An Efficient Hybrid Burst Retransmission and Burst Cloning Scheme Over Star OBS Networks

Salek Riadi, Abdelilah Maach
Computer Science Department
Mohammadia School of Engineers
Mohammed V-Agdal University
Rabat, Morocco
salek.riadi@gmail.com, maach@emi.ac.ma

ABSTRACT: *Optical Burst Switching (OBS) has become a mature technology to support the current and next generation Internet over optical network backbones. In OBS networks contention is the main source of burst loss that can reduce the performance of the higher layers. Burst retransmission is a reactive loss recovery mechanism, which attempts to resolve contention by retransmitting the contended burst at the OBS layer. Burst cloning is a proactive loss recovery mechanism, which attempts to prevent burst loss by sending two copies of the same burst; if the first copy is lost, the second copy may still be able to reach the destination. Burst retransmission is better suited when the load is low, however burst cloning is better suited when the load is high. In order to combines the advantages of burst retransmission at low load with the benefits of burst cloning at high load, we propose a hybrid scheme for star OBS networks that aims to control the extra load due to the both loss recovery mechanisms. Simulation results show that our hybrid scheme can achieve better overall network performance than both burst retransmission scheme and burst cloning scheme.*

Keywords: Optical Star Network, Optical Burst Switching, Burst Retransmission, Burst Cloning, Hybrid Scheme

Received: 28 May 2012, Revised 27 July 2012, Accepted 29 July 2012

1. Introduction

Optical Burst Switching (OBS) has become a mature technology to support the current and next generation Internet over bufferless Wavelength Division Multiplexing (WDM) networks. It can be considered as intermediate solution between Optical Circuit Switching (OCS), which has a low complexity but it suffers from low bandwidth utilization, and Optical Packet Switching (OPS), which has advantage of high bandwidth utilization but at the cost of high complexity [1], [2].

In OBS networks, a data burst, which consists of multiple packets, is created at ingress node and switched by one or more core nodes along the network all-optically until it reaches its destination egress node. Before the start of the burst transmission, the ingress node sends a control packet to reserve a wavelength for the burst at each core node, where the control packet is subject to Optical-Electric-Optical (OEO) conversions. If the wavelength reservation fails due to the contention with another burst at a core node, which lacks of wavelength conversion capability and optical memory, then the burst is lost.

Loss recovery represents one of the major challenges that face researchers. Loss recovery mechanisms have been classified into two categories, called reactive and proactive; reactive mechanisms are better suited when burst loss is rare and bandwidth utilization needs to be optimized, however proactive mechanisms are better suited when contention loss is high and delay needs to be optimized [3].

Burst retransmission is a reactive loss recovery mechanism that attempts to resolve contention by retransmitting the contended burst at the OBS layer [4]. The most studies on burst retransmission focused on mesh topologies. Whereas optical star networks have received considerable attention from researchers and industrials, only three studies, to our knowledge, have investigated the burst retransmission for OBS over star networks. In the first study [5], the authors proposed a combination of bursts retransmission and contention reduction through congestion control. In the second study [6], the authors defined two retransmissions schemes; in both schemes, when a core node processes a control packet and cannot reserve the necessary resources, it reschedules the burst transmission and sends a special control packet, called Core Reserve Packet, to inform the edge node for a suitable time to transmit the burst. In the third study [7], a controlled retransmission scheme has been proposed to control the extra load, due to the retransmissions, to a required value.

Burst cloning is a proactive loss recovery mechanism; the idea is to replicate a burst and send duplicated copies of the burst through the network simultaneously; if the original burst is lost, the cloned burst may still be able to reach the destination [8]. To our knowledge, this paper is the first study on burst cloning focus on star topologies.

In this paper, we propose a hybrid scheme for star OBS networks that combines the advantages of burst retransmission at low load with the benefits of burst cloning at high load through a decision algorithm that controls the extra load due to the both loss recovery mechanisms. According to the current state of the arrival of bursts and the retransmission buffer, the decision algorithm decides for each burst whether or not it will be processed by one of the both loss recovery mechanisms and, if so, which of the both mechanisms will be used.

The remainder of this paper is structured as follows: in section II, we describe the OBS network under study; in section III, we present the burst retransmission scheme; in section IV, we present the burst cloning scheme; in section V, we show the hybrid scheme; in section VI, we evaluate the performance of our hybrid scheme through simulations; finally, our conclusion is provided in section VII.

2. The OBS Network Under Study

In this paper, we focus on a class of OBS networks that use an overlaid-star topology (also called composite-star topology) [9], [10], [11], [12], [13]. The overlaid-star topology, as shown in Figure 1, forms a logical mesh, where each edge node is a member of two or more stars in order to have at least one recovery path in the event of core or fiber link failure. A burst traversing the network only passes through one core node, resulting in a major simplification of the control problem of ensuring that contention is rare; furthermore, each of the stars can be managed independently of the others [13].

We consider a star OBS network topology, as shown in Figure 2, where each edge node is connected to a core node using two fibers, one in each direction. All the fibers have the same number of transmission wavelengths, W of them are used for burst transmission, called data wavelengths, while the remaining wavelengths are used to transmit control packets, called control wavelengths.

The core node includes an electronic control unit and an Optical cross Connect (OxC) that can optically switch an arriving burst on a data wavelength w of an input port i to the same data wavelength w of any output port j, in other words, the core node has not wavelength conversion capability. Also, there are no fiber delay line buffers available at the core node and thus the burst will be dropped if the data wavelength w of the output port j is busy.

Each edge node functions as both the ingress and egress node. It has a burst assembler, which assembles multiple data packets from clients into a burst for transmission to other edge node. When the burst is ready, the edge node assigns a data wavelength for the duration of the burst. When the destination edge node receives the burst through its burst-mode receiver, it disassembles the burst into multiple data packets and then sends them to appropriate clients.

