A Novel QoS Provisioning Scheme for OBS Networks

Shavan K. Askar, Georgios Zervas, David K. Hunter, and Dimitra Simeonidou

School of Computer Science and Electronic Engineering
University of Essex, Colchester CO4 3SQ
skaske3@essex.ac.uk

Abstract. This paper presents Classified Cloning, a novel QoS provisioning mechanism for OBS networks carrying real-time applications (such as video on demand, Voice over IP, online gaming and Grid computing). It provides such applications with a minimum loss rate while minimizing end-to-end delay and jitter. ns-2 has been used as the simulation tool, with new OBS modules having been developed for performance evaluation purposes. Ingress node performance has been investigated, as well as the overall performance of the suggested scheme. The results obtained showed that new scheme has superior performance to classical cloning. In particular, QoS provisioning offers a guaranteed burst loss rate, and delay unlike existing proposals for QoS implementation in OBS which use the burst offset time to provide such differentiation. Indeed, classical schemes increase both end-to-end delay and jitter. It is shown that the burst loss rate is reduced by 50% reduced over classical cloning.

Keywords: QoS provisioning, Optical Burst Switching, Cloning.

1 Introduction

Optical Burst Switching (OBS) is an effective technology for the next generation optical Internet that aims to address the increasing bandwidth required by Internet users [1]. OBS is a good tradeoff between traditional Optical Circuit Switching (OCS), which is relatively easy to implement but suffers from poor bandwidth utilization and coarse granularity, and Optical Packet Switching (OPS) [2], which has a good bandwidth utilization and fine granularity but is difficult to implement because of the immaturity of current optical technologies [3]. In OBS networks, the basic switching entity is a burst. Prior to transmission of a burst, a control packet is created and immediately sent toward the destination in order to set up a buffer-less optical path for the corresponding burst. After an offset delay time, the data burst is transmitted without waiting for an acknowledgement from the destination node. The optical path exists only for the duration of a burst [4].

There has been a rapid increase in the volume of traffic from new applications (such as video on demand, Voice over IP, online gaming or Grid computing) which have real-time and/or bandwidth constraints. Hence, service differentiation must be provided for such applications in order to reduce the loss rate while maintaining the lowest possible end-to-end delay. Accordingly, the high burst loss probability evident

I. Tomkos et al. (Eds.): BROADNETS 2010, LNICST 66, pp. 418–428, 2012.

in OBS networks has become a critical issue that must be addressed in order to enable real deployment of OBS networks [5-7]. Most existing research in this area can be categorized into one of the mechanism shown in Fig. 1.

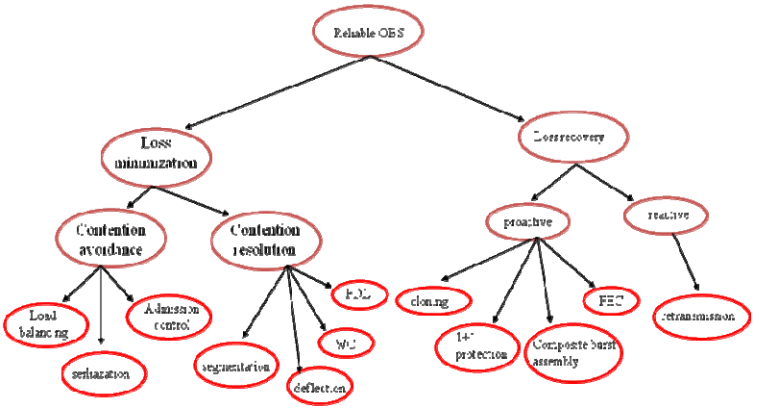

Fig. 1. OBS Mechanisms for loss reduction

The mechanisms shown in Fig. 1 are used for loss reduction in OBS networks – they are categorized into loss recovery and loss minimization techniques. Loss recovery mechanisms can be divided into sub-categories – reactive and proactive mechanisms – while loss minimization approaches are subdivided into contention resolution and contention avoidance schemes. In fact, each of these techniques has its advantages as well as its disadvantages, but all of them seek to reduce the loss rate in OBS networks. Contention avoidance aims at preventing the occurrence of contention, while contention resolution focuses on resolving contention that already exists. The most well-known contention resolution schemes are wavelength conversion [8, 9], fiber delay line (FDL) buffering [10] and deflection routing [11, 12]. Another technique called burst segmentation proposes segmentation of contended bursts [13, 14], and dropping only part of each one.

