Joint Congestion Control and Burst Contention Resolution in Optical Burst Switching Networks

Won-Seok Park, Minsu Shin, Hyang-Won Lee, and Song Chong

Abstract— **This paper revisits burst contention resolution problems in optical burst switching (OBS) networks from the viewpoint of network utility maximization. Burst collision occurs when two or more bursts access the same wavelength simultaneously, and the occurrence becomes more frequent as the offered load increases. In particular, when the network is overloaded, no contention resolution scheme would effectively avoid the collision without the help of congestion control. We formulate a joint optimization problem where two variables, the length and the time at which each burst is injected into the network, are jointly optimized in order to maximize aggregate utility while minimizing burst loss. A distributed algorithm is also developed, which explicitly reveals how burst contention resolution and congestion control must interact. The simulation results show that the joint control decouples throughput performance from burst loss performance so that burst loss ratio does not increase as network throughput increases. This is not the case in conventional contention resolution schemes where burst loss ratio increases as network throughput increases so that achievable network throughput is limited. Our work is the first attempt to the joint design of congestion and contention control and might lead to an interesting development in OBS research.**

I. Introduction

Optical Burst Switching (OBS) [1]–[7] has been proposed as a future high-speed switching technology that combines the advantages of optical circuit switching and optical packet switching. In OBS network, multiple IP packets with the same destination are assembled into a burst at an ingress OBS node and the burst is transmitted through the network core entirely in the optical domain. The ingress OBS node also sends a corresponding control packet for each data burst on a separate control channel. Because this control packet leads the data burst by an offset time and reserves wavelength resources for its data burst, OBS eliminates the need for buffering of the data burst while processing of the control packet and configuring the switch.

Although OBS has been considered as a promising solution for optical networks due to its bandwidth efficiency and implementation simplicity, it has a serious problem on the aspect of the burst contention. Since most of the current optical core nodes are not supporting optical buffers, or if any, with highly limited storage capacity, burst contention is unavoidable when two or more bursts are competing for the same wavelength resource at a time. Hence, various contention resolution schemes have been proposed from time deflection [5], space deflection [6], and wavelength conversion [1]. Unfortunately, resolving the burst contention depending solely on these reactive schemes has a fundamental limitation in that if the network is congested by increased input traffic, they cannot avoid the burst loss by the congestion and consequently their performance will be unpredictably deteriorated. An alternative approach to reduce the network contention is to proactively prevent an OBS network from entering the congestion state that invokes lots of burst contentions. Several recent schemes [4], [7] including our earlier work are designed to follow the congestion control approach [8]–[12], which is limiting the individual burst length based on the network utilization status.

However, the possibility of contention always exists when there is no congestion at all. For instance, suppose that two bursts demand the same wavelength at exactly the same time. In this case, one of them should be dropped even though there is sufficient bandwidth capacity in the core. So, the contentions are the problem of not only congestion, but also the time overlap of consecutive bursts at an optical link. For that reason, the proactive congestion control schemes [4], [7] can achieve low burst losses only when sacrificing their network utilization. If they increase target utilization threshold, the burst loss ratio is also increased. Therefore, we argue that both the network congestion and the time overlap of consecutive bursts should be jointly resolved to reach the OBS goal, i.e., maximum utilization with minimal burst loss. We can choose two controllable variables in the OBS network, burst length and additional offset time. These variables can be easily adjusted at ingress OBS nodes, for instance, the additional offset time can be implemented by using the electrical buffer.

Focusing on these two variables, we propose a joint congestion control and burst contention resolution algorithm that supports fair and efficient sharing of optical resources without burst overlap. We consider a network utility maximization problem with link capacity and burst non-overlapping constraints, where utility is defined as an increasing and concave function of the burst length. The proposed joint control algorithm can be implemented in a distributed manner where edge nodes are cooperative, and explicit signaling by the

This work was supported by the center for Broadband OFDM Mobile Access (BrOMA) at POSTECH through the ITRC program of the Korean MIC, supervised by IITA. (IITA-2006-C1090-0603-0037).

Won-Seok Park is with LG Electronics Inc., Seoul, Korea (e-mail: pwsspm1@gmail.com). Minsu Shin is with the Department of Computer Science, North Carolina State University, NC, USA (e-mail: minsu.shin@gmail.com). Hyang-Won Lee and Song Chong are with the Department of Electrical Engineering, KAIST, Daejeon, Korea (e-mail: {mslhw@netsys, song@ee}.kaist.ac.kr).

