Dynamic Traffic Grooming in Optical Burst-Switched Networks

Farid Farahmand‡**, Qiong Zhang**†**, Qingya She**‡**, and Jason P. Jue**†
‡*Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX 75083*
†*Department of Computer Science, The University of Texas at Dallas, Richardson, TX 75083*
{*ffarid, qzhang77, qxs032000, jjue*}@utdallas.edu

Abstract: We address the problem of burst grooming in optical burst-switched networks. We present an edge node architecture for enabling burst grooming and we develop two grooming heuristic algorithms.

OCIS codes: (060.4250) Networks

1 Introduction

An important issue in OBS networks [1] is data burst assembly. Burst assembly is the process of aggregating IP packets with the same destination into a burst at the edge node. A common burst assembly approach is *timer-based,* in which a burst is created and sent into the optical network when a time-out event is triggered [2]. When a burst is created, it needs to be longer than some minimum burst length. The minimum burst length is primarily determined by the core node switching time. If a timer-based assembly mechanism is used and the arrival rate of incoming IP packets with the same destination is such that, upon a time-out, the aggregated IP packets do not meet the minimum burst length requirement, then padding overhead must be added to the transmitted data burst. Excessive padding results in low throughput and higher data burst blocking probability.

One way to reduce padding overhead is to increase the timer duration to allow more packets to arrive. However, if the arrival rate is low, then some packets may experience longer delays. Another approach to minimize the amount of padding overhead, as well as the average end-to-end IP packet delay, is to groom multiple short *sub-bursts* with different destinations into a single burst. A sub-burst consists of IP packets headed to the same destination, whose total length is less than the minimum required data burst duration, L^{MIN}. We refer to the problem of aggregating and routing sub-bursts together as the *data burst grooming problem*. Heuristic algorithms that attempt to solve the data burst grooming problem are referred to as *burst grooming algorithms*. These algorithms differ depending on their aggregation and routing criteria. Furthermore, each algorithm may perform differently depending on the network assumptions and constraints. Clearly, if core nodes could process and switch incoming data bursts as fast as they arrive with no granularity constraint *or* if the edge node could buffer the IP packets until a data burst with minimum switching granularity is created, there will be no need to perform data burst grooming.

In this paper we address the problem of data burst grooming in OBS networks, when the core node switching time is much larger than the average IP packet size. This paper presents an edge node architecture that enables burst grooming and and introduces two data burst grooming heuristic algorithms for dynamically arriving IP packets. Using simulation, we examine the performance of our proposed grooming algorithms under specific network conditions. We compare our results with those obtained without any burst grooming in terms of blocking probability and average end-to-end IP packet delay.

2 Node architecture

Fig. 1(a) shows the basic architecture of an edge node supporting data burst grooming. An ingress edge node, which generates and transmits data bursts to core nodes, performs the following operations: (a) burst assembly: aggregating incoming IP packets with the same destination (or other similar characteristics) in a virtual queue (VQ); (b) sub-burst grooming: combining multiple sub-bursts from different VQs into a single burst.

In the egress path, upon receiving a data burst, the edge node initially disassembles the burst. The extracted sub-bursts which need to be retransmitted are sent to the assembly unit, while the remaining sub-burst is directed to the IP-routing unit. We assume each IP packet can only tolerate a maximum end-to-end delay of T_e in the network. Note that, in this architecture, when an incoming disassembled sub-burst requires immediate retransmission and is routed to the assembly unit, it will be treated as a timed-out sub-burst and must be released immediately.

3 Problem formulation and description of grooming algorithms

In an OBS mesh network, data burst grooming is performed at the edge node. Thus, each individual edge node must decide how to aggregate individual sub-bursts with durations smaller than the minimum length requirement, in order to optimize the throughput and reduce the probability of burst dropping. We formulate the data burst traffic grooming problem as follows. *Given* the entire network information, the minimum required data burst duration (which is a function of the core node switching time), the maximum end-to-end delay that each IP packet can tolerate, and a given timed-out sub-burst with duration smaller than the minimum required length, *find* the available sub-bursts which can be aggregated with the timed-out sub-burst in order to minimize blocking probability and average end-to-end packet

---

[1] This work was supported in part by the National Science Foundation (NSF) under grant ANI-01-33899.

