Dynamic Routing of Reliability-Differentiated Connections in Optical Burst Switched Networks

N Sreenath
Department of Computer Science and Engineering
Pondicherry Engineering College
Pondicherry, India
nsreenath@pec.edu

N Devendra
Department of Computer Science and Engineering
Pondicherry Engineering College
Pondicherry, India
devendra_ndr@yahoo.co.in

Balaji Palanisamy
College of Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
pbalaji@cc.gatech.edu

ABSTRACT

Optical burst switching (OBS) is one of the most promising next-generation all-optical data transport paradigms. As networks become increasingly distributed and autonomic, Optical Burst Switching becomes the right choice for the next generation optical Internet. In this paper, we propose a mechanism for Dynamic Routing of Reliability-Differentiated connections (DRRDC) in Optical Burst Switched networks. The proposed mechanism consists of two subschemes namely Adaptive Routing, a loss minimization mechanism that selects the least congested route for burst scheduling and Adaptive Burst Cloning, a technique for providing loss recovery in an OBS network. We develop a network simulation model to investigate the proposed DRRDC scheme and compare its performance with the existing prioritized burst scheduling QoS scheme. Our results show that the proposed service differentiation mechanism has a significantly low packet loss compared to the existing prioritized burst scheduling scheme in an optical burst switched network.

Keywords

WDM Optical Networks, Optical Burst Switching, Quality-of-Service, Burst Cloning, Service Differentiation

1. INTRODUCTION

Optical burst switching (OBS) is an emerging solution to achieve all-optical WDM networks [1-2]. It combines the advantages of optical circuit switching and optical packet switching [3-4]. In the past few years, various solutions have been proposed and analyzed in an attempt to improve the performance of OBS networks [5-7]. In OBS networks, the basic switching entity is a burst. Prior to transmitting a burst, a control packet is created and sent towards the destination to set up a buffer-less optical path. 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. OBS provides a huge bandwidth which could alleviate the increasing demands of Internet traffic; however, challenges remain on how to provide Quality of Service (QoS) for Internet applications in such a network. For example, applications such as Internet telephony and video conferencing require a higher QoS than electronic mail and general web browsing. In an IP network, many methods have been proposed to implement QoS such as fair queuing, weighted fair queuing, frame-based fair queuing, etc. However, all of these methods are based on employing buffers at the network nodes. To implement the existing QoS mechanisms to differentiate services, all intermediate nodes should have a certain amount of buffer space. However, the use of electronic buffer necessitates O/E and E/O conversions which sacrifice the data transparency in addition to having increased latency.

Optical burst-switched networks are typically connectionless in nature; thus, it is likely that there will be contention for resources in the core network, leading to packet loss. Contention resolution is an important research issue in the context of QoS provisioning in OBS networks. When two or more bursts are destined for the same output port at the same time, contention occurs. When a contention cannot be resolved, one of the contenting burst is lost. If the dropped burst cannot be recovered at the OBS layer, higher layers (such as TCP) will need to handle the retransmission of the lost data at a later time. In this paper, we propose a novel service differentiation mechanism that uses two subschemes namely Adaptive Routing and Adaptive Burst Cloning as techniques for providing service differentiation and reducing the over-all Burst Loss Rate (BLR). Our proactive data loss reduction scheme namely Adaptive burst cloning replicates a burst and sends duplicated copies of the burst through the network simultaneously. In case the original burst gets lost, the cloned burst may still be able to reach the destination. Adaptive burst Cloning is different from the conventional Burst cloning [9] in two aspects, one in selecting the node where to clone the burst and in choosing the number of cloned bursts for each original burst i.e. number of duplicate copies. The two vital factors to be considered in burst cloning include selection of the optimal nodes at which cloning needs to be done and preventing the contention of the cloned bursts with the original bursts. We address both these issues in our work. We develop a simulation model to investigate the proposed schemes namely Adaptive Routing loss minimization mechanism and Adaptive Burst Cloning loss recovery mechanism and quantify their performance in providing service differentiation among the traffic classes. We compare the performance of our proposed QoS mechanism with the prioritized burst scheduling QoS scheme [8]. Three classes of services, class0, class1 and class2, are considered in our work. Class0 is assumed to have the highest priority and class 2 is of low priority where as

---

This research is sponsored by the All India Council for Technical Education (AICTE). Ref:8023/RID/BOR/RPS-76/2005-06

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MCDES 2008, May 27–31, 2008, Bangalore, India.

Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...$5.00.

class1 is intermediate between class0 and class1. In our work, we assume that no buffers are used in the optical layer, which is highly desirable in all-optical networks. The paper is organized as follows. Section 2 describes the proposed Adaptive Routing loss minimization scheme. Section 3 introduces and discusses the issues involved in Burst cloning. Section 4 presents the proposed Adaptive Burst Cloning loss recovery mechanism. Section 5 describes the service differentiation scheme which combines Adaptive Routing and Adaptive Burst Cloning schemes. Section 6 discusses the simulation experiments and presents the simulation results and section 7 concludes the paper.

2. ADAPTIVE ROUTING

Adaptive Routing routes the bursts along the least congested path between the source-destination pair. It chooses the least congested path based on the current link and route loss probabilities of the various candidate paths. R-candidate routes are pre-computed to every destination and are available at all source nodes. The estimate of the loss probability could be made on a per-link basis and route loss probabilities can be estimated using the link loss probabilities. We use a technique similar to the one used in [10] to compute the link loss probabilities of the individual links. In our policy, link loss probability is computed at all nodes in the network, one for every outgoing link from the node. The loss probability estimate is initially set to zero. When a burst is successfully transmitted over a link, a positive feedback is generated, and when a burst is dropped on the link, a negative feedback is generated. Based on the updating scheme, the loss probability estimate for the link is computed at regular intervals. The source of every flow periodically sends a probe packet along the shortest routes to collect the loss probability estimates on all links along the route. The loss probability for the entire route is calculated from the loss probability estimates on each link. The estimation of loss probability of links and loss probability of routes are described in section 2.1 and section 2.2 respectively.

2.1 Link loss probability estimation

Initially the loss probability of all links is set to zero. For each link, we record two parameters, the number of bursts arrived into that link and the number of bursts dropped on that link. Initially, for each link the number of bursts arrived and the number of bursts dropped is set to zero. Based on the feedback received, these two parameters of the link are updated. For a positive feedback (i.e. successful burst transmission), the number of bursts arrived on the link is incremented and for a negative feedback (i.e. for burst drop on the link), both the parameters (i.e. number of bursts arrived and number of bursts dropped) are incremented by one. Loss probability of the link is the ratio of the total number of bursts dropped on that link to the total number of bursts arrived on that link.

2.2 Route loss probability estimation

The source of every flow periodically sends a probe packet along the pre-computed shortest route to collect the loss probability estimates on all links along the route. The loss probability for the entire route is defined as the maximum of the loss probabilities of all the individual links in that route. For routing a burst, the source node selects the least congested route (the one having the least route-loss probability) in order to minimize burst contention. Dynamic selection of the least congested route ensures very low burst drop in the network.