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

control packet is available. The most interesting feature of our proposed algorithm is that the burst loss ratio does not increase when the overall throughput is enhanced. Moreover, at steady state, the burst loss ratio converges to zero while keeping sufficiently high utilization. On the contrary, the conventional OBS systems should sacrifice overall utilization to maintain acceptable burst loss ratio, or vice versa.

II. OBS Network Model

We consider an OBS network which consists of a set of fiber links, $L = \{1, \cdots, L\}$, where every fiber link $l \in L$ has a set of channels with different wavelengths. We assume that the capacity of each wavelength is equal to C in the network. The network resources are shared by a set of edge-to-edge optical burst flows, $S = \{1, \cdots, S\}$, each of which is assumed to have a predefined single route between its ingress edge and egress edge, denoted by a set $L(s) \subset L$. We assume that no wavelength converter resides in the network due to its cost and complexity. Hence each flow uses the same wavelength in its route. Therefore, each wavelength plane of the network can be modeled separately. For this reason, we consider a single wavelength plane of the network solely without loss of generality. Let $S(l) \subset S$ be the set of flows whose route include link l.

We assume a timer-based burst assembler [1] so that each flow s generates one burst in every τ second interval. The burst assembly interval τ is assumed to be fixed and common for all flows in the network. Let b_s be the variable specifying the allowed burst length (in seconds) and W_s be the actual burst length (in seconds) being injected into the network. Obviously, $0 \le W_s \le b_s \le \tau$, $W_s = b_s$ if there is sufficient data backlog and W_s can be zero if there is no data backlog.

As for the channel reservation, we assume one-way reservation protocol called *Just-Enough-Time* [1]. As shown in Fig. 1, source s having a data burst to transmit first sends a control packet to its egress edge along its route $L(s)$ but using a separate signaling channel. After O_s seconds, the source transmits the data burst, the burst follows the control packet and the control packet reserves channels in advance along the route for the burst. We call O_s as offset time. The control packet carries two values, offset time O_s indicating when the burst will arrive and actual burst length W_s indicating how long the burst will be. The offset time O_s of the flow s consists of fixed base offset time e_s for compensating its control packet's processing time along its lightpath and variable extra offset time d_s for other purpose (e.g. contention resolution in this paper).

Let x_s (bytes/sec) and d_s (in seconds) be the allowed burst assembling rate (allowed edge-to-edge aggregated data transmission rate) and variable offset time of source (edge-to-edge burst flow) s respectively, and τ be the period of burst assembly as shown in Fig. 1. Then, the allowed burst length b_s of s (in seconds) can be expressed as $b_s = \frac{x_s \tau}{C}$ where C (bits/sec) is the single wavelength capacity. Let t_{sl}^{st} and t_{sl}^{end} denote the start time and end time of a burst at link l, respectively. We also assume that the guard time g is used

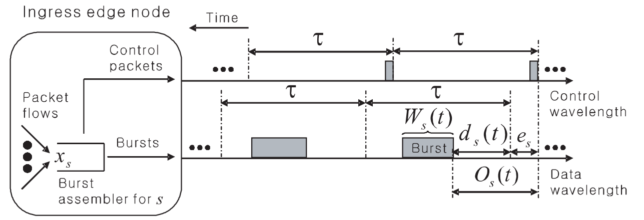


Fig. 1. The time diagram for departure of both control packets and bursts of the flow $s \in S$ in ingress edge node: The burst with its duration $W_s(t)$ (sec) and its control packets are sent every burst assembly time τ (sec).

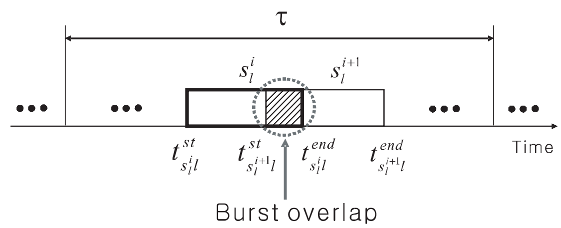


Fig. 2. An example of burst overlap: In a single wavelength in link l, burst overlap occurs when the end time of a burst is later than the start time of its consecutive burst. The burst flow sources denoted by s_l^i and s_l^{i+1} are named by the order of arrival times of their control packets.