Contention resolution appears to be a very tempting solution to the problem of contention in OBS networks, however there are a number of implementation problems: 1) wavelength conversion is an immature technique which is still very expensive to implement, 2) FDL's are bulky and they merely offer fixed delays which generally reduce channel utilization because they generate voids between scheduled bursts, 3) deflection routing suffers from the problem of endless loops as well as the possibility of insufficient offset time for rerouted bursts, 4) burst segmentation is still very complicated to implement. Reactive loss recovery is a retransmission scheme where burst retransmission is possible in the event of contention [15, 16]. Many factors hinder its implementation; firstly, very large buffers are required in ingress nodes in order to implement retransmission. Also, although retransmission may be practical in LAN's, it is not useful in MAN's or WAN's because of their higher latency, which also requires larger buffers in order to

implement retransmission. Finally, a notification protocol is required to notify edge nodes of burst losses, which generates additional load on the control channel. To overcome these problems we propose a new scheme for QoS provisioning with real-time applications – Classified Cloning – which is inspired by the basic cloning scheme [17]. In this paper we investigate the use of cloning to reduce packet loss. Research in this area is limited, with contradictions in the results from different studies [17-20]. However, it has been shown that the existing drawbacks of burst retransmission, such as the large buffer size and increased control traffic, can be avoided through cloning, yielding lower mean packet delay.

The rest of this paper is organized as follows: Section 2 provides a brief overview of the existing cloning scheme. Section 3 introduces our proposed scheme and a novel ingress node design. Section 4 evaluates the performance of the proposed Classified Cloning Scheme and compares the performance with the existing Basic Cloning Scheme. Section 5 concludes the paper.

2 Existing Cloning Schemes

In this section the existing Burst Cloning Scheme [17] is referred to as the Basic Cloning Scheme (BCS), while our proposed scheme will be referred as the Classified Cloning Scheme (CCS); the latter can provision traffic with higher priority QoS. In both the original BCS and our proposed Classified Cloning Scheme (CCS), the original copy of a burst is referred to as the "original burst", and the duplicate copy as the "cloned burst". Similarly, the traffic corresponding to the original and cloned bursts is referred as "original" and "cloned" traffic respectively. The node at which cloning is performed is referred to as the "cloning node".

In BCS, one or more cloned bursts can be made from each original burst and sent simultaneously; if one or more of these bursts arrive at the destination, the original burst is considered to be successful. On one hand, if more copies are made for a particular burst then it is less likely to be lost. On the other hand, if more copies are made overall, more cloned traffic is added to the network, which then actually increases the overall probability of burst loss.

A comparison has been made between a retransmission recovery scheme and a cloning scheme [18]. It was found that the drawbacks of the existing retransmission mechanism such as the use of large buffers and increased control channel traffic can be avoided through the use of cloning. Accordingly, lower average packet delay value was delivered.

In [19] each core node has the burden of determining whether cloned bursts have been lost or not, in order to decide whether another cloned burst must be produced. To do this, two assumptions are made. Firstly, it is assumed that each cloned burst arrives at a particular core node before its corresponding original burst. Secondly, it is assumed that there is enough time between receiving the original BHP and receiving the corresponding data to check the status of the received cloned bursts. In fact, implementing cloning in the core nodes is not recommended due to the complexity of

implementing a database there, which must be accessed when every burst is received to determine whether each cloned burst has been lost. Indeed, cloning was originally proposed as a low cost alternative to the solutions mentioned in the Introduction, which require expensive hardware.

3 The Proposed Classified Cloning Scheme

The major side effect of burst cloning is increased network load, although BCS introduces a traffic isolation mechanism which allows original bursts to preempt reservations made by cloned bursts. The optical links on average carry twice the original load, or more, since some studies suggest making more than one copy of each original burst. However, having as many cloned bursts can be counterproductive because the probability of contention often actually increases due to the overall increase in traffic. To the best of the authors' knowledge, the use of cloning for QoS provisioning has not yet been suggested, and all research in this area has involved cloning all traffic in the network [17].

In our proposed CCS we seek to avoid cloning all incoming traffic because otherwise, the network will be heavily loaded by cloned traffic without much effective reduction in burst loss. The consequent low reduction in loss rate with BCS arises because of the low priority assigned to cloned traffic in order to provision class isolation. Many studies attempt to overcome this by implementing cloning in a core node, or by making many copies of each original burst.

By using the edge nodes for cloning and applying cloning only to UDP traffic, which possibly makes up 10% of the total, there is not a major effect on the network load; however the loss rate is reduced considerably. The reduction in loss rate benefits UDP-based applications because they are time-critical, and recovery from burst loss should therefore be immediate. Furthermore, the ETE delay is maintained because unlike BCS, no extra offset time is added before each burst. Fig. 2 shows the proposed edge node design; we have designed a classifier in the ingress node which classifies incoming IP packets depending on their type of service into either Serv 1 or Serv 2 packets (Serv 1 is for best effort while Serv 2 is for real-time applications). There are two buffers: the primary buffer aggregates all traffic (Serv 1+ Serv 2) while the secondary buffer aggregates traffic from real-time applications only (Serv 2). The secondary (cloning) buffer receives IP packets forming Serv 2 traffic, but only when the offered load is low or medium. This is implemented through the write enable (WE) signal, which goes low to enable writing if RT (Real-Time) and TRG (TRiGger) signals both become high. RT goes high if the IP packet belongs to Serv 2 traffic while the TRG signal is activated if the offered load is low or medium. The classifier classifies incoming IP packets according to their destination egress node. After aggregating IP packets in this way, there are two types of burst (namely original bursts and cloned bursts), which both have the same priority and are sent to the egress node which then segregates received bursts and drops duplicates where necessary.

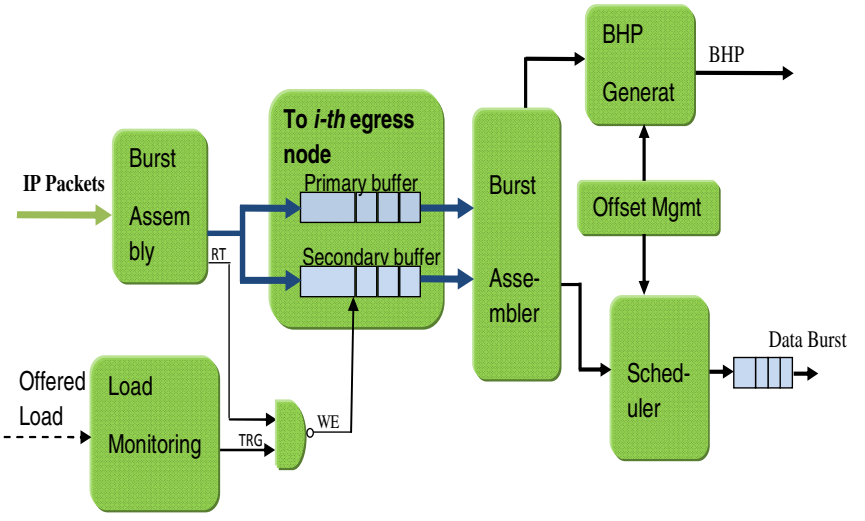

Fig. 2. Edge-ingress node model for Classified Cloning Scheme

4 Simulation and Analysis

The ns-2 simulator has been used to evaluate our proposed cloning scheme. New OBS modules have been developed to support the proposed concepts. Simulations took place on the 19-node NSF network topology of Fig. 3. A list of functionalities and simulation decisions made are shown below:

- the mean burst size is 125KBytes,
- the core nodes are bufferless,
- the wavelength continuity constraint is applied, and
- the OBS control plane supports JET (Just Enough Time).

Furthermore, in the absence of a detailed traffic model, we assume that the bursts which are generated at the network edge are described by a Poisson process, the traffic is distributed over the network uniformly, and all routes are established by a shortest path routing algorithm with the number of hops as the metric. In addition, bursts are assembled using hybrid threshold-timeout, with both timeout and size thresholds being used to obtain the best of both schemes.

Fig. 4 shows burst loss in the ingress node versus offered load. Fig. 5 shows the average delay and jitter versus the offered load, which arises because of aggregation in each edge node. Existing research does not consider loss at the ingress node when evaluating burst loss, moreover, many existing publications don't show jitter at the edge node; in fact, the edge node aggravates jitter, thus influencing the performance of the whole network. The jitter and delay values in Fig. 5 obtained are the average values for the corresponding offered load.

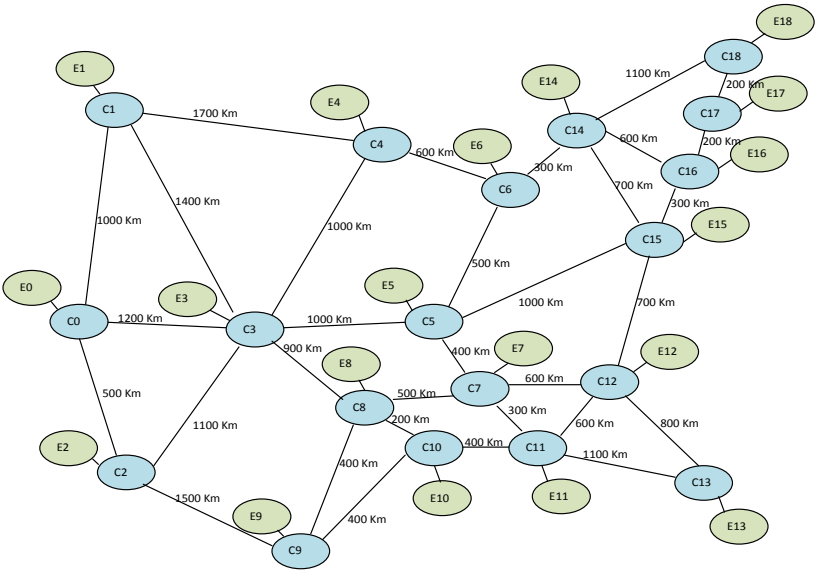

Fig. 3. A 19-node NSF network topology with real physical distances between the nodes

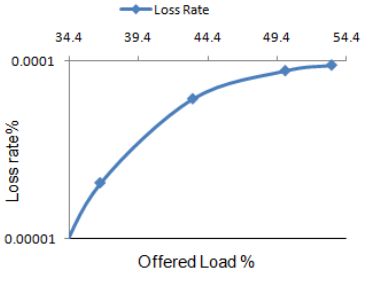

Fig. 4. Edge node loss rate versus the offered load

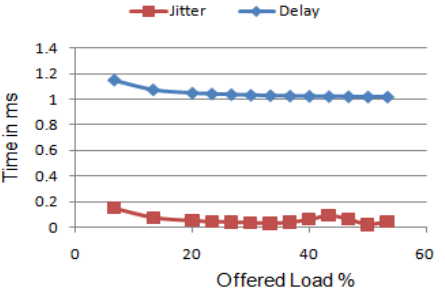

Fig. 5. Edge node delay and jitter versus offered load

Fig. 6 shows the Cumulative Distribution Function (CDF) of the jitter values with offered loads of 6.66%, 19.9%, 26.6%, and 33.3%, while Fig. 7 will shows offered loads of 39.9%, 46.6%, and 53.3%; two figures are provided in the interest of clarity. With a 6.66% offered load, the jitter is greater than with the other scenarios; with such a low level of incoming traffic, bursts are shaped with long interarrival times, even although hybrid aggreation is used. Hence with a low volume of incoming traffic, most of the generated bursts arise after the burst generation algorithm has timed out, even with hybrid aggregation – this is confirmed by Fig. 6, where no burst is of the maximum size. In Fig. 6, 93% (73%) of the generated bursts have jitter values below

0.4 ms (0.2 ms), with an offered load of 6.66%, which is acceptable with this low load. However, as the level of incoming traffic increases, jitter decreases (i.e. when the offered loads are 19.9%, 26.6%, and 33.3%). Out of these three scenarios, there is the greatest jitter with a load of 19.9% for the reasons discussed above, so that 98%, 87%, and 62% of the generated bursts have jitter vlaues of 0.2 ms, 0.1 ms, and 0.05 ms or less respectively. The CDFs of the jitter values at 26.6% offered load show that 93%, 73%, and 49% of the generated bursts have jitter values of 0.1 ms, 0.05 ms, and 0.025 ms or less respectively, while with a 33.3% offered load, 96.5%, 81%, and 56% of the generated bursts have jitter values of 0.1 ms, 0.05 ms, and 0.025 ms or less. Indeed, Fig. 6 shows that in general, jitter decreases as offered load decreases. However, there is a trade-off to be made when choosing burstification parameters, because a higher load will, in consequence, increase the edge loss rate as shown in Fig. 4.

Fig. 7 shows the CDFs for jitter at offered loads of 39.9%, 46.6%, and 53.3%. For jitter less than 1 ms, the CDFs are 92.4%, 93%, and 98.14% for offered loads of 39.9%, 46.6%, and 53.3% respectively while for jitter values less than 0.05 ms, the CDFs are 43%, 29.7%, and 73% with offered loads of 39.9%, 46.6%, and 53.3%.

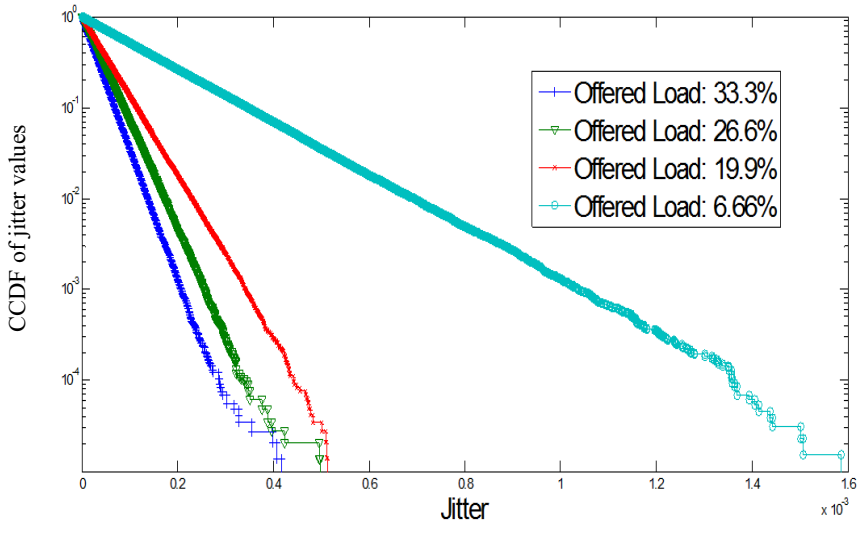

Fig. 6. CCDF (i.e. 1 - CDF) of Burst Jitter at 6.66% , 19.9%, 26.6%, and 33.3% offered loads

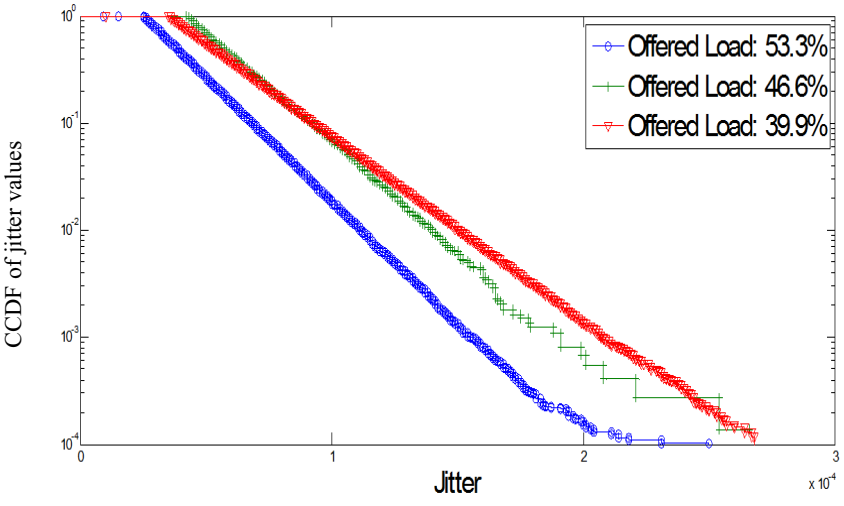

Fig. 7. CCDF (i.e. 1 - CDF) of Burst Jitter at 39.9%, 46.6%, and 53.3% offered loads

In order to conduct a comprehensive investigation of ingress node performance, the generated burst sizes must be studied in detail. Fig. 8 shows the CDFs of the generated burst sizes under different offered load scenarios, in order to provide further insight into the burst generation process. The Figure shows how the burst sizes increase as incoming load increases, and these results can be used to facilitate choosing suitable burstification parameters in order to enhance ingress node performance. Existing studies on choosing burstification parameters appear to be contradictory when specifying how to choose the maximum threshold value. Some studies suggest generating longer bursts in order to reduce control packet processing in core nodes, hence reducing the loss rate, while others argue that longer bursts increase the probability of preemption by other bursts because they occupy the link for longer It has also been claimed that bursts of equal length generated by ingress nodes will reduce the probability of loss due to contention [7, 21, 22]. As we already mentioned, we used hybrid aggregation because we sought a trade-off between loss, delay, and jitter at the ingress node when considering the burst loss rate arising in the network overall.

Fig. 9 shows the improvement in loss rate with real-time applications under the proposed Classified Cloning Scheme. The loss rate for Serv 2 applications has been reduced by more than 50% over the Basic Cloning Scheme. Because cloned traffic is sent with same priority as the original traffic, the loss rate for Serv 1 traffic increases due to the increased probability of contention arising from the additional cloned Serv 2 traffic. However real-time applications typically produce 10% of the total traffic, therefore, the loss rate with Serv 1 increases very slowly with CCS as shown in Fig. 10.

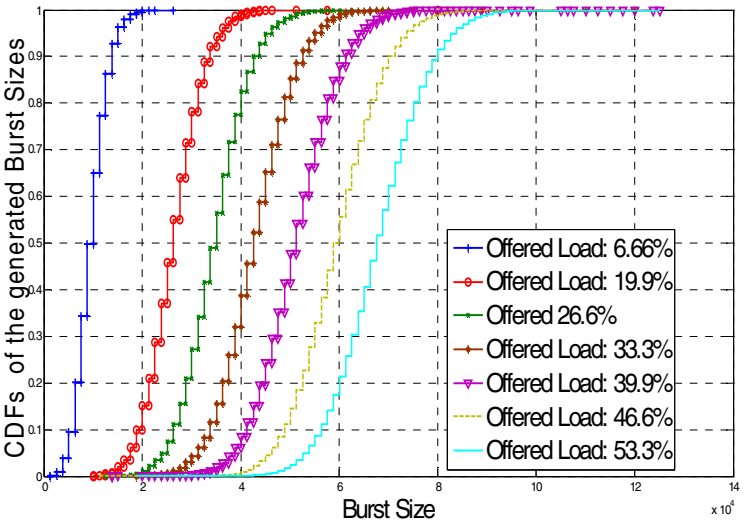

Fig. 8. CDF of generated burst sizes at different offered loads

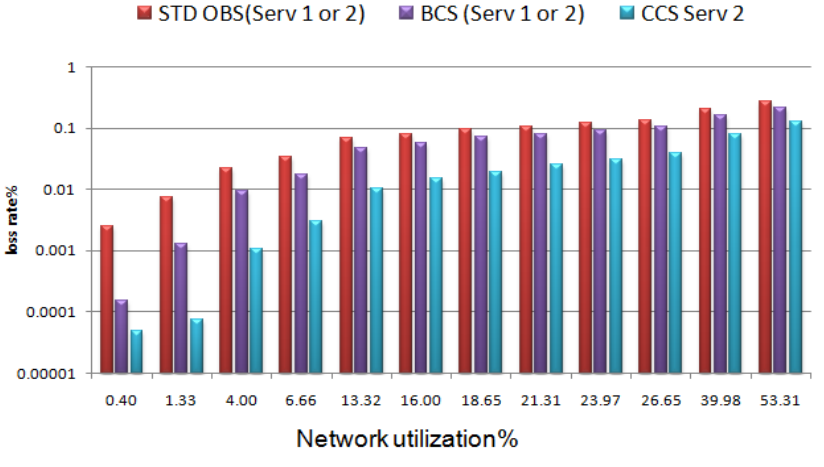

Fig. 9. Loss rate versus network load for standard OBS, BCS, and CCS

However, Serv 1 applications don't have jitter, delay and loss rate constraints. Providing that no extra offset time is added to the original traffic with CCS, Serv 2 bursts have lower ETE delay than with BCS. Applying cloning at the edge node to real-time UDP traffic does not affect network load appreciably, while it nevertheless reduces the burst loss rate considerably. The reduced loss rate benefits UDP-based applications by providing immediate burst loss recovery through CCS. Furthermore, the ETE delay is preserved because unlike BCS, no extra offset time is added to the original traffic.

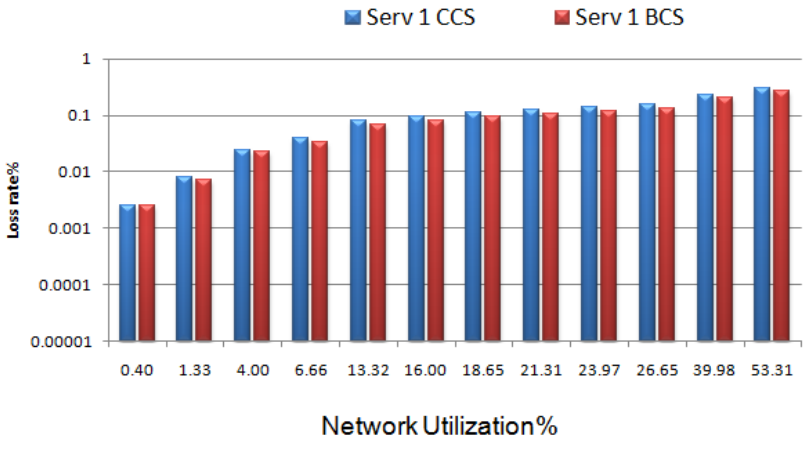

Fig. 10. Comparison of Serv 1 loss rate between BCS and CCS

5 Conclusion

This paper introduced a novel QoS provisioning scheme –Classified Cloning – for OBS networks. The ns-2 simulator was used as the simulation tool, with OBS modules being developed and compiled into the ns-2 simulator to evaluate the proposed mechanism. For real-time traffic, the results show a 50% reduction in burst loss rate over BCS. Additionally, evaluation of ingress node performance has shown the jitter, delay and loss rate values produced due to aggregation inside the edge-ingress node. The Classified Cloning Scheme outperforms BCS and the classical QoS provisioning mechanisms in OBS for three reasons: firstly, it retains the same delay as without cloning because the Classified Cloning Scheme does not use an extra offset time for class isolation; secondly, it implements immediate loss recovery for real-time applications; thirdly, it does not need extra hardware or optical splitting because classified cloning is implemented in the ingress node. We conclude that the proposed CCS scheme is a viable and realistic alternative to QoS provisioning schemes because it doesn't add extra offset time, and furthermore it offers significant improvements in reduction of burst loss rate.

References

1. Qiao, C., Yoo, M.: Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet. Journal of High Speed Networks 8(1), 69–84 (1999)
2. Bianco, A., et al.: Design of optical packet switching networks. In: GLOBECOM, New York, vol. 3, pp. 2752–2756 (2002)
3. Perros, H.G., Rouskas, G.: Techniques for optical packet switching and optical burst switching. IEEE Communications Magazine 39, 136–142 (2001)

4. Luo, J., et al.: ROBS: A novel architecture of Reliable Optical Burst Switching with congestion control. Journal of High Speed Networks 16, 123–131 (2007)
5. Simeonidou, D., et al.: Dynamic Optical-Network Architectures and Technologies for Existing and Emerging Grid Services. Journal of Lightwave Technology 23, 3347–3357 (2005)
6. Xu, H., Fan, G.: Analysis of the burst loss rate in OBS rings with depth limited optical buffers. Journal of High Speed Networks 16, 341–351 (2007)
7. Vargas, T.R., Guerri, J.C., Sales, S.: Optimal Configuration for Size-Based Burst Assembly Algorithms at the Edge Node for Video Traffic Transmissions over OBS Networks. Building, 130–133 (2008)
8. Olmos, J.J.V., et al.: Optical node with time-space-and-wavelength domain contention resolution, deflection and dropping capability. Optics Express 14, 11545–11550 (2006)
9. Rosberg, Z., et al.: Analysis of OBS Networks With Limited Wavelength Conversion. IEEE/ACM Transactions on Networking 14, 1118–1127 (2006)
10. Pedro, J., et al.: Efficient Optical Burst-Switched Networks using only Fiber Delay Line Buffers for Contention Resolution. Networks (2006)
11. Hsu, C., Liu, T., Huang, N.: Performance analysis of Deflection Routing in Optical Burst Switched Networks. In: Proceedings, IEEE INFOCOM (2002)
12. Wang, X., Morikawa, H., Aoyama, T.: Deflection routing protocol for burst switching WDM mesh networks. In: Proc. SPIE/IEEE Terabit Optical Networking: Architecture, Control, and Management Issues, pp. 242–252 (2000)
13. Vokkarane, V.M., Member, S., Jue, J.P.: Prioritized Burst Segmentation and Composite Burst-Assembly Techniques for QoS Support in Optical Burst-Switched Networks. IEEE Journal on Selected Areas in Communications 21, 1198–1209 (2003)
14. Vokkarane, V.M., Jue, J.P., Sitaraman, S.: Burst Segmentation: An Approach For Reducing Packet Loss In Optical Burst Switched Networks. Time, 2673–2677 (2002)
15. Hirota, Y., Tode, H., Murakami, K.: A Study on RWA Cooperation Method Considering Retransmission. IEIC Journal 106(208), 55–60 (2006)
16. Kwak, K.J., Coffman, E.: Retransmission in OBS networks with fiber delay lines. In: Fourth International Conference on Broadband, Communication, Networks, and Systems, BROADNETS 2007 (2007)
17. Huang, X., Vokkarane, V.M., Jue, J.P.: Burst Cloning: A Proactive Scheme to Reduce Data Loss in Optical Burst-Switched Networks. Science, 1673–1677 (2005)
18. Vokkarane, V.M., Zhang, Q.: Forward Redundancy: A Loss Recovery Mechanism for Optical Burst-Switched Networks. Wireless and Optical Communications Networks, IFIP, 5 (2006)
19. Um, T.-W., et al.: priority based duplicate burst transmission in obs. ETRI 30 (2008)
20. Maach, A., Hafid, A.S., Belbekkouche, A.: Burst Loss Reduction schemes in Optical Burst Switching Networks. Network, 256–262 (2008)
21. Watagodakumbura, C., Praluyanto, H.: Composite Burst Assembly and Lower Real-Time Class Utilization Threshold in Optical Burst Switching As Means of Overcoming Effects of Self-Similarity. ITNG (2007)
22. Casoni, M., Luppi, E., Merani, M.L.: Impact of Assembly Algorithms on End-to-End Performance in Optical Burst Switched Networks with Different QoS Classes. In: IEEE/SPIE Third Workshop on Optical Burst Switching 2004, San Jose, CA (October 2004)