We assume that each edge node sends the bursts from W sources (i.e. number of data wavelengths) to the other edge nodes with same probability. We assume also that all sources in the edge nodes are modeled with an identical IDLE-ON process [14]. In

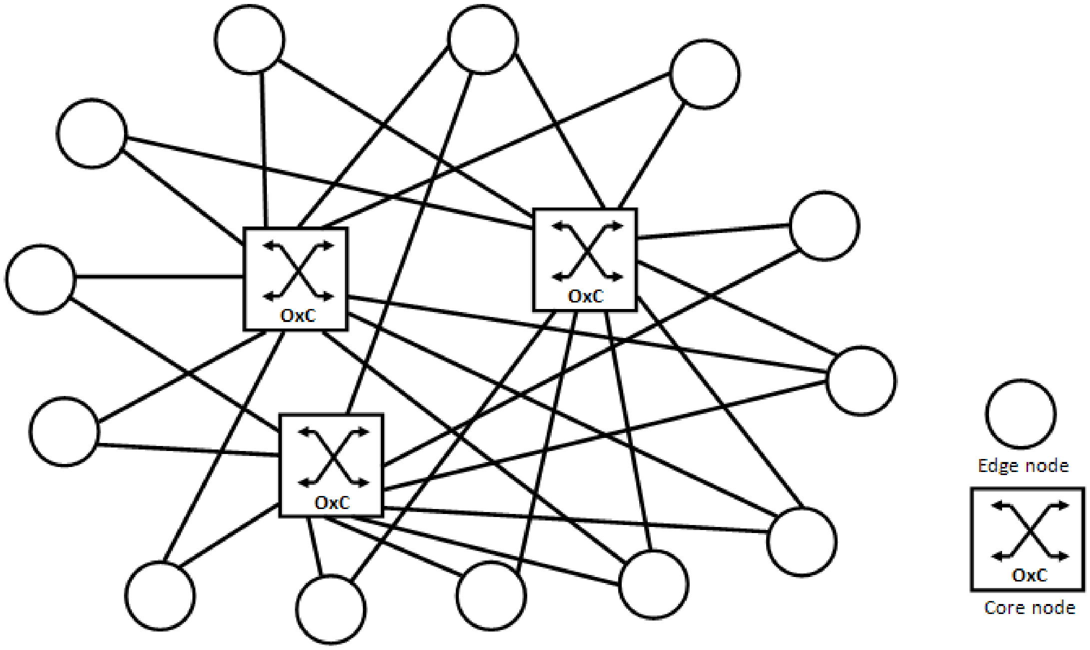

Figure 1. Overlaid-star topology

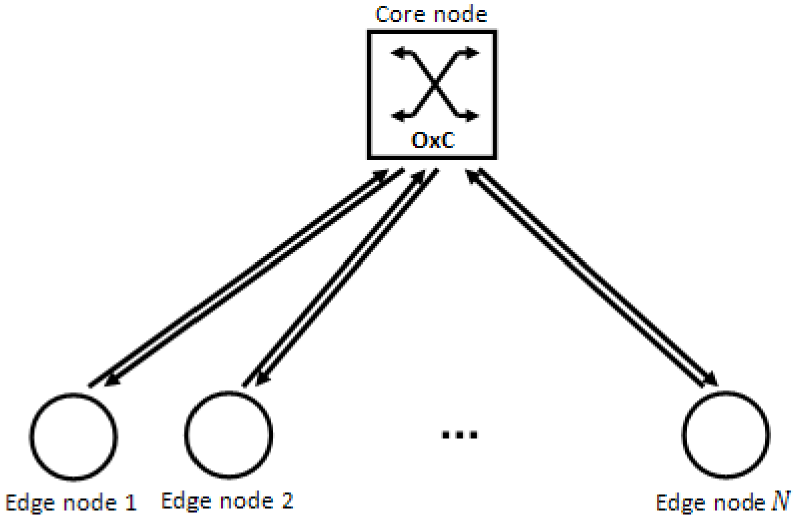

Figure 2. Star OBS Network

IDLE-ON process, the source sends one burst in the ON state and then it moves to the IDLE state. The source stays in ON state for the duration of the burst, which is exponentially distributed with a mean of $1/\mu$. The IDLE state is also exponentially distributed with a mean of $1/\alpha$. The IDLE-ON process captures the burst transmission more accurately than the Poisson process because it considers that burst transmission takes time and it is not instantaneous [14].

3. Burst Retransmission Scheme

Burst Retransmission Scheme (BRS) for OBS over star networks has been proposed in [5] and studied in [6]. As shown in Figure 3, in the BRS, each edge node stores a copy of the transmitted burst for possible retransmissions in an additional queue, called the retransmission buffer, if it has sufficient buffer space. A timer controls the lifetime of each copy; it is set to the Round Trip Time (*RTT*) between the edge and the core node. Each stored burst should be assigned a unique burst id (i.e. sequence number) in order to uniquely identify the contending burst, which needs to be retransmitted. If the channel reservation fails due to burst contention and if the contending burst is saved in retransmission buffer, the core node sends a negative acknowledgement packet (NACK) with burst id to the concerned edge node in order to report the reservation failure. When the edge node receives the NACK packet, it retransmits a duplicate of the contending burst. The timer is set again to its initial value. The edge node still keeps the copy until the burst reaches its destination.

For low load the BRS can improve significantly the throughput with reasonable delay; however, for high load, the BRS leads to overload of the network. In the overloaded network the burst contention probability is significantly higher and the retransmitted bursts probably will be lost again, therefore the retransmission of contended bursts is not recommended in such conditions.

4. Burst Cloning Scheme

In Burst Cloning Scheme (BCS), two copies of the same burst sent simultaneously through the network; if the first copy is lost, the second copy may still be able to reach the destination [8]. We adopt the BCS with star topology. As shown in Figure 4, the edge node in star network is responsible to do cloning. In order to avoid edge congestion, the edge node should not do cloning of a burst that has queuing delay (i.e. waiting time is not null). When an edge node has a new burst that has not queuing delay, the edge node sends a special control packet for two copies of the same burst with two offset times and two data channels. The first copy sent after the first offset time and the second copy sent after the second offset time. When the core node receives the special control packet, it tries to reserve a channel for the first copy and if the channel reservation fails, then it tries for the second copy. One copy can leave the core node to its destination edge node.