between consecutive bursts to compensate the configuration and switching time of core node and to avoid a burst contention. So, if the arrival time of control packet from source s at link l is defined as T_{sl}, then, the two variables can be written as $t_{sl}^{st} = T_{sl} + O_s (= e_s + d_s)$ and $t_{sl}^{end} = t_{sl}^{st} + b_{sl} + g$. For each link l, $S(l)$ is defined to be the set of sources which share the same wavelength of link l and pass through the link. The sources $S(l)$ are ordered according to the arrival times of their control packets. This order is fixed because the arrival times are determined by both the time-invariant propagation delay and the start time of its control packet at the ingress edge. Formally, we write $S(l) = \{s_l^1, s_l^2, \cdots, s_l^{N_l}\}$ where s_l^i denotes the source whose control packet arrives at link l at the i-th earliest time of all the control packets traversing the same wavelength at link l, and N_l is the cardinality of $S(l)$, i.e., $N_l = |S(l)|$. Based on this, the burst overlap can be modelled as shown in Fig. 2 where the burst overlap between burst flows s_l^i and s_l^{i+1} is represented by the period $[t_{s_l^{i+1}l}^{st}, t_{s_l^i l}^{end}]$. To avoid the overlap, we have to set $t_{s_l^i l}^{end} \le t_{s_l^{i+1}l}^{st}$.

III. Problem Formulation

For the greedy burst assembler where b_s is set to τ, the actual burst length W_s can grow up to τ when there is sufficient backlog, which is obviously an unacceptable situation since any burst which shares a link with flow s would collide. Thus, it is necessary to have a mechanism to limit b_s and control d_s such that the resource τ is fairly and efficiently shared by contending flows without burst overlap. To do this, we will propose a joint congestion and burst contention control algorithm by using the following problem (**P**):

$$\max_{b,d} \quad \sum_{s \in S} U_s(b_s) \tag{1}$$

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

subject to $$\sum_{s\in S(l)} b_s \le \tau, \forall l \quad (2)$$

$$t^{end}_{s^i_l l} \le t^{st}_{s^{i+1}_l l}, i = 1, \cdots, N_l, \forall l \quad (3)$$

$$m_s \le b_s \le M_s, d_s \ge 0, \forall s \quad (4)$$

where $b = [b_s, s \in S]$, $d = [d_s, s \in S]$ and $t^{st}_{s^i_l l}$ with $i = N_l+1$ is defined to be τ, i.e., $t^{st}_{s^{N_l+1}_l l} = \tau$. In Eq.(1), each source s is associated with utility function $U_s(b_s)$ of its burst length b_s, which is assumed to be twice continuously differentiable, increasing and strictly concave. Eq.(2) is the link capacity constraint which means that the aggregate bandwidth usage by burst flows must be less than or equal to the capacity of a wavelength at each link at anytime. We formally represent this constraint as follows: for each link $l \in L$, $\sum_{s\in S(l)} x_s \le C$, or equivalently, $\sum_{s\in S(l)} b_s \le \tau$, assuming persistent sources. Eq.(3) is the burst non-overlap constraint which guarantees overlap free at steady state. So, it is easy to see that the solution (b^*, d^*) to the problem (**P**) maximizes the sum of utility functions while ensuring that the sum of the burst lengths at a wavelength of each link does not exceed the period τ and there is no burst overlap (or loss) at steady state. In Eq.(4), m_s and M_s respectively denote the minimum required burst length and the maximum possible burst length of source s.

IV. Proposed Algorithm

The objective function in Eq.(1) is strictly concave in b but not in (b, d). So its dual function is nondifferentiable [13] and we cannot directly apply a dual method to our problem for the development of a distributed algorithm. In order to circumvent this difficulty, we use an augmented Lagrangian method that can handle nondifferentiable dual functions [13]. After converting the inequality constraints in Eq. (3) to equality constraints using the additional quadratic variables $z^2_{il}, i = 1, ..., N_l, \forall l \in L$, the objective function in Eq.(1) is augmented as follows:

$$f(b,d,z) \triangleq \sum_{s\in S} U_s(b_s) - \frac{1}{2}\kappa \sum_l \sum_{i=1}^{N_l} \left(t^{end}_{s^i_l l} - t^{st}_{s^{i+1}_l l} + z^2_{il} \right)^2 \quad (5)$$