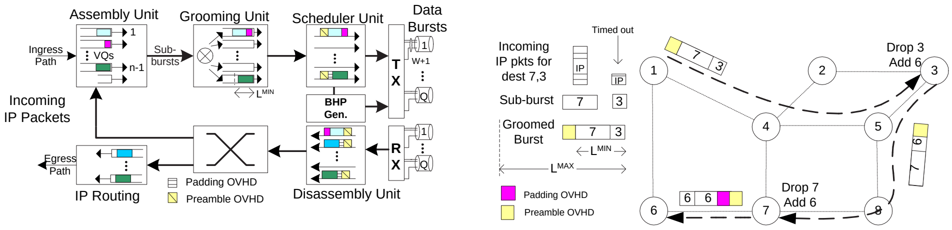

Fig. 1. (a) Edge node architecture; (b) Burst grooming concept.

delay. We assume that a data burst may reach its destination node using multi-hop routing and that data bursts with durations shorter than the minimum burst length requirement will be subject to padding overhead.

We now describe the general grooming concept in OBS networks. We denote a sub-burst as b. Each sub-burst b consists of multiple IP packets with the same destination and can be characterized by its source, destination, and length: S_b, D_b, and L_b. As soon as an IP packet with destination D_b arrives to a node, a timer is set for sub-burst b. The sub-burst will be released when it is timed out. The time-out value for data bursts in each virtual queue is bounded by the difference between the maximum tolerable end-to-end packet delay, T_e, and the sum of source-destination propagation delay and node processing delays, which includes the burst disassembly time at the destination node.

We represent a groomed data burst by $C_k = \{b_0, b_1, b_2, ...\}$, which is constructed by aggregating a number of sub-bursts with different destinations. We consider the first element (sub-burst) in the grooming set (b_0) as the timed-out sub-burst, which must be routed on a single hop. Hence, the first hop for all sub-bursts in C_k will be the node corresponding to the destination D_{b_0}. Depending on the routing decision made by D_{b_0}, other sub-bursts may be routed on a multi-hop path.

We define a hop-delay as the delay time imposed on an incoming sub-burst due to electronic processing. In our study, we only consider the maximum hop-delay, expressed as T_h, and assume it is always the same in all nodes. It is clear that the timed out sub-burst can only be groomed with any other sub-burst, b_i, whose remaining tolerable end-to-end delay, or *remaining slack time*, denoted as δ_{b_i}, satisfies the following expression: $\delta_{b_i} \geq T_p(S_{b_0}, D_{b_0}) + T_p(D_{b_0}, D_{b_i}) + 2 \cdot T_h$. In this expression, $T_p(s,d)$ is the propagation delay from node s to node d. Note that δ_b for any given sub-burst is bounded by T_e.

When C_k reaches its first destination node, D_{b_0}, sub-burst b_0 is dropped. Furthermore, each remaining sub-burst, b_i, in the grooming set C_k, is directed to its proper virtual queue and its slack time is reduced by $T_h + T_p(S_{b_0}, D_{b_0})$. Incoming sub-bursts may be aggregated with the existing IP packets waiting in the corresponding virtual queue. In this case, the remaining slack time of the *combined* sub-burst is set to the remaining slack time of the earliest packet in the queue. These concepts are illustrated in Fig. 1(b).

When a sub-burst b_0 is timed out, the burst grooming algorithm finds the appropriate C_k ($b_0 \in C_k$) among all possible grooming combinations. Selection of the grooming set is based on the optimization objective of the grooming algorithm. In general, aggregating multiple sub-bursts can reduce the *padding* overhead and thus improve the network throughput, which in turn, improves the blocking probability. However, this can potentially result in routing the groomed sub-bursts over longer physical paths. This phenomena, referred as the *routing overhead*, can considerably impact the network throughput.

3.*1* *Grooming algorithms*

The grooming algorithms are distinguished in the way the source node calculates the padding and routing overheads due to burst grooming. Since the source node has no knowledge about the traffic between other node pairs, its padding overhead calculations are based on worst case *local* estimations. We consider two grooming algorithms.

Grooming with no routing overhead (NoRoh): The main objective in this grooming algorithm is to select the grooming set, $C_k = \{b_0, b_1, b_2, ...\}$, such that there is no routing overhead. The relative routing overhead for each sub-burst b_i in the grooming set C_k is calculated as follows:

$$Roh_{b_i} = \frac{H_p(S_{b_0}, D_{b_0}) + H_p(D_{b_0}, D_{b_i})}{H_p(S_{b_0}, D_{b_0}) + H_p(S_{b_0}, D_{b_i})}, \tag{1}$$

where b_0 is the timed-out sub-burst and $H_p(s,d)$ represents the number of physical hops on the shortest path between node pair (s,d). Having $Roh = 1$, indicates that the destination of the timed-out sub-burst, D_{b_0}, is on the shortest path to the destination of the groomed sub-burst, D_{b_i}. The total relative routing overhead for each grooming set C_k will be

$$TRoh_k = \sum_{b_i \in C_k, b_i \neq b_0} Roh_{b_i}. \tag{2}$$

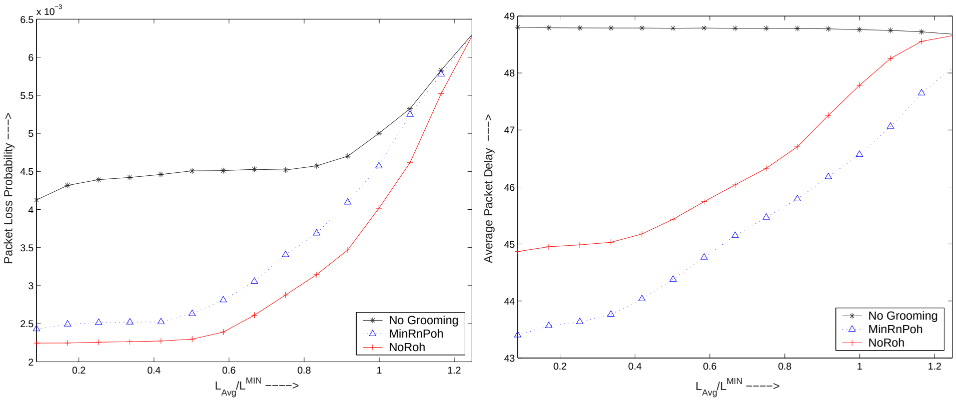

Fig. 2. Plotting the (a) packet blocking probability and (b) average end-to-end packet delay as a function of L^{AVG}/L^{MIN}.

Consequently, the NoRoh algorithm only considers the grooming sets with $TRoh_k = |C_k|$-1.

Grooming with minimum routing and padding overhead (MinRnPoh): The NoRoh algorithm is very *strict* in the sense that it only allows grooming along the shortest paths. This routing constraint can be relaxed by allowing sub-bursts to be groomed if the combined routing and padding overhead is maintained below the case without grooming. We define the relative routing and padding overhead, $RPoh$, for a grooming set, C_k, as follows:

$$RPoh_k = \frac{[max(L^{MIN}, L_k) \cdot H_p(S_{b_0}, D_{b_0}) + \sum_{b_i \in C_k}^{b_i \neq b_0} max(L^{MIN}, L_k - L_{b_i}) \cdot H_p(D_{b_0}, D_{b_i})]}{\sum_{b_i \in C_k} max(L^{MIN}, L_{b_i}) \cdot H_p(S_{b_0}, D_{b_i})}. \quad (3)$$

Hence, the MinRnPoh algorithm selects the grooming set with the lowest $RPoh$ as long as it is not greater than 1. Otherwise, no grooming will be performed and the timed-out sub-burst will be transmitted with padding.

4 Simulation results

We have chosen the NSFNet backbone with 14 nodes as our test network. We assume the end-to-end allowed IP packet delay is 50 ms and the switching time at the core node is 1 ms, requiring a minimum duration of 250 μs (or 250 IP packets) for each data burst. The results presented in this section are based on $|C_k| = 2$, indicating that the timed-out sub-burst can be groomed with only one other sub-burst.

Fig. 2(a) compares the packet loss performance of NoRoh and MinRnPoh to the case with no data burst grooming. In this figure, we plot the blocking probability as a function of average burst length, L^{AVG}, when no grooming is applied. Clearly, as the load increases, L^{AVG} approaches L^{MIN}, in which case grooming tends to become less effective.

Fig. 2(b) presents the corresponding end-to-end average delay when different grooming algorithms are implemented. This figure suggests that the proposed grooming algorithms can significantly reduce the average end-to-end IP packet delay when the network load is low. This is due to the fact that, under the low loading scenario, IP packets are no longer required to wait until they are timed out. Note that, although NoRoh performs better than MinRnPoh in terms of packet blocking, it results in longer average end-to-end packet delay. Relaxing the routing constraint in the MinRnPoh algorithm results in more grooming *opportunities* and hence, packets tend to reach their destinations faster.

5 Conclusion

In this paper we discussed the problem of data burst grooming in optical burst-switched networks. We presented an edge node architecture supporting burst grooming capacity and developed two grooming algorithms in order to aggregate multiple small sub-bursts together. We demonstrated that even limited aggregation of short sub-bursts can improve the packet blocking probability while decreasing the average end-to-end packet delay throughout the OBS network.

References

1. 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, pp.69-84, Jan. 1999.
2. V.M. Vokkarane, K. Haridoss, and J. P. Jue, “Threshold-Based Burst Assembly Policies for QoS Support in Optical Burst-Switched Networks,” *Proceedings, SPIE OptiComm 2002*, Boston, MA, vol.4874, pp. 125-136, July 2002.