The BCS can improve the throughput but not significantly; however, it keeps very low delay even at high load. The burst cloning mechanism has been suggested, in [8], as a complementary and not as an alternative mechanism to existing reactive loss recovery mechanisms.

5. Hybrid Scheme

We propose a Hybrid Scheme (HS) for star OBS networks that combines the advantages of burst retransmission mechanism at low load with the benefits of burst cloning mechanism at high load. According to the current state of the arrival of bursts and the retransmission buffer, the edge node in the HS must decides, for each burst, whether or not it will be processed by one of the both loss recovery mechanisms and, if so, which of the both mechanisms will be used. To make the decision, we propose a decision algorithm that controls the extra load, due to the both loss recovery mechanisms. Since all sources in the edge nodes are modeled with an identical IDLE-ON process, the required extra load (E) should be:

$$E = \frac{\mu}{\alpha + \mu} \tag{1}$$

Therefore, the mean of duration (R) of the burst, which will be retransmitted or cloned, should be less than 1D *á*. For this reason the main goal of our algorithm is to maintain the following constraint:

$$\sum_{i=1}^{k} (n_i \,.\, on\_time_i) \le \sum_{i=1}^{k} idle\_time_i \tag{2}$$

Where k is the number of bursts that are transmitted; $idle\_time_i$ is the idle time of a source before it begin to send the burst i; on_time_i is the duration of burst i; n_i ($n_i = 0, 1, 2, \ldots$) is the number of times when the burst i is retransmitted or cloned. We let G be:

$$G = \sum_{i=1}^{k} idle\_time_i - \sum_{i=1}^{k} (n_i \,.\, on\_time_i) \tag{3}$$

G is initialized to 0, and the constraint (2) will be as follows:

$$G \ge 0 \tag{4}$$

As shown in Figure 5, when an edge node has a new burst i to transmit, the edge node recalculates G:

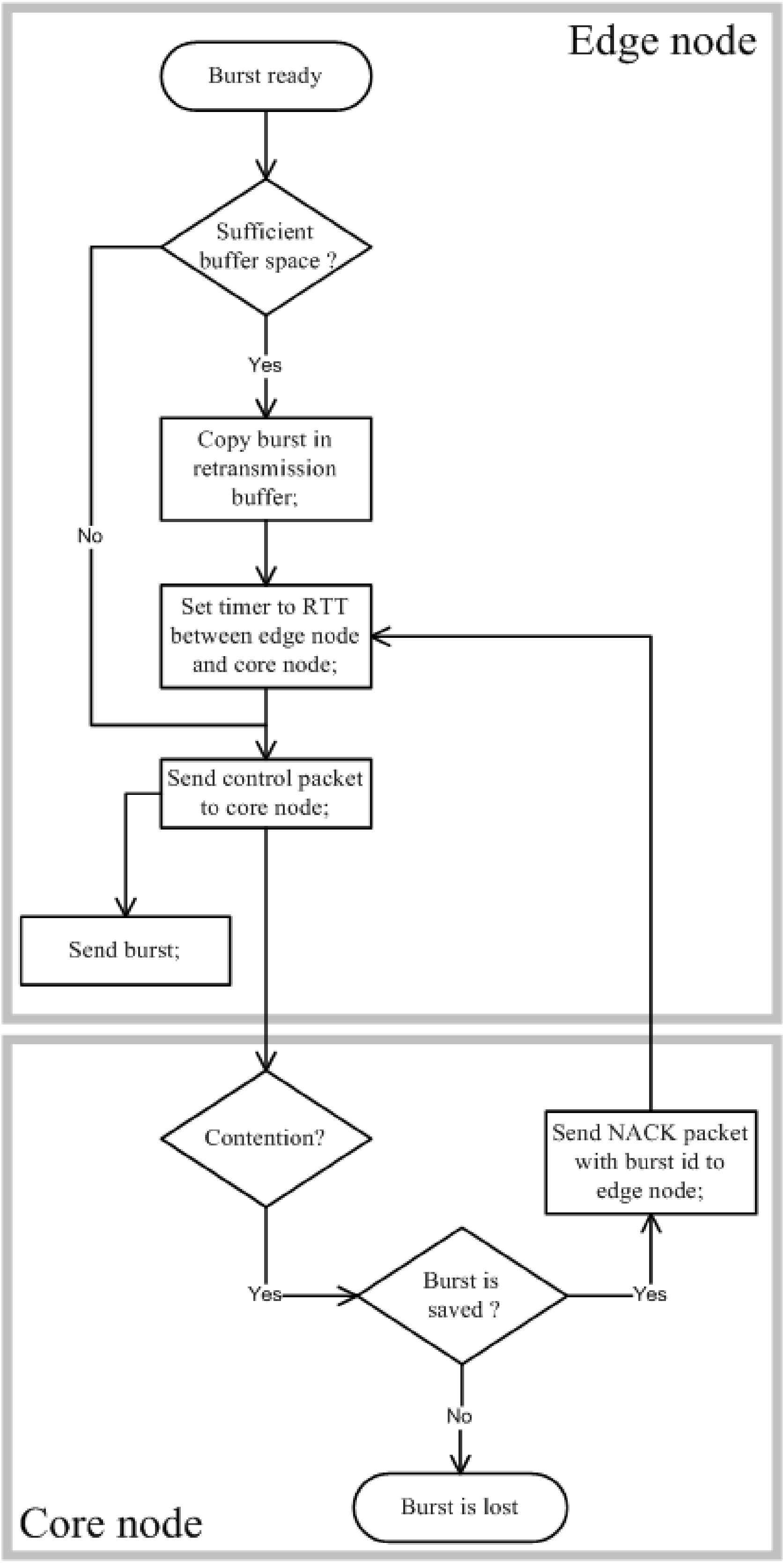

Figure 3. Burst Retransmission Scheme

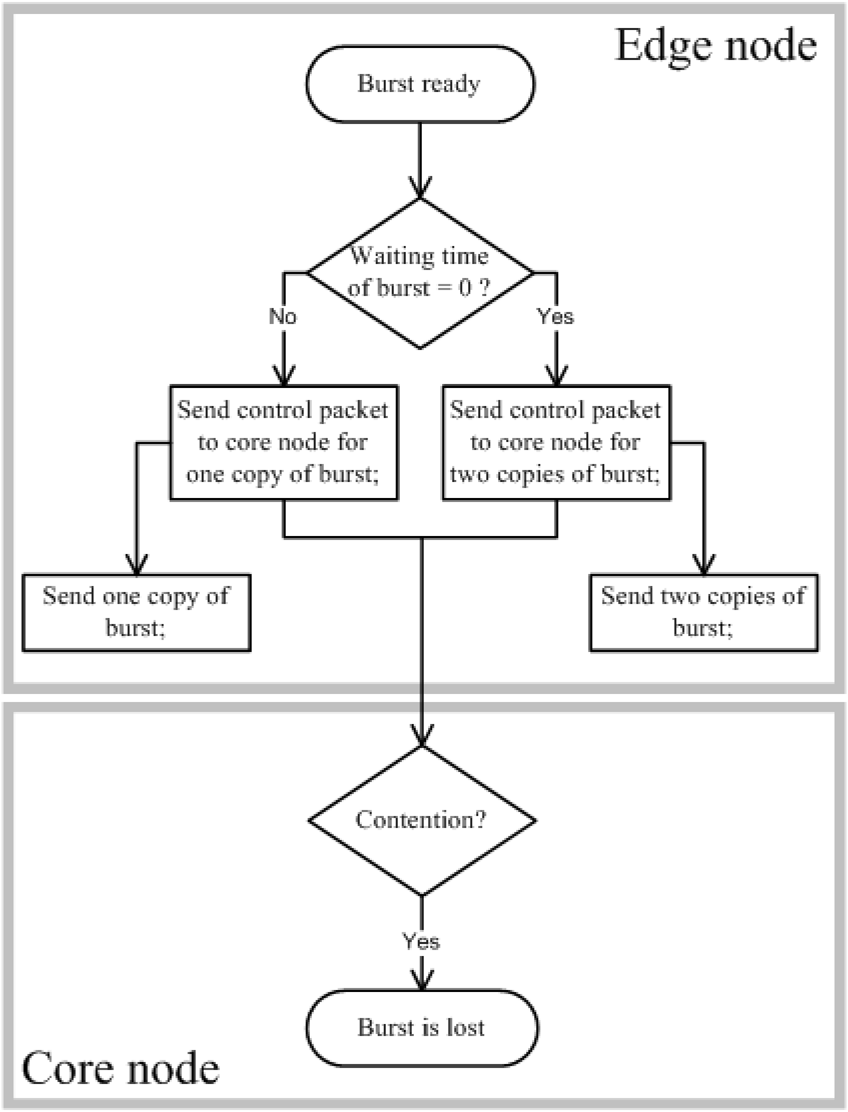

Figure 4. Burst Cloning Scheme

$$G = G + idle\_time_i$$

And then decides, for the burst *i*, whether or not will be processed by a loss recovery mechanism. If the on_time_i of the burst *i* is greater than *G*, then the edge node disables the recovery loss of the burst. Else, the edge node decides either to keep the burst for possible retransmission or send two copies of the same burst:

- If the edge node has sufficient buffer space, then the edge node keeps the burst for possible retransmission.
- Else, the edge node send two copies of the same burst, and recalculates *G*:

$$G = G - on\_time_i$$

When the channel reservation fails due to burst contention, the core node sends a negative acknowledgement packet (NACK) only if the burst *i* is saved in retransmission buffer, else, the core node disables the retransmission of the contending burst.

When the edge node receives the NACK packet, it recalculates *G*:

$$G = G - on\_time_i$$

And then decides whether or not to keep the burst i for possible retransmission for the burst. If the on_time_i of the burst i is greater than G, then the edge node deletes the burst from retransmission buffer. Else, the edge node keeps again the burst for possible retransmission. The edge node still keeps the copy of the burst until the burst reaches its destination or until the on_time_i becomes greater than G.

When an edge node is in overloaded state its G should be much lower than the on_time_i. In this situation, the both loss recovery mechanisms should be disabled. However, the loss recovery should be enabled, when G is higher than the on_time_i. In this case, the edge node enables burst cloning mechanism only if the retransmission buffer space is insufficient, that means the edge node is near to overloaded state, else it enables burst retransmission mechanism. Consequently, this way can combine the advantages of burst retransmission mechanism at low load with the benefits of burst cloning mechanism at high load.

6. Performance Evaluation

In order to compare the performance of our HS with BRS and BCS, we implemented these schemes over Optical Burst Switching - network simulator (OBS-ns) that developed at the Optical Internet Research Center (OIRC) [15] on the basis of ns-2 [16].

6.1 Performance Metrics

The performance metrics of interest here are normalized throughput and normalized delay, that are more convenient to evaluate the performance of the HS with variable burst size. The normalized throughput defined as the absolute throughput divided by the total available bandwidth. The normalized delay defined as the sum of the size of burst i ($size\_of\_burst_i$) multiplied by the transfer delay of burst i ($delay\_of\_burst_i$), then divided by the sum of the size of burst i:

$$ND = \frac{\sum_{i=1}^{k} (delay\_of\_burst_i \,.\, size\_of\_burst_i)}{\sum_{i=1}^{k} size\_of\_burst_i} \qquad (5)$$

Where, M is the number of bursts, which are successfully received.

6.2 Performing Simulations

In performing the simulations, we consider a star OBS network with 15 edge nodes (i.e. $N = 15$). A dual-fiber was established between each edge node and the core node. We assume that all the dual-fibers are 200 Km in length, in other words, the Propagation Delay (PD) of each dual-fiber is 10^{-3}s. The number of data wavelengths is 8 per single fiber (i.e).

$W = 8$). The number of control wavelengths is large enough to ensure no losses for all control packets. The capacity of each wavelength (C) is 10Gbps. The core node has not wavelength conversion capability and has not fiber delay lines. The Processing Time (PT) of control packet and NACK packet is 10^{-6}s. The initial offset time between the control packet and its corresponding burst is 10^{-5}s. The sources generate bursts with the IDLE-ON process with an average burst length of 1Mbytes. Consequently, we can set the input traffic rate ($Load$) of each source by adjusting the average idle time $1/\alpha$:

$$Load = \frac{\alpha}{\alpha + \mu} \qquad (6)$$

The retransmission buffer size (RBS) at each edge node is:

$$RBS = a.\ B.\ \overbrace{2\,(PD + PT)}^{RTT} \qquad (7)$$

Where, a is a real number and B is the bandwidth of each fiber. B is given by:

$$B = W.\ C \qquad (8)$$

6.3 Simulation Results

In Figure 6 and Figure 7, we plot respectively the normalized delay and the normalized throughput as function of offered load for HS with small retransmission buffer $a = 0.5$, BRS with small retransmission buffer $a = 0.5$, BCS, and conventional OBS without using any

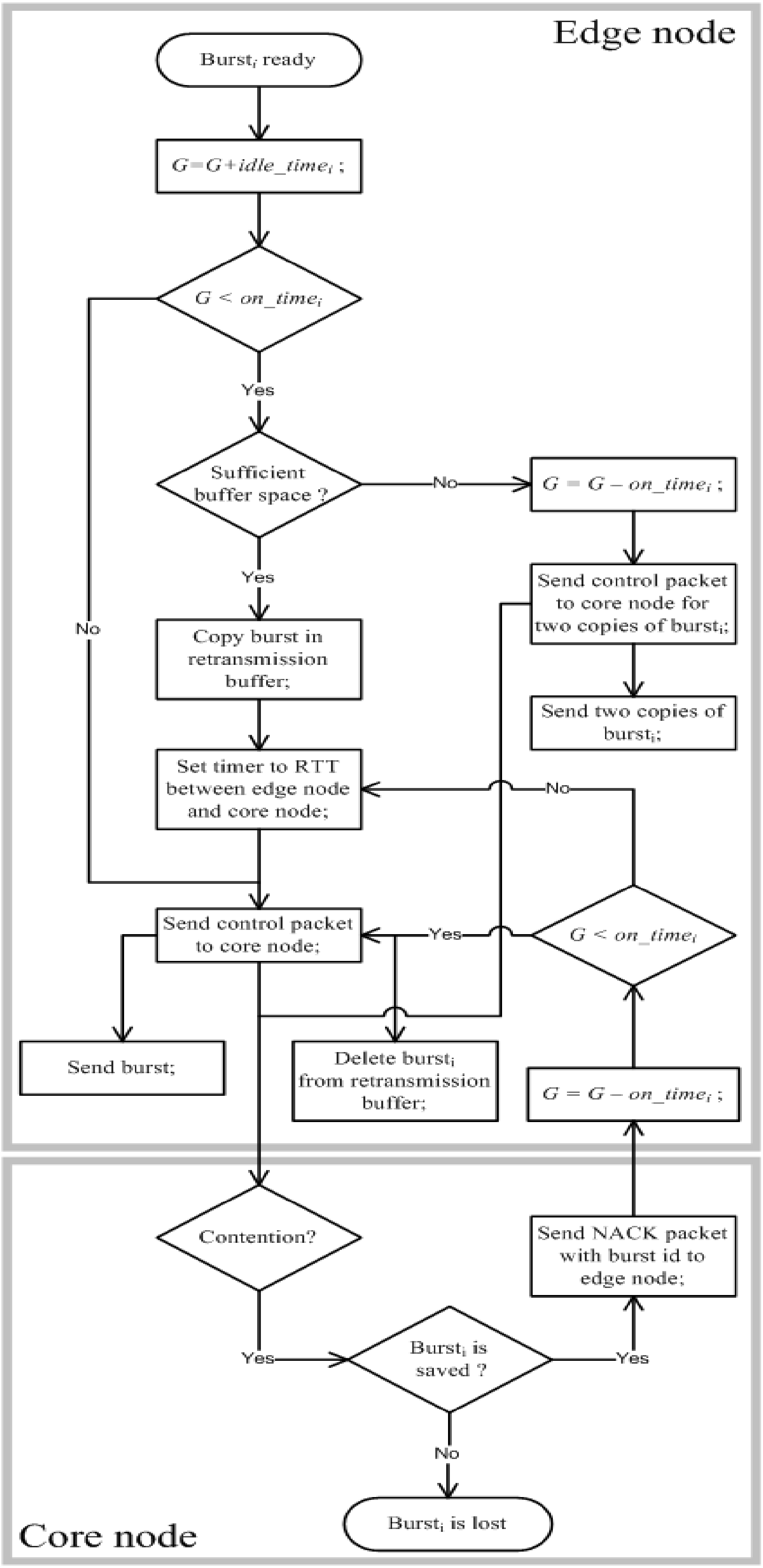

Figure 5. Hybrid Scheme

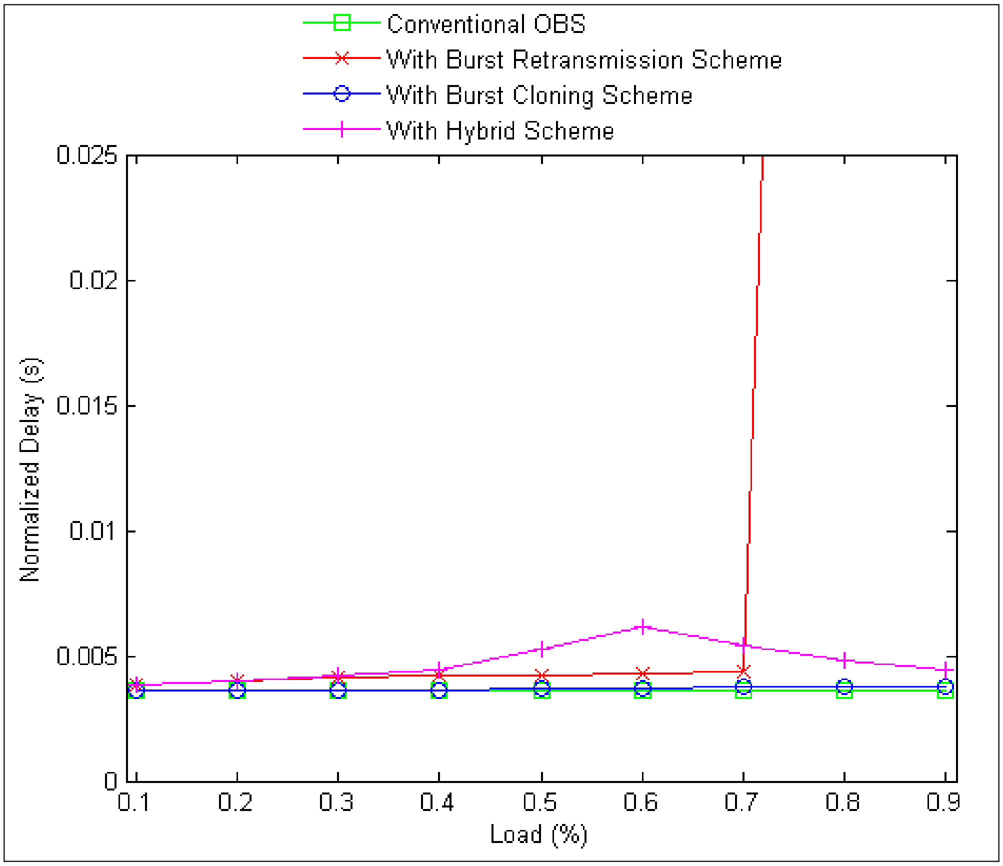

Figure 6. Normalized delay for very low to very high load with small retransmission buffer $\alpha = 0.5$

loss recovery mechanism.

In Figure 6, we observe that when the load is below 0.4, the normalized delay for the HS is very similar to that for the BRS and slowly increases as we increase the load. When the load is above 0.4 the normalized delay for the HS rapidly increases until it reaches its maximum value when the load is 0.6 and then begins to decrease as we increase the load. However, the normalized delay for the BRS slowly increases as we increase the load until the load is 0.7, after this point the normalized delay tends to infinite. For the BCS and the conventional OBS, the normalized delay remains almost constant and less than that of both the BRS and the HS for every load.

In Figure 7, it is clearly that the HS can achieve a better normalized throughput than the conventional OBS, the BRS and the BCS.

In Figure 8 and Figure 9, we plot respectively the normalized delay and the normalized throughput as function of offered load for HS with large retransmission buffer $a = 1$, BRS with large retransmission buffer $a = 1$, BCS, and conventional OBS without using any loss recovery mechanism.

In Figure 8, we show that when the load is below 0.4, the normalized delay for the HS is very similar to that for the BRS and slowly increases as we increase the load. When the load is above 0.4 the normalized delay for the HS rapidly increases until it reaches its maximum value when the load is 0.5 and then begins to decrease as we increase the load. However, the normalized delay for the BRS slowly increases as we increase the load until the load is 0.5, after this point the normalized delay tends to infinite. For the BCS and the conventional OBS, the normalized delay remains below that of both the BRS and the HS for every load.

In Figure 9, we show that when the load is below 0.5, the normalized throughput for the HS is very similar to that for the BRS and the both schemes (the HS and the BRS) can improve the normalized throughput more than the BCS. When the load is above 0.5,

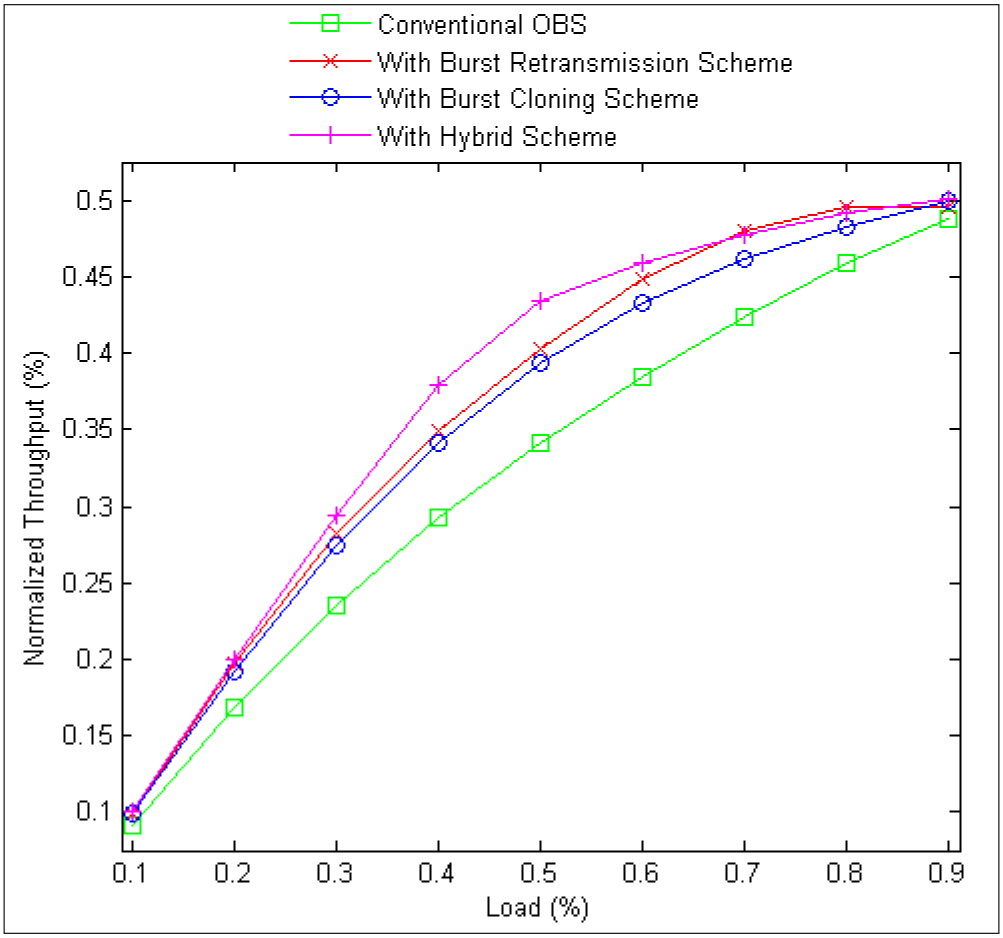

Figure 7. Normalized throughput for very low to very high load with small retransmission buffer $\alpha = 0.5$

the normalized throughput for the BRS continues to increase and reaches its maximum value when the load is 0.6 and then remains constant, but for the BCS, as we increase the load the normalized throughput continues to increase and becomes greater than that of BRS when the load is very high. For the HS, when the load is above 0.5 the normalized throughput continues to increase slowly as we increase the load and becomes greater than that of BRS when the load is very high.

6.4 Interpretation

The above simulation results can be interpreted as follows. For the conventional OBS and the BCS, the normalized delay depends mainly on the propagation delay, the offset time, and the average transmission time of bursts, that are successfully received. However, for the HS and the BRS, the normalized delay depends, in addition to the above factors, on delay introduced by waiting time in an $M/M/W$ queuing model and retransmission. For the BRS, as the load increases the burst contention probability increases and the extra load, due to the burst retransmission mechanism, increases until the input links to the core node are almost full. After this point, the network is overloaded, the normalized throughput reaches its maximum value and the normalized delay tends to infinite. Therefore, the BRS can improve significantly the normalized throughput while keeping low normalized delay only at low load, especially with large retransmission buffer. For the BCS, as the load increases the extra load, due to the burst cloning mechanism, increases until the input links to the core node are almost full. After this point, the BCS disables the burst cloning mechanism. Therefore, the BCS can improve the normalized throughput, but not significantly; however, it keeps very low normalized delay even at high load. The HS combines the advantages of BRS at low load with the benefits of BCS at high load through the decision algorithm. At the low load, the decision algorithm enables the burst retransmission and replaces the burst retransmission by the burst cloning in the case of insufficient the buffer space. When the load is high, the decision algorithm prevents the network overload by controlling the extra load and disables the both loss recovery mechanisms only when it is necessary. The decision algorithm can improve significantly the normalized throughput

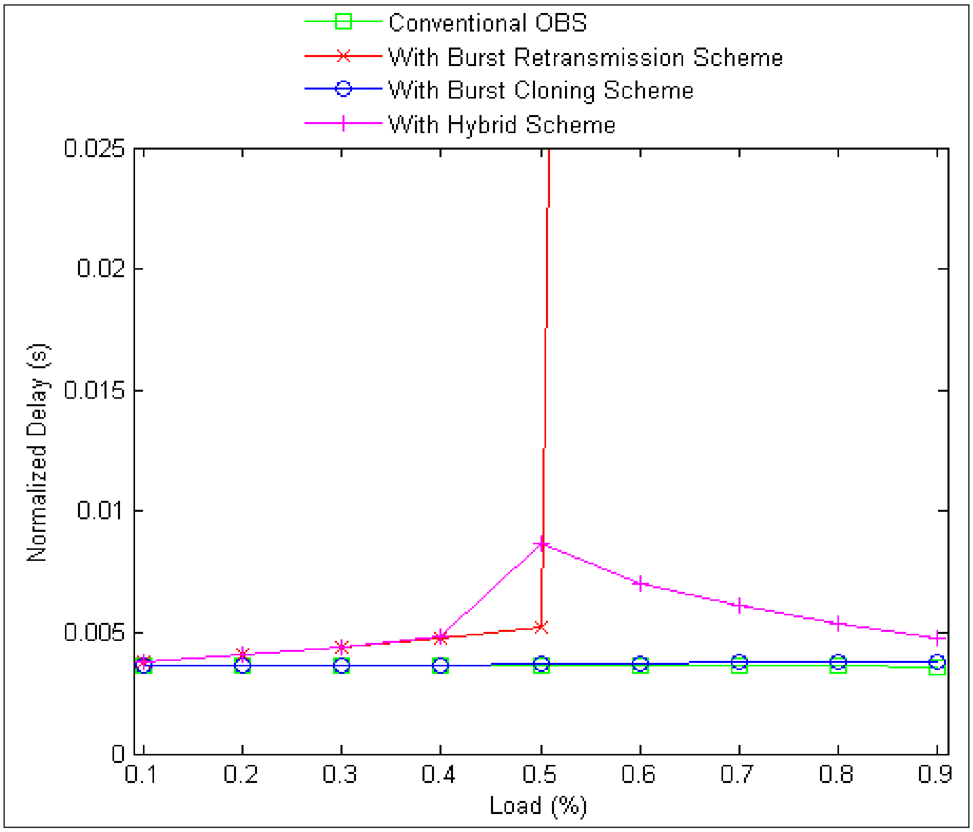

Figure 8. Normalized delay for very low to very high load with large retransmission buffer $a = 1$

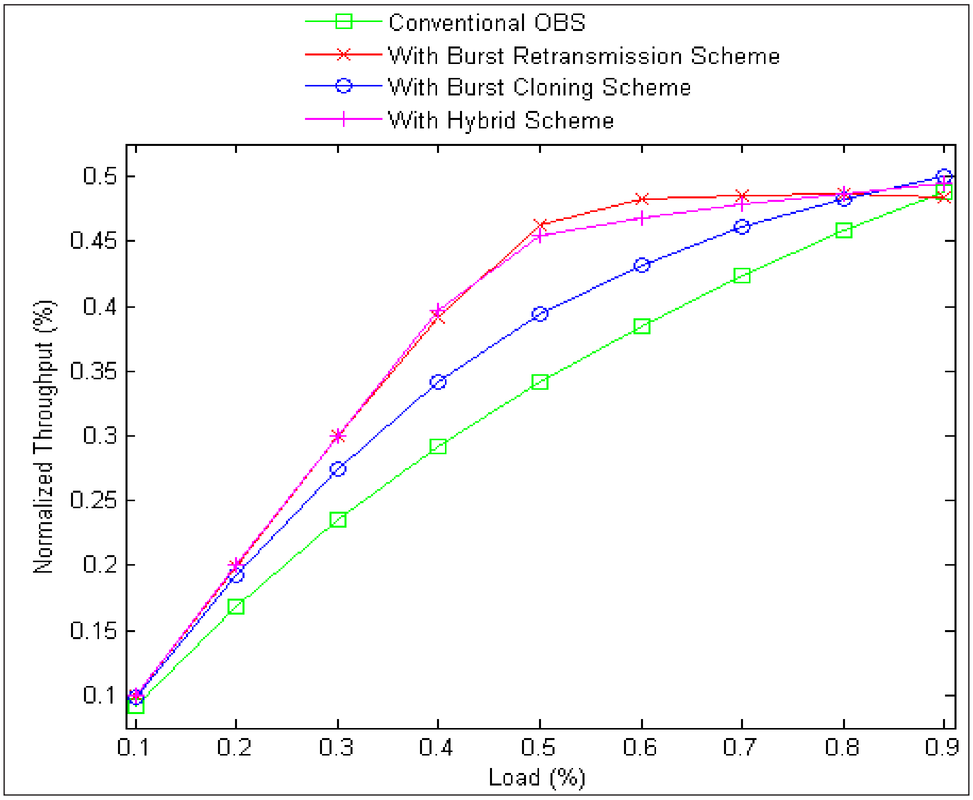

Figure 9. Normalized throughput for very low to very high load with large retransmission buffer $\alpha = 1$

while keeping reasonably low normalized delay even at high load. Consequently, the HS achieves better overall network performance than the both BRS and BCS.

7. Conclusion

In this paper, we proposed a hybrid burst retransmission and burst cloning scheme for star OBS networks. In order to take the advantages of the burst retransmission mechanism at low load with the benefits of the burst cloning mechanism at high load, the hybrid scheme controls the extra load due to the both loss recovery mechanisms. When an edge node is in overloaded state, the both loss recovery mechanisms are disabled; however, when the edge node is not in overloaded state, the edge node enables burst cloning mechanism only if the retransmission buffer space is insufficient that means the edge node is near to overloaded state, else it enables burst retransmission mechanism. Simulation results show that our hybrid scheme can achieve better overall network performance than both burst retransmission scheme and burst cloning scheme.

References

[1] Qiao, C., Yoo, M. (1999). Optical burst switching (OBS) — A new paradigm for an Optical Internet, *Journal of High Speed Networks*, 8 (1) 69–84, January.

[2] Yao, S., Mukherjee, B., Dixit, S. (2000). Advances in photonic packet switching: an overview, *IEEE Communications Magazine*, 38 (2) 84–94, February.

[3] Vokkarane, V. M., Zhang, Q. (2006). Forward redundancy: A loss recovery mechanism for optical burst-switched networks, Proceedings of the third IEEE/IFIP International Conference on Wireless and Optical Communications Networks, WOCN, Bangalore, India, April.

[4] Zhang, Q., Vokkarane, V. M., Wang, Y., Jue, J. P. (2005). Evaluation of burst retransmission in optical burst-switched networks, Proceedings of IEEE BROADNETS, p. 276–282, October.

[5] Maach, A., Bochmann, G., Mouftah, H. T. (2004). Robust optical burst switching, *In*: Proceedings of Networks, Vienna, Austria, p. 447-452, June.

[6] Agustí-Torra, A., Bochmann, B., Cervelló-Pastor, C. (2005). Retransmission schemes for optical burst switching over star networks, *In*: Proceedings of the 2nd IFIP *International Conference on Wireless and Optical Communications Networks*, WOCN, Dubai, UAE, March.

[7] Riadi, S., Maach, A. (2006). A controlled retransmission scheme for optical burst switching over star networks, *International Conference on Multimedia Computing and Systems*, ICMCS'12, Tangier, Morocco, May.

[8] Huang, H., Vokkarane, V. M., Jue, J. P. (2005). Burst cloning: A proactive scheme to reduce data loss in optical burst-switched networks, *In*: Proceedings of IEEE International Conference on Communications (ICC), Seoul, South Korea, May.

[9]Vickers, Beshai, M. (2000). PetaWeb architecture, Networks Symposium, Toronto, Canada, September.

[10] Blouin, F. J., Lee, A. W., Lee, A. J. M., Beshai, M. (2003). Comparison of two optical-core networks, *Journal Optical Networking*, 1 (1) 56–65, January.

[11] Mountrouidou, X., Puttasubbappa, V. S., Perros, H. G. (2006). A zero burst loss architecture for star OBS networks, Net-Con'06, (part of the IFIP World Computer Congress 2006), Sandiago, Chile, August.

[12] Reinert, T., Sansò, B., Secci, S. (2009). Design optimization of the Petaweb architecture, *IEEE/ACM Transactions on Networking*, 17 (1) 332–345, February .

[13] Bochmann, G., Coates, M. J., Hall, T. J., Mason, L., Vickers, R., Yang, O. (2004). The agile all-photonic network: An architectural outline, Queen's 22nd Biennial Symposium on Communications, May.

[14] Battestilli. T., Perros, H. G. (2004). End-to-end burst loss probabilities in an OBS network with simultaneous link possession, *In*: Proceedings of the Third International Workshop on Optical Burst Switching, WOBS3, San Jose, CA, October.

[15] OIRC OBS-ns Simulator, [Online]. Available: http://wine.icu.ac.kr/~obsns/index.php.

[16] ns-2 Network Simulator, (2009). [Online]. Available: http://www.isi.edu/nsnam/ns.