where κ is a positive penalty parameter. Then, the augmented Lagrangian function is given by

$$L_A(b,d,z,\lambda,\mu) = f(b,d,z) + \sum_l \lambda_l \left(\tau - \sum_{s\in S(l)} b_s \right) - \sum_l \sum_{i=1}^{N_l} \mu_{il} \left(t^{end}_{s^i_l l} - t^{st}_{s^{i+1}_l l} + z^2_{il} \right) \quad (6)$$

where $\lambda = [\lambda_l, \forall l \in L]$ and $\mu = [\mu_{il}, i = 1, ...N_l, \forall l \in L]$ are Lagrangian multiplier vectors, and $z = [z_{il}, i = 1, ..., N_l, \forall l \in L]$. Let $m = [m_s, \forall s \in S]$ and $M = [M_s, \forall s \in S]$. By the method of multipliers [13], we have the following successive maximization of the form

$$(b(t+1), d(t+1)) = \arg_{(b,d)} \max_{m \le b \le M, d \ge 0, z} L_A(b, d, z, \lambda(t), \mu(t)) \quad (7)$$

followed by updates of the vectors $\lambda(t)$ and $\mu(t)$ according to

$$\lambda_l(t+1) = \left[\lambda_l(t) - \gamma \left(\tau - \sum_{s\in S(l)} b_s(t) \right) \right]^+ \quad (8)$$

$$\mu_{il}(t+1) = \left[\mu_{il}(t) + \kappa \left(t^{end}_{s^i_l l}(t) - t^{st}_{s^{i+1}_l l}(t) \right) \right]^+. \quad (9)$$

The maximization in Eq.(7) at each time t is separable and can be solved by the following iterations:

$$\begin{aligned} b_s(t+1) &= \left[U'^{-1}_s \left(p^b_s(t) \right) \right]^{M_s}_{m_s} \\ p^b_s(t) &= \sum_{l\in L(s)} \left\{ \lambda_l(t) + \mu_{I(s,l)l}(t) \right\}, \end{aligned} \quad (10)$$

and

$$\begin{aligned} d_s(t+1) &= \left[d_s(t) - \kappa p^d_s(t) \right]^+ \\ p^d_s(t) &= \sum_{l\in L(s)} \left\{ \mu_{I(s,l)l}(t) - \mu_{\bar{I}(s,l)l}(t) \right\} \end{aligned} \quad (11)$$

where $[\cdot]^{M_s}_{m_s}$ denotes the projection onto the interval $[m_s, M_s]$, $L(s)$ is the set of links which flow s traverses, $I(s,l)$ indicates the order of source s at a wavelength of link l and $\bar{I}(s,l) = I(s,l) - 1$. So if the control packet's arrival time of s is the i-th earliest at link l, then $I(s,l) = i$ and $\bar{I}(s,l) = i-1$. In Eq.(11), $\mu_{0l}(t)$ is defined to be zero for every t.

A. Algorithm and Its Interpretation

Based on Eqs.(8)-(11), we propose a joint flow and burst offset time control algorithm as follows. Each egress edge node sends a backward control packet (BCP) to carry the prices p^b_s and p^d_s to its corresponding ingress edge node.

Algorithm 1

- *Ingress Edge's Algorithm (for each $s \in S$)*
  1) receive p^d_s and p^b_s through BCP.
  2) adapt the burst offset time of source s by
  $$d_s \leftarrow \left[d_s - \kappa p^d_s \right]^+.$$
  3) adapt its data transmission rate by
  $$b_s \leftarrow \left[U'^{-1}_s \left(p^b_s \right) \right]^{M_s}_{m_s}.$$
  4) generate an FCP containing t^{st}_{sl}, t^{end}_{sl} and b_s, and transmit the FCP every τ seconds.
  5) transmit a data burst $O_s (= e_s + d_s)$ seconds after an FCP was transmitted.
- *Egress Edge's Algorithm (for each $s \in S$)*
  1) receive an FCP.
  2) send the BCP containing p^b_s and p^d_s to its ingress edge node.
- *Link l's Algorithm*
  1) receive b_s, t^{st}_{sl} and t^{end}_{sl} from every source $s \in S(l)$ through FCP.

