A Framework for Optical Burst Switching Network Design

Jolyon White, *Student Member, IEEE*, Moshe Zukerman, *Senior Member, IEEE*, and Hai Le Vu

***Abstract*—In this letter, we analyze optical burst switching (OBS) systems. The analysis leads to a framework which provides guidelines for OBS design. We identify conditions for OBS feasibility and the relationship between burst size, or equivalently burst assembly delay, and throughput, taking into consideration control packet processing and the number of available wavelengths per fiber.**

***Index Terms*—Burst size, optical burst switching, optimization.**

I. Introduction

OPTICAL burst switching (OBS) [1] is a step toward the ultimate goal of optical packet-switching in next-generation IP-over-WDM optical transport networks. In OBS, data packets are aggregated into much larger bursts before transmission through the network. This allows amortization of the switching overhead across many packets. The burst is preceded in time by a control packet which is sent on a separate control wavelength and requests resource allocation at each switch. When the control packet arrives at a core cross-connect (or switch), capacity is reserved in the cross-connect for the burst. If capacity can be reserved, the burst can then pass through the cross-connect. The benefit of OBS over circuit switching is that there is no need to dedicate a wavelength for each end-to-end connection. OBS is more viable than optical packet switching because the burst data does not need to be buffered or processed at the cross-connect, so that the strengths of optical switching technologies can be leveraged effectively and the problem of buffering in the optical domain (for which technology does not yet exist) is circumvented.

Optical burst switching schemes may be based on either two-way or one-way reservation protocols. Tell-And-Go (TAG) and Just-Enough-Time (JET) are examples of one-way protocols [1]. In these protocols the data burst follows the control packet after a predetermined offset time without waiting for acknowledgment of resource reservation from switches along the path. In this paper we only consider JET-based OBS where channel capacity is reserved for no longer than the time required to accommodate the duration of the data burst. In this way, JET achieves better channel utilization than schemes which reserve the channel from the time the control packet arrives at the switch (such as the Horison scheme [2]), or in which the channel is reserved until a release message is received (e.g., Just-In-Time (JIT) [3]).

Manuscript received December 11, 2001. The associate editor coordinating the review of this letter and approving it for publication was Prof. J. Evans.

The authors are with the ARC Special Research Center for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Vic. 3010, Australia (e-mail: j.white@ee.mu.oz.au; m.zukerman@ee.mu.oz.au; h.vu@ee.mu.oz.au).

Publisher Item Identifier S 1089-7798(02)05796-4.

Since data bursts are sent out without waiting for acknowledgment, the burst could be blocked and dropped due to resource contention. Therefore, one should keep the burst blocking probability (which can be modeled well using the $M/M/k/k$ queueing system [4]) under a certain predefined value when designing the OBS network.

It is important to note that there is a crucial assumption in JET-based OBS systems: as soon as the control packet arrives at the switch it is assumed that it will be processed and that an adequate resource reservation will be made for the corresponding burst. However, if several control packets arrive at the switch at the same time, or if the processing time of a particular control packet is too long, then some of the control packets must be queued at the switch. However, queueing is not desirable in this system, since the waiting time of a queued control packet will be nondeterministic and the control packet has to provide the switch with the exact time at which it expects the burst to arrive.

In OBS networks, packets are collected into a burst at the source before being sent into the network. Here, we assume that as soon as the burst is ready, the source will send the corresponding control packet. We call this delay due to aggregating packets into a burst the *burst assembly delay*. Clearly, if the burst assembly delay is too small, many control packets are sent. Thus queueing at the control processor increases, which adversely affects OBS performance.

Dolzer *et al.* [5] analyzed blocking in the data path of OBS. They used the forward recurrence time of the transmission time of low priority bursts to account for contention between low and high priority bursts in a two class system. An extension to multiple classes is found in [6]. In [7] an $M/M/1$ model is used to compute an approximation for the complementary distribution of the control packet processing delay. This is used to choose an additional compensating fixed fiber delay placed at the inlet to the switch.

In Section II of this paper we introduce a framework based on a throughput optimization problem for analyzing OBS networks. Using this framework, we optimize the burst size and identify feasibility conditions for OBS networks. This leads us in Section III to two enhanced design approaches for which the benefits are evaluated and fundamental performance limits are obtained.

II. OBS Throughput Optimization

We define our OBS framework as a constrained optimization problem. Let B be the average time it takes to transmit a burst using one wavelength and let λ be the burst arrival rate into a given output port of an optical burst switch. We assume the switch has full wavelength conversion capability. We aim to maximize $A(B) = \lambda B$, the throughput for this port. We will now consider the constraints resulting from OBS-JET operation as well as quality-of-service requirements. We begin by considering constraints imposed by the switch operation and then we follow up by discussing constraints related to burst assembly.

A. Switching Constraints

To avoid queuing of control packets we require $\Pr\{Q > 0\} \leq \epsilon_1$ where Q is the number of packets in the control processor. We model the control processor as a $G/G/1$ queue for which $\Pr\{Q > 0\} = \rho = \lambda/\mu$ (where λ will be the same for both the control packets and data bursts) and $1/\mu$ is the mean control packet service time. Assuming Poisson arrival of bursts we also limit the burst blocking probability for this port (modeled as an $M/G/k/k$ system) to be lower than a certain level ϵ_2. Thus, we can summarize the constraints as

$$\frac{\lambda}{\mu} \leq \epsilon_1 \tag{1}$$

$$E_k(A(B)) \leq \epsilon_2 \tag{2}$$

where $E_k(A(B))$ is the burst blocking probability obtained by the Erlang B formula for k servers (in this case wavelengths) and an offered load of $A(B)$.

Let $A^*(B) = \max A(B)$ subject to (1) and (2). Substituting $A(B) = \lambda B$ into (1) gives

$$A(B) \leq \mu\epsilon_1 B. \tag{3}$$

Defining $E_k^{-1}(\epsilon)$ as the inverse of the Erlang B formula, in particular it is the A value which gives a blocking probability of ϵ, we obtain by (2):

$$A(B) \leq E_k^{-1}(\epsilon_2). \tag{4}$$

Using the constraints (3) and (4) replacing (1) and (2), it is now more convenient to obtain $A^*(B)$ given by

$$A^*(B) = \begin{cases} \mu\epsilon_1 B, & B \leq B^* \\ E_k^{-1}(\epsilon_2), & B > B^*, \end{cases} \tag{5}$$

where

$$B^* = \frac{1}{\mu\epsilon_1} E_k^{-1}(\epsilon_2). \tag{6}$$

By (5) and (6) it is evident that $A(B)$ is maximized for $B \geq B^*$, and that its maximum value is

$$A_{\max} = E_k^{-1}(\epsilon_2). \tag{7}$$

B. Burst Assembly Constraint

In the previous section, we optimized the OBS system throughput subject to constraints at an output port of a particular switch. However, there is another constraint at the source due to assembling a burst from the incoming packets before sending them into the network. In the following, the burst assembly delay is considered as an additional constraint in our optimization problem. For simplicity, we consider N homogeneous sources, each with the average offered traffic $(A(B)/N)$ forwarding to the same output port in our switch.

Let t^* be the maximum allowable burst assembly delay. The value of t^* depends on the delay requirement of applications that send packets across the network. Furthermore, given the capacity per wavelength, we define $B_{\max}$ to be the maximum burst length in seconds, i.e., the maximum allowable time for one wavelength to transmit the burst [9]. In other words, we can express $B_{\max}$ as a maximum allowable transmission delay for each burst. Using t^* and $B_{\max}$ the burst assembly mechanism at the source can be described as follows.

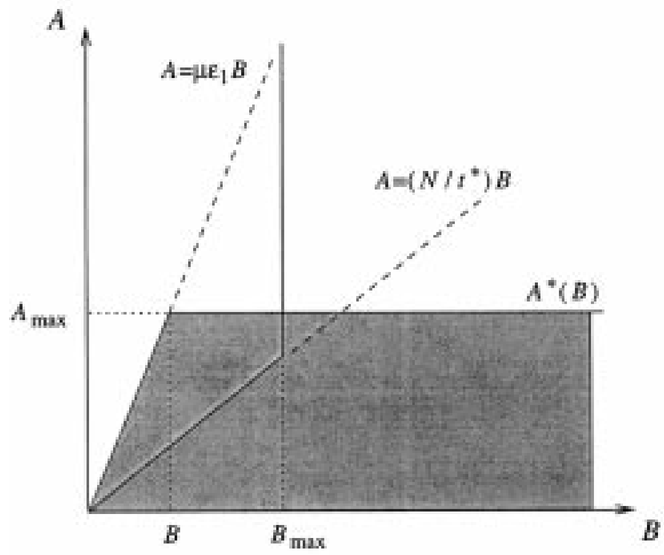

Fig. 1. Optimization of throughput versus burst size for a feasible scenario. The shaded area represents the set of feasible solutions.

In the case of light traffic, the allowable burst assembly delay t^* will expire before a burst of size $B_{\max}$ has been assembled. Then the source will send the control packet of this burst and begin collecting a new one.

In the case of heavy traffic, however, during t^* the source can collect several bursts of size $B_{\max}$. In this situation, the source does not wait until t^* expires, but sends the control packet immediately once the burst reaches its allowable maximum length $B_{\max}$. After sending the control packet, the source resets the burst assembly timer and starts to collect the new burst. Under this heavy traffic condition, the source may send several control packets and assigns more than one wavelength for the corresponding bursts during t^* time period.

As mentioned before, $(A(B)/N)$ is the mean offered traffic per source. Therefore, $(A(B)t^*/N)$ is the mean amount of traffic load added per t^* time units at each source.

We will use the following approximation for B:

$$B = \min\left(\frac{A(B)t^*}{N}, B_{\max}\right). \tag{8}$$

Clearly B is bounded by $B_{\max}$. However, under light traffic conditions $B = A(B)t^*/N$.

We have confirmed the accuracy of this approximation using a Poisson traffic model for a wide parameter range for t^* and $(A(B)/N)$, especially in the neighborhood of $(A(B)/N) = 1$.

Fig. 1 shows $A^*(B)$ together with the operating curve defined by (8). The feasible region defined by the constraints (3) and (4) is shaded in the figure. Note that the expression $A = (N/t^*)B$ is equivalent to $B = At^*/N$, corresponding to the light traffic case of (8). This figure shows that the optimal operating point is achieved when $B = B_{\max}$ and $A = A_{\max}$. At this point, throughput is maximized, as is the average burst size. This corresponds to a heavily loaded case. If $B_{\max} < B^*$ then the optimum B is $B_{\max}$ again, however we do not achieve the maximum throughput; in this case the throughput is $A = \mu\epsilon_1 B_{\max}$. Therefore, a designer should aim at $B_{\max} \geq B^*$.

Consider the case when $B_{\max} > B^*$ and $A_{\max}/B_{\max} < N/t^* < \mu\epsilon_1$. In this case, $A^*(B) = A_{\max}$. However, B is less than $B_{\max}$.

It is very important to observe that if

$$\frac{N}{t^*} > \mu\epsilon_1 \tag{9}$$

the OBS system is not feasible. In this situation, the operating curve defined by (8) is outside the feasible area.

III. Further Design Approaches

The previously defined optimization problem gives rise to certain concerns regarding the feasibility and performance of OBS. For example, let us set ϵ_1 to be very small, say, 10^{-8}, so that we have no queueing in the control section of the switch. Supposing our control processor is capable of processing a control packet every 100 ns, and suppose we have 100 sources and a burst assembly delay no more than 1 ms. Then $N/t^* = 10^5$ and $\mu\epsilon_1 = 10^{-1}$, so our system is not feasible. In this case we could relax the value of ϵ_1 by several orders of magnitude (e.g., 10^{-5}) and still not have a workable system. The only other alternative is to increase the burst assembly delay to levels which may be unacceptable for some services.

We will now provide two possible design approaches that can significantly improve OBS performance: 1) allowing queuing in the control processor and 2) synchronization of sources.

A. Queueing Control Packets

Suppose we allow a significant amount of buffering of control packets in the switch. In this case, ϵ_1 can be significantly increased. Setting $\epsilon_1 = 0.25$ (which, assuming an $M/M/1$ queue as an approximation, corresponds to $\Pr\{Q \geq 10\} = \epsilon_1^{10} \approx 10^{-6}$, i.e., a negligible chance of having a large queue) gives $\mu\epsilon_1 = 2.5 \times 10^6 > N/t^* = 10^5$ which is feasible. If we let $\epsilon_2 = 10^{-8}$ and $k = 100$ wavelengths then we have $B^* = 21.8$ μs and the optimum value of B can be anywhere between this and $B = A_{\max}t^*/N = 545$ μs, depending on the value of $B_{\max}$.

However, OBS requires strict timing of bursts relative to their control packets and if queueing is likely, then the corresponding queueing delay must be accounted for. To resolve this problem we propose that on entry to the switch each control packet must be timestamped to ensure that the offset time it contains can be interpreted meaningfully by the control processor. Further, on exit from the control processor the control packet should again be timestamped to enable the next-hop switch to determine its propagation delay. This requires accurate synchronization of all network clocks by some mechanism. Only in this way can the burst be successfully assigned without collision. In this scheme the control packet size gradually increases as it moves through the network collecting timestamps. This may cause the service rate μ to decrease, which may narrow the set of feasible solutions.

B. Synchronization

It is well known that the $D/D/1$ and $D/D/k/k$ queues perform better than their $M/M/1$ and $M/M/k/k$ counterparts. In this section we will use our framework to evaluate how much better performance can be achieved if all sources are synchronized so that our single server queue for the control packets is a $D/D/1$ queue and our $M/G/k/k$ for the bursts is a $D/D/k/k$ system. This provides an insight into the fundamental performance limits of OBS.

If this change is made then there will be no loss in the $D/D/k/k$ system as long as $A \leq k$. Therefore, $A_{\max} = k$. For the control packets, an arrival will always find the control processor idle unless $\lambda > \mu$, so constraint (1) becomes $\lambda/\mu < 1$ (i.e., in this system $\epsilon_1 = 1$). Given these values of $A_{\max}$ and ϵ_1, we have in the deterministic system $B^* = k/\mu$ and $B = \min(B_{\max}, kt^*/N)$.

In this system, we require that $N/t^* \leq \mu$ for feasibility, and we aim for $B_{\max} \geq B^*$ in order to achieve the maximum throughput of k. If $B_{\max} < B^*$ then the throughput achieved will be $\mu B_{\max}$ which is less than k. However since $\epsilon_1 = 1$ in this system (compared to 0.25 in the previous case) we still achieve significantly higher throughput than in the stochastic system, for both cases of $B_{\max}$.

For instance, with all other parameters unchanged, in this system $B^* = 10$ μs, less than half that of the stochastic case, and the maximum throughput is 100 compared with 54.5 in the stochastic system. Thus we can set $B_{\max}$ smaller than before and still achieve maximum throughput. This is a strong motivation for engineering the OBS network to have synchronized sources.

IV. Conclusion

We have identified the burst assembly delay, the control packet queueing delay and the burst blocking probability as the main constraints in JET-based optical burst switching networks. We have formulated an optimization problem aiming to maximize network throughput subject to these constraints. This optimization problem has provided insight into the performance of various OBS design options and has led to the identification of feasibility conditions and fundamental limits for OBS performance.

References

[1] C. Qiao, "Labeled optical burst switching for IP-over-WDM integration," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 104–114, Sept. 2000.

[2] J. Turner, "Terabit burst switching," *J. High Speed Networks*, vol. 8, no. 1, pp. 3–16, Jan. 1999.

[3] J. Wei, J. Pastor, R. Ramamurthy, and Y. Tsai, "Just-in-time optical burst switching for multiwavelength networks," in *IFIP Broadband Commun.*, Hong Kong, Nov. 1999, pp. 339–352.

[4] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.

[5] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *Int. J. Electron. Commun.*, vol. 55, no. 1, pp. 18–26, 2001.

[6] K. Dolzer and C. Gauger, "On burst assembly in optical burst switching networks—A performance evaluation of just-enough-time," in *Proc. 17th Int. Teletraffic Congress*, vol. 4, Salvador da Bahia, Brazil, Sept. 2001, pp. 149–160.

[7] F. Callegati, H. C. Cankaya, Y. Xiong, and M. Vandenhoute, "Design issues of optical IP routers for Internet backbone applications," *IEEE Commun. Mag.*, vol. 37, no. 12, pp. 124–128, Dec. 1999.

[8] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Control architecture in optical burst switched WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1838–1851, Oct. 2000.

[9] A. Ge, F. Callegati, and L. Tamil, "On optical burst switching and self similar traffic," *IEEE Commun. Lett.*, vol. 4, pp. 98–100, Mar. 2000.