is lower than the burst i length [1]. Assume that preemption occurs and let (t_*, l_*) denote the arrival time and length of the burst that wins the contention. For example, assume that bursts $i = 1, \ldots, n-1$ do not fulfill (1) but burst n does. Then, bursts $i = 1, \ldots, n-1$ are dropped from the switch and burst n wins the contention, i. e. $t_* = t_n$, $l_* = l_n$ and $l_n > l_0 - \sum_{j=1}^{n}(t_j - t_{j-1})$.

We now show that the service time distribution of burst (t_*, l_*) is shifted to larger values in comparison to the burst 0 counterpart. Intuitively, the preempting burst has a larger probability of high service times, in comparison to the burst in service. This is shown in the following theorem.

Theorem *1:* Let (t_0, l_0) be the arrival time and service time of the first burst in a busy period. Let F be the common burst service time distribution. Let us assume that preemption occurs and let (t_*, l_*) represent the arrival time and length of the burst that wins the contention, then $P(l_* > x) > P(l_0 > x)$ for all $x > 0$.

Proof: Let us consider the denumerable set $\Phi_n = \{(t_0, l_0), (t_1, l_1), \ldots, (t_n, l_n)\}, n = 1, 2, \ldots$ of all possible arrivals of n bursts *during* burst 0 service time, namely in the interval $(t_0, t_0 + l_0)$, such that burst n wins the contention. On the other hand, since F is the service time distribution, l_0 is distributed according to F. Let L be a random variable with distribution F. First, note that

$$P(l_* > x) = \sum_{n=0}^{\infty} \sum_{\Phi_n} P(l_n > x | \Phi_n) P(\Phi_n) \pi_n \qquad (2)$$

where the operator $\sum_{\Phi_n}$ represents the sum over all possible sets of n arrivals Φ_n and π_n is the probability of n arrivals in $(t_0, t_0 + l_0)$. Note that π_n is a Poisson measure and $\sum_{n=0}^{\infty} \sum_{\Phi_n} P(\Phi_n)\pi_n = 1$, since preemption is assummed to occur a. s. On the other hand,

$$P(l_n > x | \Phi_n) = P\left(L > x | L > l_0 - \sum_{j=1}^{n}(t_j - t_{j-1})\right)$$
$$> P(L > x | L > 0) = P(l_0 > x) \qquad (3)$$

Then, substitute (3) into (2) to obtain $P(l_* > x) > P(l_0 > x) \sum_{n=0}^{\infty} \sum_{\Phi_n} P(\Phi_n)\pi_n = P(l_0 > x)$ and the Theorem is proved $\square$.

Note that *the theorem does not make any assumption about the burst length distribution*. Therefore, the theorem applies to any burst length distribution, regardless of possible truncation in the nodes upstream. Intuitively, equation (3) implies that l_* is not distributed according to F but, on the contrary, the distribution of l_* provides larger service time values. On the other hand, the time instants t_0 are renewal epochs of the system. Thus, the preemption probability is only affected by the system dynamics during the busy period. From theorem 1 it turns out that preemption is less likely to occur for the burst that wins the contention than for the first burst in a busy period (burst 0). Note that the latter is (distributionally) shorter than the burst that wins the contention and, thus, it is easier for incoming bursts to preempt the wavelength. Hence, *the preemption probability reaches a maximum with the first burst in a busy period*. Such probability is given by

$$P(L > A) = \int_0^{\infty} P(L > x) dF_A(x) \qquad (4)$$

where A is a random variable that provides the residual life of the server (wavelength). In the following subsections, an upper bound for the preemption probability will be derived for an Exponential, Pareto and Gaussian distribution. Note that the density of the residual life of the service time L is given by [12, pp. 172, vol I]

$$f_A(x) = \frac{P(L > x)}{E[X]}, x > 0 \qquad (5)$$

This expression will be used frequently in the following sections. Recall that the burst length will be represented by L.

A. Exponential service times

For exponentially-distributed service times the residual life is also exponential due to the memoryless property and, obviously:

$$P(L > A) = \int_0^{\infty} e^{-\lambda x} \lambda e^{-\lambda x} dx = 1/2 \qquad (6)$$

B. Pareto service times

A Pareto distribution has the form

$$P(L > x) = 1 \qquad x < K$$
$$P(L > x) = K^{\alpha} x^{-\alpha} \qquad x \geq K \qquad (7)$$

thus, from (5) the residual life is given by (see [13] for details)

$$P(A > y) = \frac{(\alpha - 1)(K - y) + K}{\alpha K}, 0 \geq y \geq K$$
$$P(A > y) = \frac{K^{(\alpha-1)}}{\alpha} y^{(1-\alpha)}, y > K \qquad (8)$$

and, from (4) it turns out that

$$P(L > A) = \frac{2(\alpha - 1)}{2(\alpha - 1) + 1} \qquad (9)$$

Note that the result depends on the distribution parameter α, in contrast to the exponential case.

C. Gaussian service times

Following our previous work in [13], the residual life of a Gaussian service time [1] with mean μ and variance σ^2 is given by

$$P(A > y) = \frac{1}{\mu} \int_y^{+\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$
$$- \frac{1}{\mu} \int_y^{+\infty} y \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx \qquad (10)$$

Direct calculation of the preemption probability following (4) is unfeasible. However, note that (4) can be written as

$$P(L > A) = \frac{1}{\mu} \int_0^{\infty} P(L > x)^2 dx \qquad (11)$$

since, following (5), $dF_A(x) = f_A(x)dx = (P(L > x)/\mu)dx$.

Now, it must be noted that the mean of the first order statistic of a sample of two i.i.d random variables is given by $\mu_{(1)} = \int_0^{\infty} P(L > x)^2 dx$ [14, chapter 1]. Thus,

$$P(L > A) = \frac{\mu_{(1)}}{\mu} \qquad (12)$$

---

[1]It will be assumed truncation to positive values.

The mean of the first order statistic $\mu_{(1)}$ can be approximated using Taylor expansion techniques [15]. Let $p_1 = 1/3$ and $q_1 = 2/3$ and assume that L has a density f and a distribution F. Let $G_1 = F^{-1}(p_1)$ and $f_1 = f(G_1), f_1' = f'(G_1), \ldots$, then

$$\mu_{(1)} = F^{-1}(p_1) - \frac{f_1' p_1 q_1}{2f_1^3(n+2)} +$$
$$\left\{ \frac{3(f_1')^2 - f_1 f_1''}{3f_1^5} \frac{p_1 q_1 (q_1 - p_1)}{(n+2)(n+3)} \right\} +$$
$$\left\{ \frac{10 f_1 f_1' f_1'' - f_1^2 f_1''' - 15(f_1')^3}{8f_1^7} \right\} \frac{n p_1^2 q_1^2 + p_1 q_1 (2 - 5 p_1 q_1)}{(n+2)(n+3)(n+4)}$$
$$+ \ldots \tag{13}$$

Now, substitute (13) into (12) to obtain an approximation for the preemption probability. The numerical experiments show that such preemption probability only depends on the coefficient of variation $c_v = \sigma^2/\mu^2$. Note that the use of order statistics provides a novel methodology for performance evaluation with generic burst size distribution.

III. RESULTS AND DISCUSSION

Extensive numerical and simulation experiments have been performed in order to verify the analysis. The simulation parameters are selected in such a way that traffic load remains constant regardless of the distribution type. In all cases, the mean burst transmission time is made equal to 12 μs, which corresponds to an average burst size of 15 Kbytes (average file size in the Internet [16]) in a 10 Gbps wavelength. Note that switching times for SOAs are in the vicinity of nanoseconds, as reported by [6], thus making the switching time negligible. By adequately choosing the burst arrival rate, different wavelength utilization factors (ρ) are considered (low load, $\rho = 0.2$, and high load, $\rho = 0.8$). Simulation parameters are summarized in table I where c_v is the coefficient of variation. In any case, since analytical expressions are provided, the upper bound for the preemption probability can be obtained for any combination of arrival rate and burst size distribution parameters.

Burst size distribution	Parameters	Utilization
Exponential	$mean = 12\mu s$	$\rho = (0.2, 0.8)$
Pareto	$mean = 12\mu s, \alpha = [1.5, 2]$	$\rho = (0.2, 0.8)$
Gaussian	$mean = 12\mu s, cv = [0.01, 0.2]$	$\rho = (0.2, 0.8)$

TABLE I

SIMULATION PARAMETERS

First, following theorem 1, figure 3 shows the distribution of l_* in comparison to the distribution of l_0, i.e. the figure shows the service time distributions corresponding to preempting bursts and bursts arriving when the wavelength is unoccupied. As can be observed from figure 3, theorem 1 is confirmed by the simulation results. The same conclusion is obtained with other distribution parameters.

Second, figure 4 shows preemption probability for Pareto and Gaussian service times. The upper bound is given by
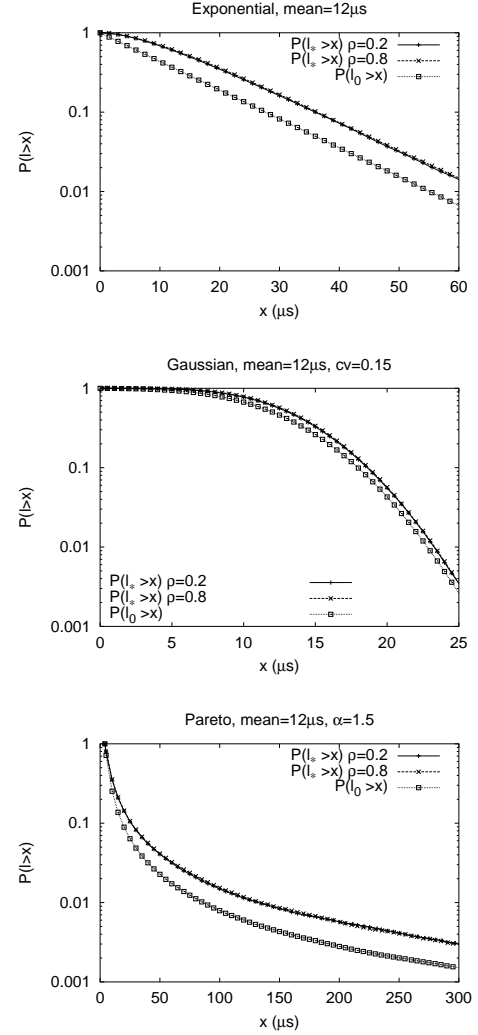

Fig. 3. Comparison between $P(l_* > x)$ and $P(l_0 > x)$ (Exponential -top-, Gaussian -center-, Pareto -bottom-)

(9) and (12). Note that simulation results provide probability values below the upper bound, for both Pareto and Gaussian cases. Concerning the exponential distribution, the preemption probability does not depend on the distribution parameter and is equal to 0.48 and 0.43 for utilization factors equal to 0.2 and 0.8. This is also smaller than the upper bound of 0.5 provided by (6).

Figure 4 shows that simulation results match the upper bound analytical expressions. As the utilization factor decreases, the busy periods tend to be shorter. Thus, the system behavior is closer to the best case that was assumed for the upper bound derivation, i.e. preemption of the first burst in a busy period. Hence, the upper bound becomes closer to the simulation results.

The above results give raise to the following discussion. Preemption probabilities are highly dependent on the burst length distribution. Hence, for the same traffic load, the preemption probability depends on the burst assembly technique that is used at the network edges. Furthermore, it must be noted that the preemption probability increases as the length variability decreases (higher α for the Pareto distribution and lower c_v for
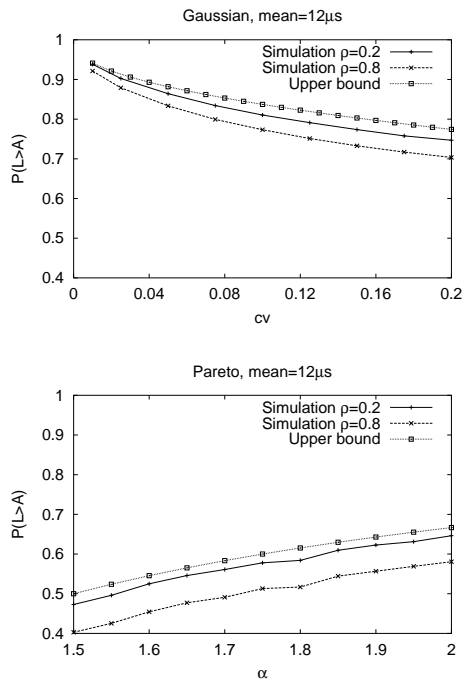
Fig. 4. Preemption probability (Gaussian -top- and Pareto -bottom-)

the Gaussian distribution). As a conclusion, the burst assembly algorithm has a strong impact on the burst segmentation dynamics in the optical network core. Furthermore, theorem 1 and figure 3 show an interesting *low-pass filtering* effect on the preempting bursts, due to the fact that the size distribution is shifted to larger values. On one hand, the average burst size tends to decrease as intermediate switches are traversed [11], due to truncation. On the other hand, the preempting bursts tend to have larger sizes. Thus, comparison of the burst size distribution at the input and output of a switch featuring burst segmentation appears as a promising research topic.

IV. CONCLUSIONS AND ONGOING WORK

In this paper, an evaluation of the burst preemption probability has been performed, for preemption techniques based on maximization of the residual life. However, this analysis is restricted to the case of no wavelength conversion. Our current research focuses on the wavelength conversion case, and upper bounds for the preemption probability are being obtained. On the other hand, the fact that the size distribution of the preempting burst is skewed to higher values deserves further analysis, due to the performance impact on the downstream OBS switches. This is also subject of our current and future research.

REFERENCES

[1] V. Vokkarane and J. Jue, "Prioritized burst segmentation and composite burst assembly techniques for qos support in optical burst-switched networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 7, pp. 1198–1209, September 2003.
[2] C. Qiao and M. Yoo, "Optical burst switching (OBS) - A new paradigm for an optical Internet," *Journal of High-Speed Networks*, vol. 8, no. 1, 1999.
[3] M. Yoo and C. Qiao, "Supporting multiple classes of services in IP over WDM networks," in *Proceedings of GLOBECOM 1999*, Rio de Janeiro, Brazil, 1999.
[4] H. Buchta, C. Gauger, E. Patzak, and J. Saniter, "Limits of effective throughput of optical burst switches based on semiconductor optical amplifiers," in *Proceedings of Optical Fiber Communication conference OFC2003*, Atlanta, March 2003.
[5] C. M. Gauger, "Trends in optical burst switching," in *Proceedings of SPIE ITCOM 2003*, Orlando, September 2003.
[6] X. Ma and G.-S. Kuo, "Optical switching technology comparison: Optical mems vs. other technologies," *IEEE Communications Magazine*, vol. 41, no. 11, pp. S16–S23, November 2003.
[7] M. Izal and J. Aracil, "On the influence of self-similarity in optical burst switching traffic," in *Proceedings of GLOBECOM 2002*, Taipei, Taiwan, 2002.
[8] A. Ge, F. Callegati, and L. Tamil, "On optical burst switching and self-similar traffic," *IEEE Communications Letters*, vol. 4, no. 3, March 2000.
[9] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP over WDM networks," *IEEE Journal of Selected Areas in Communications*, vol. 18, no. 10, pp. 2062–2071, October 2000.
[10] K. Dolzer, C. Gauger, J. Spath, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *International Journal of Electronics and Communications (AE)*, vol. 55, no. 1, 2001.
[11] V. Vokkarane and J. Jue, "Burst segmentation: an approach for reducing packet loss in optical burst-switched networks," *Optical Networks Magazine*, vol. 4, no. 6, pp. 81–89, Nov/Dec 2003.
[12] L. Kleinrock, *Queueing Systems*. John Wiley and Sons, 1975.
[13] D. Morato, M. Izal, J. Aracil, E. Magaña, and J. Miqueleiz, "Blocking time analysis of OBS routers with arbitrary burst size distribution," in *Proceedings of IEEE Globecom 2003*, San Francisco, CA, December 2003.
[14] B. C. Arnold and N. Balakrishnan, *Relations, Bounds and Approximations for Order Statistics*. Springer-Verlag, 1989.
[15] C. E. Clark and G. T. Williams, "Distributions of the members of an ordered sample," *Annals of Mathematical Statistics*, vol. 29, pp. 862–870, 1958.
[16] A. Downey, "Evidence for long-tailed distributions in the Internet," in *Proceedings of ACM SIGCOMM Internet Measurement Workshop*, 2001.