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

2) update its congestion price by

$$\lambda_l \leftarrow \left[\lambda_l - \gamma\left(\tau - \sum_{s\in S(l)} b_s\right)\right]^+ .$$

3) update its burst overlap prices by

$$\mu_{il} \leftarrow \left[\mu_{il} + \kappa\left(t^{end}_{s^i_l l} - t^{st}_{s^{i+1}_l l}\right)\right]^+ .$$

4) communicate all the prices to corresponding sources and edge nodes.

As well-known, λ_l is interpreted as the price for link utilization. As seen in Eq.(8), if link l is congested, then λ_l will increase, and otherwise, it will decrease. For source s, the sum of these λ_l's on its path is delivered to the source, and used in burst length adaptation Eq.(10). Since the utility function is assumed to be increasing and strictly concave, source s will increase its burst length b_s if the sum decreases, and otherwise, it will decrease b_s. Thus, λ_l is well interpreted as link congestion price.

Similarly, μ_{il} is interpreted as the price for burst overlap. As seen in Eq.(9), if the i-th and $(i+1)$-th bursts overlap, then μ_{il} will increase. Otherwise, it will decrease. The sum of these $\mu_{I(s,l)l}$'s divided by τ on its path is delivered to source s, which adapts its burst length according to Eq.(10). Thus, any overlap event associated with the burst of s will result in the decrease of b_s. On the other hand, if there is no overlap of burst on its path, b_s could increase depending on the sum of λ_l's on its path. Let us see how μ_{il} affects the adaptation of burst offset time d_s. Consider source s and link l, and suppose $I(s,l) = i$. Then, μ_{il}, which is the price associated with the burst from s and the $(i+1)$-th burst arriving at link l, contributes to the increase of p^d_s. This in turn leads to the decrease of offset time d_s, which is reasonable because burst s should be sent earlier to avoid the overlap. In contrast, if burst s overlaps with the $(i-1)$ burst, increased $\mu_{(i-1)l}$ will result in the increase of offset time d_s.

B. Convergence

The convergence of the proposed distributed algorithm to a global optimum, which is not unique, can be readily proved for the case that the algorithm is executed synchronously (refer to Chapter 4.2 in [13] and Chapter 3.4.4 in [14]). However, the convergence of its asynchronous version needs more work and we leave it for future study. Instead, we test its asynchronous convergence through simulations in the following section.

V. Simulation Results

In this section, we present some simulation results verifying the performance of the proposed joint control scheme in various optical network topologies. First we examine the convergence property of the OBS networks when the proposed scheme is used. The results show that after some transient period the system enters steady-state where the amount of burst loss is converged to zero. Next, we compare the performance of the joint control scheme with conventional burst contention resolution schemes. We adopt two performance metrics for the comparison of efficiency. One is burst loss ratio, which is the ratio of the number of lost bursts to the total number of transmitted bursts. The other is goodput, the amount of bursts per second that reached egress edge successfully. Obviously, lower burst loss ratio and higher goodput represent better efficiency in the OBS network. In the comparative simulation, we can demonstrate outstanding efficiency with the proposed scheme.

A. Simulation Environment

The simulation is performed in the NS-2 [15] environment. We implement the proposed algorithm in Section IV at ingress/egress edge and core optical node models. The congestion and burst overlap prices are conveyed by the control packets including FCP and BCP. Since our algorithm can be separately implemented over any wavelength resource, only a single wavelength is used in the simulation.

We set the timer value $\tau = 100\mu$s and burst guard time $g = 1\mu$s at every ingress edge node, and the single wavelength capacity C = 10Gbps. We also set the basic offset time equally as $e_s = 50\mu$s. The additional offset time d_s is initially set to 50μs and can be adjusted within the range of $[0, 100]\mu$s. The minimum and maximum burst lengths are fixed as $m_s = 2\mu$s and $M_s = 50\mu$s, respectively. Each edge-to-edge burst flow $s \in S$ is composed of 40 on-off traffic sources which follow Pareto distribution with shaping factor 1.2. The packet size is fixed to 1000 Bytes. We intend that these aggregate sources emulate the Internet traffic.

