Minimizing Preemption Probability to Efficiently Support Service Differentiation in Just-In-Time based OBS Networks

Guido Marchetto
Dipartimento di Automatica e Informatica — Politecnico di Torino
Corso Duca degli Abruzzi, 24 – 10129 Torino (Italy)
guido.marchetto@polito.it

Abstract—Preemptive contention resolution schemes are very effective solutions for providing service differentiation in optical burst switching networks. However, they cannot be applied together with the just-in-time signaling protocol because of the great loss in efficiency in terms of wavelength utilization and maximum achieved throughput that results when the number of preemptions becomes too large. This paper presents a preemption based service differentiation solution that is suitable for the just-in-time optical burst switching paradigm thanks to the fact that it can minimize the preemption probability (i.e., the probability of observing a preemption when a contention occurs). The proposed technique combines a conventional preemption scheme at core nodes and an improvement of the recently proposed burst cluster transmission scheme at edge nodes. In particular, bursts are created at their ingress node and combined into chains, arranging them in order of decreasing priority. Some traffic scenarios are analyzed by simulation to evaluate the performance of the proposed method.

I. INTRODUCTION

The traffic on the Internet continues to grow exponentially. This leads to an increasing demand for bandwidth that Internet Service Providers will have to satisfy by developing new network infrastructures able to support high-rate transmissions. This renewal process has to involve both the edge and the core of the network in order to provide an end-to-end broad-band service to all the network users. Optical Burst Switching (OBS) [1] is an emerging network core technology that can fulfill such requirements. It is therefore very attractive for use in the backbone of the future Internet.

OBS is based on the transport of IP packets over a Wavelength Division Multiplexing (WDM) network. Packets entering the OBS network are assembled in bursts at their ingress node. Different bursts are created for different egress nodes. After an out-of-band control packet has performed resource reservation (i.e., wavelength assignment and switching fabric configuration) at each WDM switch belonging to the burst path, each burst is transmitted through the network. If no resources are available, then the burst is dropped. This allows bursts to be transmitted over bufferless pre-allocated high-capacity channels, with low delays and zero jitters. The time that elapses from the start of the control packet to the instant at which the burst is actually sent through the network is called the offset time.

The Just-In-Time (JIT) [2] protocol is one of the signaling schemes defined for OBS networks. According to this protocol, resources for a particular burst are reserved at a node as soon as the corresponding control packet reaches the node, i.e., resources are reserved but not used during the offset time of a burst. Although other proposed signaling protocols provide more efficient wavelength utilization — e.g., Just-Enough-Time (JET) [1] operates resource reservation just before the burst arrives at the node — the JIT scheme has assumed a role of primary importance thanks to its greater simplicity.

An important open issue in JIT-based OBS networks is the support to service differentiation. This means to ensure that, in case of burst dropping for resource unavailability, higher priority packets are successfully transmitted with higher probability. One of the most effective known methods to perform such a service is the contention resolution preemptive scheme proposed in [3]. In this approach, a priority is assigned to each created burst. Packets are classified according to their service class at each ingress node and assembled in different bursts which assume the corresponding priorities. At each core node, whenever an incoming burst contends for the same wavelength with a previously arrived burst, the contention is resolved according to these two rules: assuming \( n \) and \( m \) are the priorities of the incoming burst and the previously arrived burst respectively, (i) the incoming burst preempts the transmission of the previously arrived if \( n > m \); (ii) the incoming burst is dropped if \( n \leq m \). Since \( k \)-priority bursts contain only \( k \)-priority packets, this contention resolution scheme ensures that higher priority packets are always transmitted in case of contention. However, this scheme can not be effectively used in JIT-based networks because of the low efficiency that results when it is combined with such signaling protocol. Preemptions imply a resource release and reservation overhead which results in a very low achievable average throughput (i.e., a large overall packet loss probability) when the number of preemptions becomes large. A similar technique has been studied in [4] which improves the wavelength utilization; however, it cannot avoid the bandwidth waste due to preemptions.

Other solutions have also been proposed. Burst
segmentation [5] derives from the above presented method and consists in assembling bursts with packets belonging to different service classes, arranging them in order from the highest to the lowest priority. When a contention occurs, the tail of the previously arrived burst is dropped. Since lower priority packets are arranged in the tail of bursts, this method achieves a low loss probability for high priority packets. Another solution [6] consists in prioritizing bursts by assigning them a different offset time. In particular, an extra offset time is given to high priority bursts in order to properly delay them. This allows high priority traffic to have better performance than low priority one. Both solutions are powerful if combined with the JET signaling protocol, but none of them can be applied in JIT-based OBS networks in an efficient manner: in burst segmentation a preemption occurs whenever two bursts contend for resource, while the extra offset time that is applied in [6] results in a further increase of the waste initial time during which resources at a node are reserved but no transmissions are performed. Thus, both further reduce the efficiency of JIT systems as they make the average packet loss probability larger.

The recently proposed burst cluster transmission [7] aims at improving the wavelength utilization by deploying a non-preemptive contention resolution scheme on network nodes. Service differentiation is provided by combing bursts into clusters at the network edge, arranging them from the lowest to the highest priority. This solution is suboptimal regarding both the achieved average throughput and the provided service differentiation because of the non-preemptive scheme applied to resolve contentions. In particular, there are two main issues that affect the burst cluster transmission effectiveness: (i) higher priority packets can be dropped even if they contend with lower priority ones and (ii) bursts can also be dropped when contending with burst clusters that will be dropped at other nodes.

This paper presents a service differentiation technique — referred to as High-Priority First (HPF) transmission of bursts — which improves burst cluster transmission. In fact, it provides high efficiency in service differentiation aware JIT-based OBS networks while maintaining the effectiveness of the previously described contention resolution scheme [3]. The proposed solution is based on the burst assembly algorithm proposed for the burst cluster transmission technique: bursts, assembled using a priority based classification of packets, are combined into chains for transmission. But here, unlike the burst cluster transmission scheme, the bursts are arranged in order from the highest to the lowest priority and nodes operate according to a preemption based contention resolution scheme. This particular burst transmission scheduling can minimize the preemption probability in case of contention, achieving the above mentioned properties.

The rest of the paper is organized as follows. Section II focuses on the JIT protocol by discussing its operating principles and its problems when used with a preemptive contention resolution scheme. Section III describes the HPF transmission technique. Section IV provides some simulation results, and Section V concludes the paper.

II. JUST-IN-TIME PROTOCOL OVERVIEW

In the JIT protocol, the ingress node of a burst sends a control packet (the SETUP message) before burst transmission. Such packet performs resource reservation at each node belonging to the burst path. As soon as the control packet arrives at a node, wavelength reservation and switch configuration are performed, and then the packet is forwarded to the next node. Since burst transmission needs to happen only when resources have been configured along the entire path, an initial transmission delay is necessary at the ingress node. Defining $t_p$ as the control packet processing time at each node, and $N$ as the number of hops from the ingress node to the egress node of a burst, the initial transmission delay $T_0$ can be evaluated by $(N + 1)t_p$. Typically, $t_p = 1\text{ ms}$. Therefore, at each node there is an idle time during which resources are reserved but no transmissions are performed. The resulting bandwidth waste is however generally small if compared to the average burst size.

Regarding the resource release, two methods have been proposed: (i) the explicit release method, which consists in sending a control packet (the RELEASE message) at the end of the burst transmission, and (ii) the estimated release method, where no more control packets are sent after the burst transmission and resource release is performed using burst size information that the ingress node has to put in the SETUP message. The operation of the JIT protocol in case of both explicit release and estimated release is shown in Fig. 1.

![JIT protocol operating principles](image)

When a preemptive contention resolution method is used to provide service differentiation, the SETUP message also contains the service class of the burst. When a SETUP message referred to a 1-priority burst arrives at the generic core node and there are no available resources, a contention occurs. If $n$ is higher than the priority $m$ of a previously scheduled burst, the SETUP message preempts the transmission of such burst. However, the entire offset time has
to elapse before the new burst arrives at the node. Thus, at each node there is an idle time during which resources are reserved but not utilized whenever a preemption occurs. This results in a swift rise of the wasted bandwidth when the number of preemptions becomes large. Furthermore, a resource release procedure based on explicit RELEASE messages has to be started by the node at which the preemption occurs. Depending on which algorithm is applied (one-way or two-way preemption occurs. Depending on which algorithm is applied (one-way or two-way [8]), one or two RELEASE messages are generated, with different efficiency. However, in both cases, all nodes belonging to the path of the preempted burst are reached by a RELEASE message which allows them to release the resources that have been reserved for such burst. As propagation delays are not zero, several milliseconds can elapse before these resources can be reused, with consequent further bandwidth waste. The following section discusses the HPF transmission of bursts, which can reduce bandwidth waste by minimizing the preemption probability.

III. HPF TRANSMISSION

A. Operating Principles

In a preemption based service differentiation solution, every burst that arrives at a node can potentially cause a preemption if its priority is higher than the actual lowest priority. In fact, in case of contention, an incoming burst always preempts the transmission of a lower priority one. HPF transmission of bursts aims at scheduling burst transmissions in such a way that the probability for an incoming burst to contend with a lower priority one is minimized. Thus, it aims at minimizing the preemption probability when a preemptive contention resolution method is used.

HPF transmission operates at the network edge and consists in transmitting bursts consecutively, in order of decreasing priority, so that they appear as composed into chains. Two modules are necessary at the ingress node to perform HPF transmission: a per-egress node burst assembler and a per-output port burst transmission scheduler. Furthermore, a per-egress node oriented mixed time-length based burst assembly algorithm is used to generate bursts.

As previously seen, the burst assembler operates according to the algorithm presented in [7] for burst cluster transmission. The module is composed by $M$ buffers (let $M$ be the number of supported service classes) where incoming packets are stored according to their service class. Bursts are therefore assembled with packets of the same service class and a priority $k$ is assigned to bursts composed by $k$-priority packets. Here, the per-destination burst assembly algorithm performs burst generation. Its operating principles are the same as the well-known Max-Time-Min-Max-Length based algorithm [9] developed for OBS networks, but now it is deployed in a per-egress node fashion. Max-time $T_{\text{max}}$ and max-length $S_{\text{max}}$ thresholds are related to the entire group of $M$ buffers, and are defined so that $M$ bursts are generated and passed to the output queues for transmission whenever:

1. $T = T_{\text{max}}$, where $T$ is a per-egress node timer restarted when a packet arrives at a node and assembling buffers related to its egress node are empty, or

2. $\sum_{k=0}^{M-1} S_k = S_{\text{max}}$, where $S_k$ is the size of the $k$-priority burst.

The max-time threshold $T_{\text{max}}$ is defined according to the maximum tolerable assembling delay. The max-length threshold $S_{\text{max}}$ here is fixed to $M \cdot S_{\text{max}}$, where $S_{\text{max}}$ is the maximum length threshold that would be used for a single burst in a conventional OBS network.

![Fig. 2. HPF transmission operating principles: (a) packet arrivals and buffering, (b) burst assembly, and (c) burst transmission](image)

The min-length threshold is instead referred to each single burst; if the size of a burst is under this threshold when $T = T_{\text{max}}$ or $\sum_{k=0}^{M-1} S_k = S_{\text{max}}$, the data size of the burst is increased to min-length with padding.

Once the $M$ bursts have been created, they have to be transmitted consecutively, as previously depicted. In order for the bursts to be sent out in such fashion, the burst transmission scheduler serves the output queues in a round-robin order, from the highest priority queue to the lowest priority one. This is the key point of the HPF transmission
scheme which leads to the minimization of the preemption probability. As in the burst cluster scenario, control packets could be transmitted so that the wavelength is left idle only at the end of the entire burst chain. Thus, SETUP messages are sent in such a way bursts within a chain are transmitted consecutively. Furthermore, if explicit release is used, a RELEASE message is sent after the chain, while, in the case of estimated release, the RELEASE message is not sent and resources are released when the chain transmission finishes. Fig. 2 shows how the HPF transmission method operates in an edge node of an OBS network. Packet arrivals and buffering, burst assembly, and burst transmission in case of estimated release are presented.

No modifications are required at core nodes, which operate according to the traditional preemptive contention resolution scheme [3].

B. Features and benefits

HPF transmission improves the efficiency of [3] in JIT-based OBS networks by minimizing the preemption probability. In case of contention between two bursts, an incoming higher priority burst preempts the transmission of a lower priority one. The chain which the preempted burst belongs to is called the preempted chain. The other chain is the preempting chain. As the bursts within each chain are transmitted consecutively, the preempting chain maintains the control of the wavelength until the last burst is sent. Furthermore, since bursts are transmitted in order of decreasing priority, the bursts that follow the preempted (preempting) one in the preempted (preempting) chain have a lower priority than it. Thus, assuming $M$ bursts are transmitted within each chain (i.e., $S_i \neq 0 \forall k : 0 < k \leq M - 1$) and $k$-priority bursts have comparable lengths in different burst chains:

1. no contentions occur when remaining bursts of the preempting chain arrive at the node and
2. contentions without preemptions are observed when remaining bursts of the preempted chain arrive at the node. In fact, they contend with higher priority bursts that belong to the preempting chain.

Preemption probability is therefore minimized, reducing the overhead in handling preemptions that drastically degrades the efficiency of a preemptive contention resolution scheme in JIT-based OBS networks.

Fig. 3 shows an example of contention resolution in a congested node when HPF transmission is used at the network edge. For the sake of simplicity, control packets are omitted and only one wavelength per port is shown. A service differentiation scheme based on $M=8$ service classes (priority 7 is the highest) is used. Burst chain A arrives first at the node and acquires the output wavelength 3. The first burst of chain B arrives during the 5-priority burst transmission. Since its priority is higher than 5, it preempts the transmission of the chain A: 5-priority burst tail is dropped and burst chain B assumes the control of the wavelength. The following bursts of chain A contend with higher priority bursts of chain B and are therefore entirely dropped. Note that even if chain A is cut during the contention resolution procedure, it is not distorted (i.e., decreasing priority order is maintained for bursts). Thus, subsequent core nodes can also benefit from HPF transmission of bursts.

Both effectiveness and wavelength utilization are higher than the ones obtained with burst cluster transmission. In burst cluster transmission, incoming bursts are always dropped if there are no available resources. Since resources at a node could be reserved by bursts that will be subsequently dropped along their path, it is possible that incoming bursts are blocked in favor of bursts that never reach their destination. This results in a bandwidth waste that could rapidly increase at high traffic rate. We refer to this phenomenon as bandwidth waste blocking. HPF transmission can limit the effects of bandwidth waste blocking thanks to both the arrangement of bursts into the chains and the preemptive contention resolution scheme adopted by network nodes: only the low priority tails of the burst chains could be blocked with consequent bandwidth waste. Furthermore, HPF transmission also avoids drops of high priority bursts that contend with lower priority ones, which are inevitable in a non-preemptive scheme such as burst cluster transmission.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed method, simulations are implemented and conducted over the 14-node NSF network topology showed in Fig. 4.

Packet arrivals are modeled by a Poisson process with average arrival rate $\lambda$. Furthermore, a given discrete probability distribution function models the belonging of packets to the defined service classes. In particular, an incoming packet has priority $k$ with probability $p_k$. Ingress and
egress nodes of an incoming packet are uniformly distributed among the 14 nodes composing the network topology. For simplicity, a fixed packet length of 1 KB is used. In addition, the number of wavelengths per link is \( W = 8 \), the capacity of each wavelength is \( C = 10 \text{ Gb/s} \), and the control packet processing time \( t_p \) is set to 1 ms. For the burst assembly algorithm parameters, max-time threshold \( T_{\text{max}} \) is set to 5 ms, while the per-burst max-length threshold \( S_{\text{max}} \) is assumed to be equal to 7 MB.

Network nodes are supposed to be wavelength conversion capable. For the sake of simplicity, a wavelength for preemption is randomly selected among all the wavelengths used by lower priority bursts [8]. Furthermore, the estimated release method is implemented.

Fig. 5 compares the conventional preemptive contention resolution scheme and the HPF transmission method with regards to the achieved packet loss probability when \( M=4 \) service classes are used. The packet priority probability distribution function is \( p_3=0.1, p_2=0.2, p_1=0.3, p_0=0.4 \). Priority 3 is the highest. It can be observed that both systems can effectively serve highest priority packets (packet loss probability achieved to 3-priority packets is low and comparable for both cases). However, if HPF transmission is not used, the great number of preemptions that occur at core nodes causes a gradual loss in efficiency that results in an unacceptable quality of service provided to lower priority packets. In fact, a high overall loss probability can be observed in Fig. 5 for lower priority packets, which results in an undifferentiated service offered to them at high packet arrival rate. HPF transmission of bursts is instead both efficient and effective in supporting service differentiation: \{2,1,0\}-priority packets receive an adequate differentiated service in term of burst loss probability, and the overall efficiency is sensibly higher than that of the traditional preemptive scheme without HPF transmission.

Fig. 6 shows how HPF transmission is also superior to burst cluster transmission in providing service differentiation. It is noticeable that the dotted lines, related to burst cluster transmission, are closer to each other and higher than the lines that represent HPF transmission. This results from the preemptive nature of the adopted contention resolution scheme, which enables HPF transmission to (i) guarantee high priority packets to be transmitted when contending with lower priority ones, and (ii) reduce bandwidth waste blocking.

Fig. 7 compares the overall achieved throughput in the three considered cases. It is noticeable how the efficiency is higher if HPF transmission is used. Fig. 7 also shows the effect of the bandwidth waste blocking, which causes the efficiency of burst cluster transmission to be also lower than that of the conventional preemptive scheme at high bit rate.

The performance of HPF transmission is evaluated also with a higher percentage of high priority traffic. In particular, the case of \( M=4 \) uniformly distributed service classes is considered. Fig. 8 shows how HPF transmission can efficiently provide service differentiation also with such traffic scenario.
V. CONCLUSIONS

This paper proposes a service differentiation solution for just-in-time based optical burst switching networks. In particular, we describe the High-Priority Furst (HPF) transmission scheme, which can improve the efficiency of preemption based just-in-time networks by minimizing the preemption probability. The proposed method also outperforms burst cluster transmission, a non-preemptive service differentiation solution from which the adopted burst assembly algorithm is derived. Simulation results confirm these properties and also show the effectiveness of the proposed method with a higher intensity of high priority traffic.

REFERENCES