3. BURST CLONING

In this section, we describe the details of the burst cloning technique proposed in [9] and the motivation behind the proposed Adaptive Burst Cloning mechanism. The original copy of a burst is referred to as the original burst, and the duplicated copy is referred to as the cloned burst. Similarly, the traffics corresponding to the original and cloned bursts are referred to as original and cloned traffics respectively. The node at which cloning is performed is referred to as the cloning node. The various factors to be considered in burst cloning include the number of cloned bursts for each original burst, the selection of the cloning node, and the selection of the routes for the original and the cloned bursts. In burst cloning, one or more cloned bursts can be made for each original burst. As explained in [9], on one hand, if more copies are made for a burst, the possibility of data loss for the burst is lower. On the other hand, if more copies are made, then more cloned traffic is added to the network. Cloned bursts may contend for network resources with original bursts, which may result in increasing loss for original bursts, which in turn may increase data loss instead of reducing it. Hence, traffic isolation mechanism is used in [9] in order to avoid the contention of the original bursts with the cloned burst. This priority-based pre-emptive burst scheduling hides the lowpriority cloned traffic from the high-priority original traffic thereby ensuring that the performance is at least as good with burst cloning as without it. However, if burst cloning is to be used as a technique for provisioning service-guaranteed connections, it may so happen that the cloned bursts get lost due to their contention with other original and cloned bursts. In our work, every connection is expected to have a minimum reliability. We use Burst cloning as a technique to ensure reliability. If the reliability of the original burst along its route is less, we clone the original burst at an intermediate node and ensure service guarantee. Clearly, in such a case, the cloned traffic (duplicated bursts for guaranteeing reliability) needs to be given the same priority as the original traffic. In [9], the authors have shown that in the presence of traffic isolation mechanism, the cloning technique exhibits maximum performance when source node happens to be the clone node. But in our work, having equal priorities for both the cloned and original traffics may result in increased contention and burst loss if cloning is done at the source node. Hence, in our work, we decide the clone node and the number of cloned bursts dynamically based on the current state of the network.

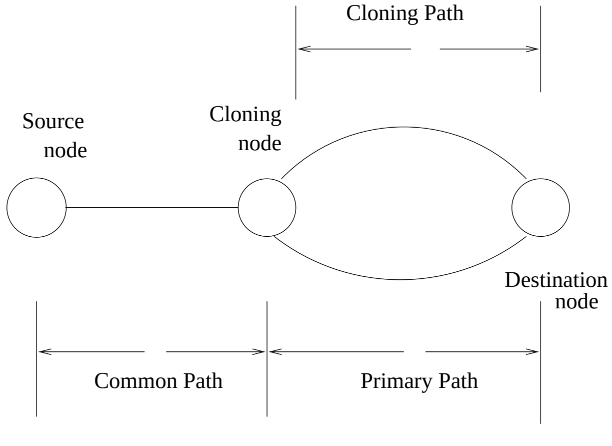

Figure 1: General Path Structure for Burst Cloning

The general path structure for burst cloning is shown in Fig. 1. The original burst is first sent along the common path. After the cloned copy is made at the cloning node, the original burst will then continue along the primary path while the cloned burst will be routed through the cloning path. As evident from the figure, the common path will be null if the source node happens to be the cloning node. In order to keep the loss of original bursts as low as possible, we choose the primary path to be on the shortest path from the source to the destination. Accordingly, the cloning node is on the shortest path between the node pair.

4. ADAPTIVE BURST CLONING

Our proposed Adaptive Burst Cloning mechanism differs from the conventional burst cloning mechanism in the following ways:

a. the selection of cloning node
b. the number of cloned bursts for each original burst, and
c. the routing for the original and cloned bursts.

4.1 The selection of cloning node

In the proposed Adaptive Burst Cloning technique, the cloning node is selected dynamically rather than fixing it at the source as it is done in the burst cloning mechanism proposed in [9]. For each source-destination pair, R-candidate routes (link disjoint routes) are pre-computed and are available at all source nodes. For scheduling a burst between a source-destination pair, the pre-computed shortest route is chosen. We then find the loss probabilities of all the links in that route using the link loss probability estimation mechanism presented in section 2. We consider a threshold loss probability value for deciding if a link is good enough to be scheduled a burst. Clone node selection is done dynamically by comparing the link loss probabilities of all the links in the selected route with the threshold loss probability value of the traffic classes. The route used for routing the original burst is called the Primary path and the part of the primary path that is common between the routes of the original and cloned bursts is referred to as the common path. The computation of the common path starts from the source node, it keeps adding links from the shortest path to the common path as long as the added links have loss probabilities less than the threshold. The common path is terminated as and when a link with loss probability greater than the threshold is encountered. Consequently, the node preceding this link is chosen as the clone node. As the loss probability of the links gets varied in accordance with the level of congestion in the links, choosing the clone node based on the loss probabilities yields high throughput. Moreover, since all the links along the common path are good enough to be scheduled (with loss probability less than the threshold), it is expected that the burst successfully reaches the end node of the common path (cloning node) without suffering a burst loss due to congestion at any intermediate node. As some of the links succeeding the clone node on the primary path may have loss probabilities greater than the threshold, it may cause burst loss due to congestion. However, if cloning is done at the clone node whereby the burst gets duplicated and sent along multiple routes, it may happen that at least one of the duplicate bursts or the original burst arrives at the destination successfully. Fixing the clone node always at the source [9] is resource consuming. In our approach, cloning is done only for that part of the route where there is likely to be a burst loss.