For utility function $U_s(b_s), \forall s$, we could use the following [16]

$$U_s(b_s) = \begin{cases} \frac{1}{1-\alpha} b_s^{1-\alpha}, & \text{if } \alpha \geq 0, \alpha \neq 1 \\ \log(b_s), & \text{if } \alpha = 1. \end{cases} \quad (12)$$

This utility function enables to select any compromise between efficiency and fairness. When $\alpha = 0$, the total burst length (or throughput) is maximized, and when $\alpha = 1$, the proportional fair [8] burst length allocation is achieved. As α increases, fairness is improved at the cost of reduced total burst length, and especially as α tends to infinity, the burst length allocation becomes max-min fair. We adopt the logarithm utility function $U_s(b_s) = \log(b_s)$ for each flow s to follow the proportional fairness criteria in OBS networks. However, it should be noted that this allocation can be different from Kelly's proportional fair allocation [8] since the set of feasible rates can be different from that of Kelly's formulation due to the extra non-overlapping constraint Eq.(3).

B. Convergence in Dumbbell Topology

First, we test the convergence of our algorithm in the dumbbell topology depicted in Fig. 3, where 8 edge-to-edge burst flows share a single link L1 and have randomly selected edge-to-edge propagation delays within [33, 36]ms. In this scenario, the equal share of optical link capacity is a proportional fair share. All the burst lengths shown in Fig. 4(a) converge to the same value after transient period. Fig. 4(b) plots the offset time of each burst flow, and we can see that they converge to

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.

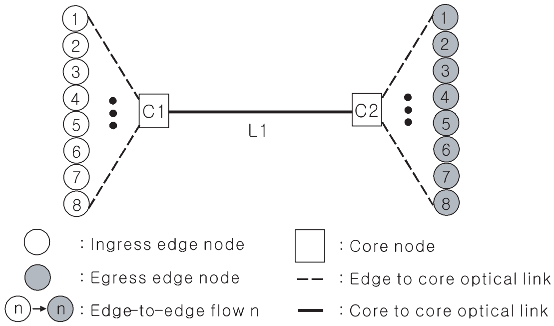


Fig. 3. The dumbbell topology

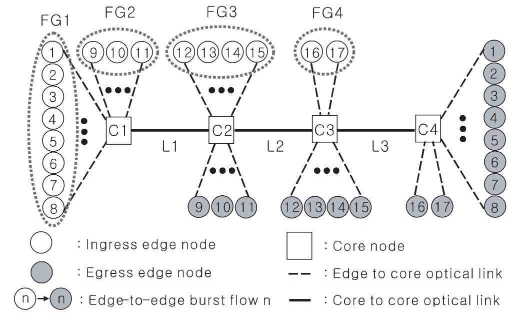
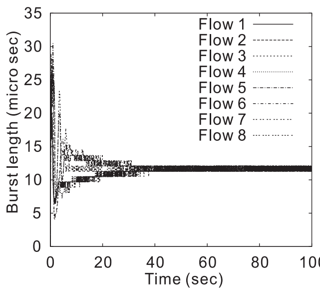
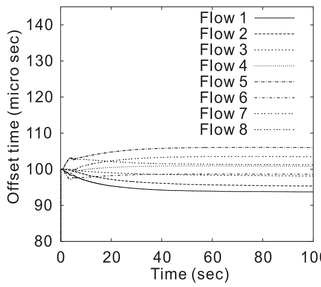


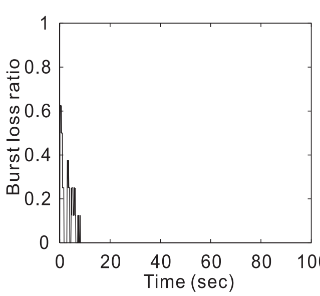
Fig. 5. The parking lot topology for multi-link scenario



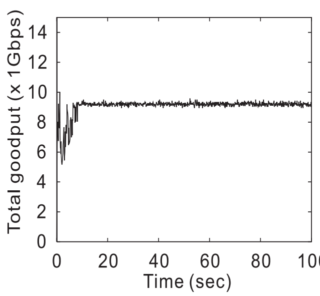
(a) Burst length b_s



(b) Offset time O_s



(c) Burst loss ratio



(d) Total goodput

Fig. 4. Simulation results in single link scenario: convergence to the state of zero loss and maximum goodput.

different values. The more interesting results are depicted in Fig. 4(c) and (d). As seen in Fig. 4(c), the burst loss ratio eventually converges to zero, which implies that at steady state, the burst contention can be perfectly resolved by our algorithm. Furthermore, as shown in Fig. 4(d), the bottleneck link is utilized to the maximum available precluding the inter-burst guard time. This result is not easily attained by using only either a flow control or a contention resolution scheme. Hence, to achieve better efficiency and fairness, it is necessary that the burst length and its placement be jointly controlled like the proposed method.

Remark: We further examined the convergence properties of the proposed algorithm in various scenarios such as i) in the dynamic scenario where the burst flows join or leave the network over time and ii) in the parking lot topology, in Fig. 5, where multiple core links can be bottlenecked. We could observe similar convergence behavior from all scenarios, but we do not show the results here due to limited space. The results can be found in our technical report [17].

C. Performance Comparison

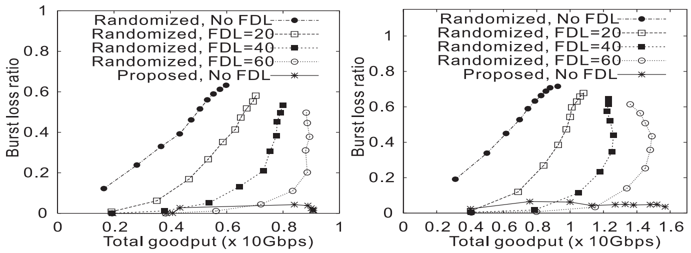
In this section, we demonstrate how much our proposed algorithm enhances the performance (i.e., burst loss ratio and total goodput) compared with those of the conventional contention resolution schemes in various topologies. The conventional scheme generates a burst by aggregating all the packets that have arrived until the burst timer is expired, which is different from our algorithm that assembles the burst according the control information b_s. Below we introduce the general contention-protecting schemes that we use in the simulation.

1) Randomized offset time [3]: To prevent repetitive collision when using timer-based burst assembler, the extra offset time $d(t)$ in Fig. 1 is randomized only when collision is detected at ingress edge node. At every τ, we update $d_s(t+1) = [0, \tau - W_s(t)]^+ \cdot u(0,1)$ if the burst loss is reported by the control packet and $d_s(t+1) = d_s(t)$ otherwise, where $u(0,1)$ is a uniformly distributed random number in [0,1].
2) Fiber Delay Lines (FDL) [2], [5]: FDL provides optical buffering capability to core nodes. The potential contentions among bursts that slightly overlap can be resolved efficiently using FDL. In some cases, we set up per port FDLs with its delay unit $D = \tau/100$, and vary its buffer size N for maximum delay ND. Apparently, if the size N is larger, more contentions can be resolved.

Note that in the comparison, the conventional scheme uses various amount of FDL as a temporal buffer, which is not required in the proposed method.

The simulations are performed at three different topologies such as dumbell network, multiple-link network in Fig. 5, and 14-node NSF mesh network in Fig. 7(a). The network like NSF topology is likely to experience serious congestion problem as the offered load increases. In order to see how the proposed and conventional algorithms react to the congestion, their performances are compared by increasing the aggregate input traffic rate of each edge-to-edge flow. We take 1Q ($\approx$ 0.24Gbps) as the amount of edge-to-edge input traffic increment by each step in the single and multi link topology, and $1\bar{Q}$ ($\approx$0.47 Gbps) in the NSF topology.

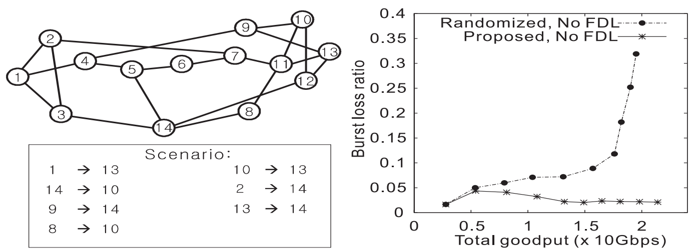
1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.



(a) Single link scenario

(b) Multi-link scenario

Fig. 6. Comparison of the proposed method and conventional methods while increasing offered load: In conventional methods, the goodput performance is increased only at the cost of high burst ratio. Proposed method is keeping very low burst loss ratio even at the highest goodput.



(a) NSF mesh topology

(b) Performance comparison at the NSF mesh topology

Fig. 7. Comparison of the proposed method and the conventional method at the NSF mesh topology. The proposed method outperforms in terms of burst loss ratio and goodput.