4.2 The number of cloned bursts for an original burst

In our proposed scheme, the number of duplicated bursts for cloning an original burst is dynamically decided based on the route loss probabilities of the routes between the selected clone node and the destination node. The route loss probability estimation mechanism is presented in the section 2. Similar to the notion of loss probability, we define success probability of a route as 1- loss probability of that route. Success probability of a route is the probability that the burst gets transmitted along the route successfully without contention. For cloning a burst, we need to choose minimum number of routes in order to minimize resource usage. However, it is also necessary to ensure high probability for at least one copy of the burst to reach the destination successfully. We choose the number of cloned bursts required for an original burst in such a way that the total(cumulative) probability of successful transmissions along the selected n-routes (for n cloned bursts) is closer to 1 for class0 service (high priority service) and 0.8 for class1 service respectively (low priority service). Dynamic selection of number of duplicated bursts not only ensures high probability of successful transmission but also significantly minimizes the resource utilization as compared to using fixed number of cloned bursts for all original bursts.

4.3 The routing for the original and cloned bursts

For routing the original and cloned bursts, we use link disjoint routes between each source destination pair in order to avoid resource-contention among the routes.

5. SERVICE DIFFERENTIATION

In our work, we consider three types of traffic classes namely class0, class1 and class2. We consider the threshold loss probabilities for class 0 and class 1 as 0.3 and 0.5 respectively. As Class2 is of best effort service, we do not associate any threshold loss probability for it. For routing class2 traffic bursts, we simply choose the best route (the one having the minimum loss probability value) among the R-candidate routes using Adaptive Routing and then schedule the burst along the chosen route. However for routing class0 and class1 traffic, after choosing the best route among the existing Rcandidate routes, we check if the loss probability of that route is within the threshold of the corresponding traffic class. If the loss probability of the selected route violates the threshold loss probability of the traffic class, the network enters into burst loss protection mechanism (i.e. Adaptive burst cloning loss recovery mechanism). In Adaptive Burst cloning protection mechanism, the threshold route loss probability is taken as the threshold loss probability of the corresponding traffic class. This ensures that Burst cloning is done along the chosen route if and when a link exceeds the threshold loss probability. Further service differentiation is ensured by choosing the number of cloned bursts based on the traffic class QoS requirements (i.e the cumulative success probability of the cloned bursts is 1 for Class0 and 0.8 for Class1) as described in section 4.2. By separating the existing routes into least congested routes, average congested routes and high congested routes, our technique ensures service differentiation by routing the high priority bursts along least congested routes, low priority bursts along high congested routes and intermediate priority bursts along average routes. Loss probability of a route reflects the level of congestion along that route, if loss probability of a route is high, congestion is more on that route. By using loss probability as a metric for service differentiation, our approach achieves better service variation among the traffic classes. For routing the bursts, we use Adaptive routing loss minimization mechanism that selects the best route among the existing routes between the given source-destination pair. As a result, burst drop is significantly minimized. For protecting the bursts that do not satisfy the connection service requirements (i.e. if all route loss probabilities violate the threshold), the network enters into Adaptive Burst Cloning loss recovery mechanism in order to reduce data loss. Thus, our approach not only ensures service differentiation among the traffic classes but also greatly improves the reliability of the OBS network.

6. DISCUSSION OF THE RESULTS

In this section, we discuss the experimental results obtained by simulating the proposed DRRDC QoS scheme in an OBS network. We compare our proposed QoS scheme with the Priority burst scheduling QoS scheme. We use NSF network topology as our test network and assume that the burst length is fixed and is equivalent to 4 seconds, containing 50,000 bytes. We also assume that all core nodes are buffer less (no FDLs) and have no-wavelength conversion capability. We adopt the first available unscheduled channel (FAUC) algorithm to schedule data bursts at the core nodes.

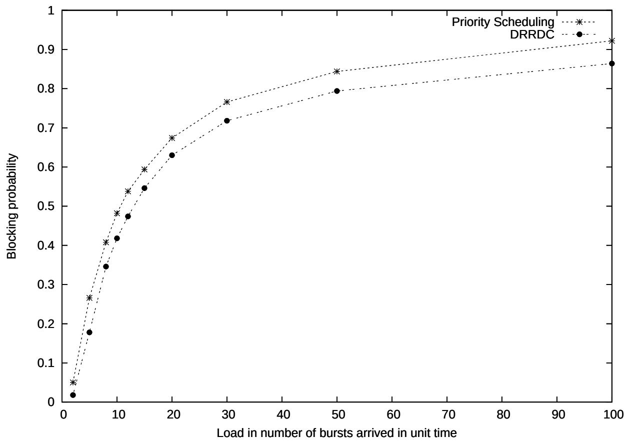

Figure 2: Load versus Burst blocking probability of total burst drop when TX=3, RX=3 and WL=5

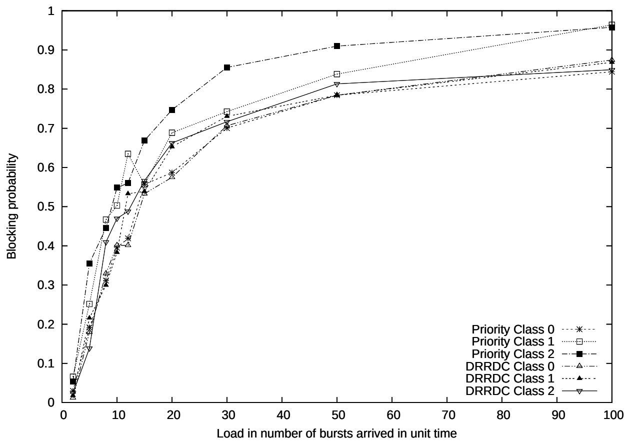

Figure 3: Load versus Burst blocking probability of different traffic classes when TX=3, RX=3 and WL=5

The conditions followed in our simulations are

- Random call distribution in network,
- Blocking probability includes source and destination busy conditions.
- Shortest Path Routing Algorithm is applied using the number of hops as the metric.
- Links in the network are bidirectional, if there exists a fiber between nodes a and b, there also exists a fiber between nodes b and a.
- The threshold loss probability value for class0 traffic is considered as 0.3 and threshold loss probability of class1 traffic is considered as 0.5.
- Burst arrivals follow a Poisson process and connection requests are randomly generated among the source-destination nodes.

The load value in each plot is the number of bursts arrived into the entire network per unit time. Figure 2 plots the network load versus the total burst blocking probability (including all traffic classes) for the DRRDC and Priority burst scheduling QoS schemes. Here, the number of transmitters (Tx) and receivers (Rx) at each node is taken as 3 and number of wavelengths on each link WL is assumed to be 5. Figure 3 plots the network load versus burst blocking probability of the individual traffic classes: class0, class1, and class2 for the same experimental conditions used in figure 2. In both figures 2 and 3, we find that the blocking probability for the traffic classes increases with increase in the network load. A closer observation reveals that the blocking probabilities for the DRRDC and priority-based schemes do not show a wide difference under low-load conditions. However, as the load increases beyond a certain value, we find that the proposed DRRDC scheme comprehensively outsmarts the priority-based burst scheduling scheme. The lack of significant difference at low-load conditions is attributed to the availability of unused network resources to satisfy the low-load demands. However at high load, resource-availability becomes scarce and hence it results in increasing burst drop in the network. DRRDC tries to identify those parts of the network where there is possibility of burst drop and tries to clone(duplicate) the bursts along such paths thereby ensuring better service guarantees than the priority-based scheme.

Figures 4 and 5 plot the network load versus Burst Blocking Probability by keeping Tx=3, Rx=3 and Wl=3. The observations are very similar to the figures 2 and 3.

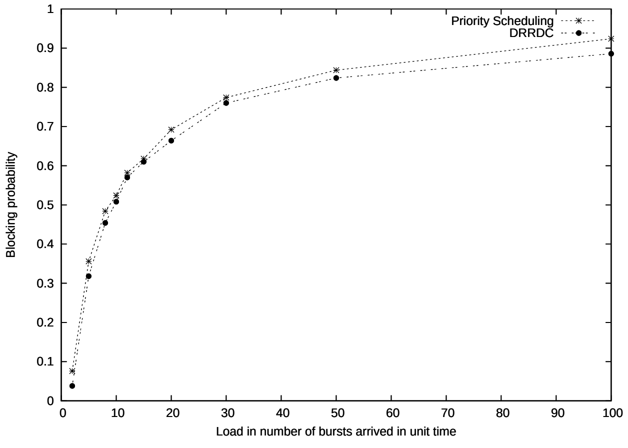

Figure 4: Load versus Burst blocking probability of total burst drop when TX=3, RX=3 and WL=3

Fig. 6 plots total Burst Blocking probability of the traffic classes as a function of the number of wavelengths available

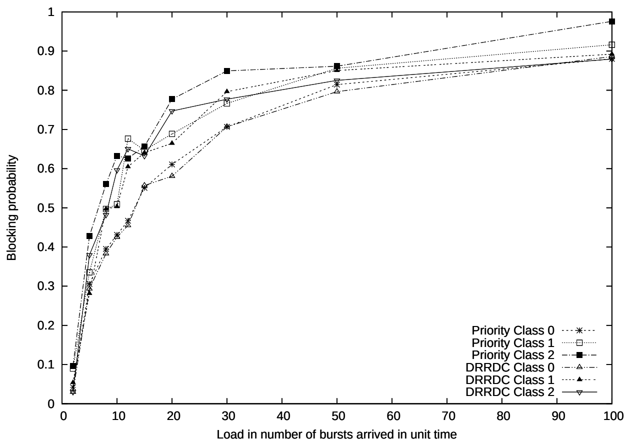

Figure 5: Load vs Burst blocking probability of different traffic classes when TX=3, RX=3 and WL=3

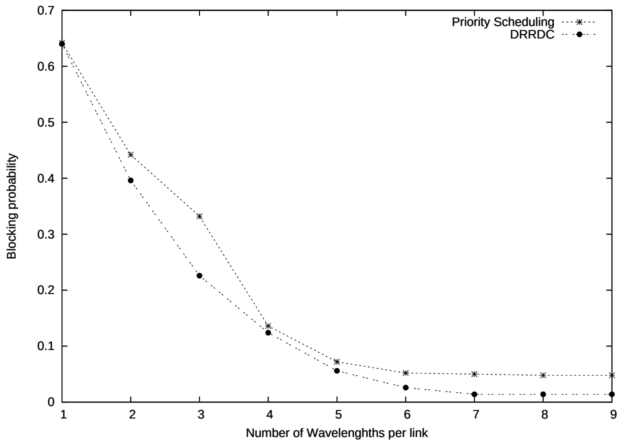

Figure 6: Number of Wavelengths per link versus Burst Blocking Probability of total burst drop when TX=5, RX=5 and WL=5

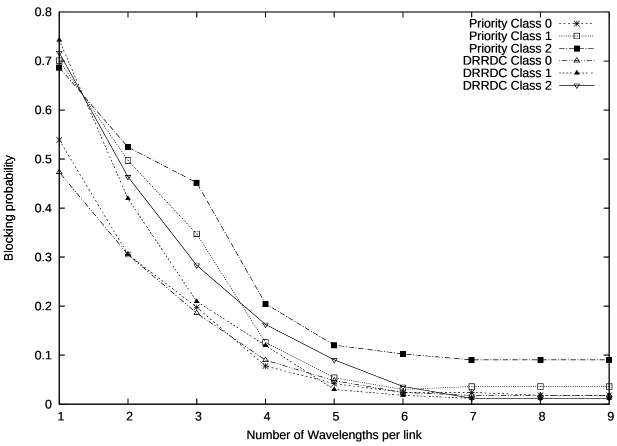

Figure 7: Number of Wavelengths per link versus Burst Blocking Probability of different traffic classes when TX=5, RX=5 and WL=5

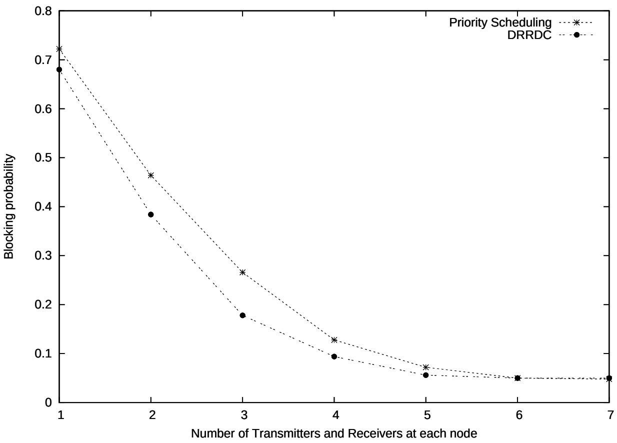

Figure 8: Number of Transmitters and Receivers at each node versus Burst Blocking Probability of total drop when WL=5

in each fiber. The load (i.e. burst arrival rate) is kept constant with Poisson arrival rate $\lambda = 5$. Fig. 7 plots individual burst blocking probability of the traffic classes as a function of wavelengths under constant network load ($\lambda = 5$). In these figures, the number of transmitters (Tx) and receivers (Rx) is assumed as 5. Here again, DRRDC is found to have low burst loss. Here, we note that the burst blocking probability of the traffic classes decreases with increase in the number of wavelengths present in the fiber. As more wavelengths are made available, more number of bursts could be successively transferred through the wavelength links.

Fig. 8 plots the number of Transmitters and Receivers on each link versus the total Burst Blocking probability of the traffic classes when the load (i.e. burst arrival rate) λ is 5 and number of wavelengths per link WL is 5.

Fig. 9 plots the number of Transmitters and Receivers available at each node versus Burst Blocking probability of the individual traffic classes under the same experimental conditions. As observed from the graphs, the blocking probability is significantly low for the proposed DRRDC scheme. Another interesting observation in our simulations is that even for class 2 (best effort service) traffic, we find a significant performance improvement with DRRDC. It is due to the underlying adaptive routing strategy used by DRRDC. As adaptive routing always chooses the least congested path, it reduces the burst loss and thereby outsmarts the prioritized burst-scheduling technique that has no knowledge of the burst loss along the candidate routes.

7. CONCLUSIONS

In this paper, the issue of QoS support in OBS networks is addressed. Two mechanisms namely, Adaptive Routing and Adaptive Burst Cloning are proposed and integration of these two mechanisms has been used to ensure service differentiation in OBS networks. The integrated Adaptive Routing loss minimization mechanism with Adaptive Burst Cloning loss recovery mechanism has showed the best performance. The proposed Dynamic Routing of Reliability Differentiated Connections (DRRDC) QoS technique further reduces the loss probability experienced by the no guaran-

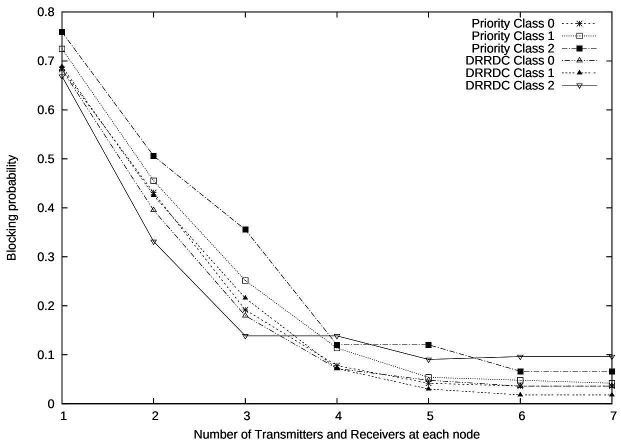

Figure 9: Number of Transmitters and Receivers at each node versus Burst Blocking Probability of different traffic classes when WL=5

teed traffic while satisfying the loss requirement of the guaranteed traffic, thereby improving the networkwide loss performance.

8. REFERENCES

[1] J.Y. Wei, “Advances in the Management and Control of the Optical Internet," *Journal of Selected Areas in Communications*, vol. 20, no. 4, pp. 768-785, May 2002.

[2] B. Mukherjee, “WDM Optical Communication Networks: Progress and Challenges," *IEEE Journal on Selected Areas in Communications*, pp. 1810-1823, Oct. 2000.

[3] N. Gnhani, S. Dixit and T. Wang, “On IP-over- WDM integration,” *IEEE Communication Magazine* pp. 72-84, March 2000.

[4] S. Yao, B. Mukherjee, and S. Dixit, “Advances in photonic packet switching: An overview," *IEEE Communications Magazine*, vol. 38, no. 2, pp. 84-94, February 2000.

[5] S. Yao, S. J. B. Yoo, B. Mukherjee, and S. Dixit,“All-optical packet switching for metropolitan area networks: Opportunities and challenges,"*IEEE Communications Magazine*, vol. 39, pp. 142-148, March 2001.

[6] C. Qiao and M. Yoo, “Optical burst switching (OBS) - A new paradigm for an optical Internet," *Journal of High Speed Networks*, vol. 8, pp. 69-84, 1999.

[7] L. Liu, P. J. Wan, and O. Frieder, “Optical burst switching: the next IT revolution worth multiple billions dollars?" *MILCOM 2000. 21st Century Military Communications Conference Proceedings*, Vol. 2, pp. 881 - 885 vol.2, 22- 25 Oct. 2000.

[8] T. Battestilli and H. Perros,“An introduction to optical burst switching," *IEEE Communications Magazine*, Vol. 41, Issue. 8, pp. S10 - S15, Aug. 2003.

[9] V. M. Vokkarane and J. P. Jue, “Prioritized routing and burst segmentation for QoS in optical burst-switched networks," *Proceedings, Optical Fiber Communication Conference (OFC)*, vol. WG6, pp. 221 - 222, March 2002.

[10] Xiaodong Huang, Vinod M. Vokkarane , and Jason P. Jue, “Burst Cloning: A Proactive Scheme to Reduce Data Loss in Optical Burst-Switched Networks," *in proceedings, IEEE International Conference on Communications (ICC)*, vol. 0, pp. 1673-1677, May 2005,

[11] J. Praveen, B. Praveen, and C. Siva Ram Murthy, “First Step Towards Autonomic Optical Burst Switched Networks," *in Proc. The Second International Conference on Autonomic Computing (ICACŠ 05)*, Seattle, Washington, June 2005.