Fig. 6 plots the burst loss ratio and goodput performance of each algorithm while increasing the offered load. The performance comparison in the single link scenario is shown in Fig. 6(a), and the multiple link scenario in Fig. 6(b). In conventional methods, we can see that while increasing the offered load, the goodput also increases but at the cost of increased burst loss ratio getting worse. This trend is maintained when adopting FDL, even though the performance is considerably improved by increasing the capacity of FDL. However, the proposed joint control algorithm does not follow this tendency but always marks very low burst loss ratio whatever the offered load is. The reason that the proposed scheme outperforms the conventional one is because it does not send bursts randomly but carefully chooses the burst length and offset time taking into account the link congestion and burst contention. Hence the proposed scheme always achieves fairly low burst loss ratio without costly FDL.

Lastly we conduct the performance comparison in the NSF mesh network topology. The topology and the set of burst flows is depicted in Fig. 7(a), and we use the shortest path algorithm to select the route of each burst flow. Fig. 7(b) shows the burst loss ratio and goodput performance curves in the NSF topology when using the conventional method and the proposed method. We can convince that the network topology does not affect the performance of the proposed algorithm and the conventional scheme experiences serious burst loss problem when offered load is much increased.

VI. Conclusions and Future Work

In this paper, we considered the burst contention problem in the OBS network. The motivation for our work is the observation that the burst contention, which is the main bottleneck of OBS performance, is caused by not only the uncontrolled wavelength resource access trials but also the uncontrolled input traffic. We proposed a joint congestion control and burst contention resolution algorithm that completely eliminates the burst contentions by explicitly controlling the network congestion and burst contention. The proposed algorithm is developed by using network utility maximization problem with optical link capacity constraints and burst contention-free constraints. Through extensive simulations, we showed that our algorithm achieves zero burst loss and high throughput at steady state, and outperforms the conventional burst contention resolution schemes.

References

[1] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: A new area in optical networking research," *IEEE Network.*, vol. 18, pp. 16–23, 2004.

[2] F. Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, pp. 292–294, Sept. 2000.

[3] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: A viable solution for terabit ip backbone," *IEEE Network.*, vol. 14, pp. 48–53, 2000.

[4] F. Farahmand, Q. Zhang, and J. P. Jue, "Feedback-based contention avoidance mechanism for optical burst switching networks," in *Proc. 3rd International Workshop on Optical Burst Switching*, San Jose, California, Oct. 2004.

[5] C. M. Gauger, "Dimensioning of fdl buffers for optical burst switching nodes," in *Proc. Optical Network Design and Modeling (ONDM 2002)*, Torino, Italy, Feb. 2002.

[6] M. Yoo, C. Qiao, and S. Dixit, "Qos performance of optical burst switching in ip-over-wdm networks," *IEEE J. Selec. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.

[7] W. Park, M. Shin, and S. Chong, "Performance enhancement in obs network with flow control and edge delay method," in *Proc. IEEE GLOBECOMM*, San Francisco, California, Nov. 2006.

[8] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. of the Operational Research Society*, vol. 49, pp. 237–252, Apr. 1998.

[9] S. H. Low, "A duality model of tcp and active queue management algorithms," *IEEE/ACM Trans. on Networking*, vol. 11, pp. 525–536, Aug. 2003.

[10] S. H. Low and D. E. Lapsley, "Optimization flow control-I: Basic algorithm and convergence," *IEEE/ACM Trans. on Networking*, vol. 7, no. 6, pp. 861–875, Dec. 1999.

[11] R. J. Gibbens and F. P. Kelly, "Resource pricing and the evolution of congestion control," *Automatica*, vol. 35, pp. 1969–1985, Dec. 1999.

[12] S. Kunniyur and R. Srikant, "End-to-end congestion control schemes: Utility functions, random losses and ecn marks," in *Proc. IEEE INFOCOMM*, vol. 3, Mar. 2000, pp. 1323–1332.

[13] D. P. Bertsekas, *Nonlinear Programming*. Massachusetts: Athena Scientific, 1995.

[14] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, Massachusetts: Athena Scientific, 1997.

[15] "ns-2 network simulator," 2000. [Online]. Available: http://www.isi.edu/nsnam/ns/

[16] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556–567, Oct. 2000.

[17] W. Park, M. Shin, H. Lee, and S. Chong, "Joint congestion control and burst contention resolution in optical burst switching networks," Extended version of this paper, 2007. [Online]. Available: http://netsys.kaist.ac.kr/∼zeus/Tech-rpt-JointOBS.pdf

1930-529X/07/$25.00 © 2007 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings.