New algorithms for QoS performance improvement in high speed networks

RADU DOBRESCU, DANIELA HOSSU, STEFAN MOCANU, MAXIMILIAN NICOLAE
"Politehnica" University of Bucharest, Faculty of Control and Computers,
313 Spl. Independentei, Bucharest
ROMANIA
radud@isis.pub.ro

Abstract. In this paper is addressed the issue how to provide QoS control and improve performance in high speed Optical Burst Switching (OBS) networks. In order to provide better proportional differentiated services are proposed new algorithms: a burst assembly algorithm (named adaptive timer-based) with traffic shaping functions and a Delayed Burst Assignment (DBA) scheme, both able to reduce the variance of assembled traffic and to improve the burst loss performance. A precise analytical model to study the performance of the assembled traffic in high speed networks is also presented. The simulation results show that the proposed algorithms have better performance in terms of edge buffering delay and can lighten burst loss ratio for TCP flows of the OBS network.

Keywords: Optical Burst Switching networks, QoS control, deterministic and statistical services, integrated and differentiated services, traffic burstiness, burst loss performance

1 Introduction

The new broadband high speed networks are designed to integrate a wide range of audio, video and data traffic within the same network. In these networks, the stringent Quality of Service (QoS) requirements of real-time multimedia applications are met by the mean of special services providing guarantees. One distinguishes between two main kinds of such services: deterministic and statistical [1].

The deterministic service, also called Guaranteed Service within the Internet framework, is aimed to provide strict guarantees on the QoS requirements. The mathematical basis of the deterministic service is a network calculus which allows obtaining deterministic bounds on delay and buffering requirements in a communication network. This model provides bounds only on the maximal delay that traffic can undergo through the different elements in the network.

The statistical service is intended to provide only statistical (probabilistic) guarantees. For services offering statistical guarantees, one is rather interested in bounding the mean delay or the probability that the delay exceeds a certain value. The main advantage of the statistical service over the deterministic service is that it can achieve higher network utilization, at the expense of some minor quality degradation.

Two notable models of statistical services are used to meet the demand for QoS: Integrated Services (IntServ) and Differentiated Services (DiffServ). The IntServ model is characterized by resource reservation; before data is transmitted, applications must set up paths and reserve resources along the path. IntServ aims to support applications with different levels of QoS within the TCP/IP (Transport Control Protocol/Internet Protocol) architecture. IntServ however, requires the core routers to remember the state of a large number of connections giving rise to scalability issues in the core of the network. It is therefore suitable at the edge network where the number of connections is limited. The DiffServ model is currently being standardized to provide service guarantees to aggregate traffic instead of individual connections. The model does not require significant changes to the existing Internet infrastructure or protocol. DiffServ does not suffer from scalability issues, and hence is suitable at the core of the network. It is therefore believed that a significant part of the next generation high speed networks will consist of IntServ at the edge and DiffServ at the core of the network.

Unfortunately, all of these algorithms have poor burst loss performances, even at low traffic loads, because the real traffic is bursty. To solve this problem, we propose in this paper a novel burst assembly algorithm with traffic smoothing functions (the traffic is smoothed by restricting the burst length to a threshold). Compared with existing algorithms, our scheme can improve network performance in terms of the burst loss ratio. The simulation results show that our proposed scheme can lighten burst loss ratio for TCP flows of a high speed network.

2 Technologies for high speed networks

In high speed networks all communications are limited by the electronic processing capabilities of the system. Although hardware-based high-speed electronic IP routers with capacity up to a few hundred gigabits per second are available now, there is still a serious mismatch between the transmission capacity of WDM (Wavelength Division Multiplexing) optical fibers and the switching capacity of electronic IP routers. With IP traffic as the dominant traffic in the networks, the traditional layered network architecture is no longer adapted to the evolution of the Internet. In the multi-layered architecture, each layer may limit the scalability of the entire network, as well as adding the cost of the entire network. As the capabilities of both routers and OXCs (optical cross-connects) grow rapidly, the high data rates of optical transport suggest bypassing the SONET/SDH and ATM layers and moving their necessary functions to other layers. This results in a simpler, more cost-efficient network that can transport very large volumes of traffic. IP over WDM is considered as a promising solution for the next generation network since it has no intermediate layer so that it can void the functionality redundancy of the ATM and SONET/SDH layers. On the other hand the processing of IP packets in the optical domain is still not practical yet, and the optical router control system is implemented electronically. Nowadays, we are mostly studying the semi-transparent optical transport networks. In optical transport networks, the control messages are processed electronically, and the data are propagated in the high-speed transparent data channels. To realize an IP-over-DWDM architecture, several approaches, such as Wavelength Routing (WR) [2], Optical Packet Switching (OPS) [3] and Optical Burst Switching (OBS) [4], have been proposed. Of all these approaches, optical burst switching (OBS) can achieve a good balance between the coarse-gained wavelength routing and fine-gained optical packet switching, thereby combining others' benefits while avoiding their shortcomings.

2.1. Characteristics of OBS networks

In an OBS network (shown in Fig. 1), the edge routers and core routers connect with each other with WDM links. The edge nodes are responsible for assembling packets into bursts and de-assembling bursts into packets. The core nodes are responsible for routing and scheduling based on the burst header packets.

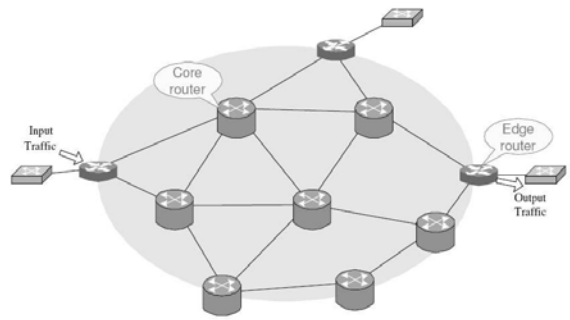

Fig. 1. OBS network model

The architecture of edge router aims to eliminate O/E/O (Optical to Electronic to Optical) conversions and electronic processing loads, which are the bottlenecks of an OBS network. In Du's approach [5] the ingress part of an edge node assembles multiple IP packets with the same egress address into a switching granularity called a burst. A burst consists of a Burst Header Packet (BHP) and a Data Burst (DB). The BHP is delivered on a control channel; its corresponding DB is delivered on a data channel without waiting for a confirmation of a successful reservation. A channel may consist of one wavelength or a portion of a wavelength, in case of time-division or code-division multiplexing. When a BHP arrives at a core node, the core node converts it into an electronic signal, performs routing and configures the optical switching according to the information carried by the BHP. The DB remains in the optical domain without O/E/O conversion when it cuts through the core node.

2.2. Burst Assembly Schemes

Burst assembly at the edge router is an important issue for OBS networks. Basically, there are two assembly schemes: *threshold-based* and *timer-based*.

In a threshold-based scheme, a burst is created and sent into the optical network when the total size of the packets in the queue reaches a threshold value

L_b. The shortcoming of the threshold-based scheme is that it does not provide any guarantee on the assembly delay that packets will experience.

In a timer-based scheme, a timer is started at the initialization of the assembly. A burst containing all the packets in the buffer is generated when the timer reaches the burst assembly period T_b. A large time-out value T_b results in a large packet buffering delay at the edge node. On the other hand, a too small T_b results in too many small bursts and a high electronic processing load.

The choice of burst assembly algorithms depends on the type of traffic being transmitted. Timer-based algorithms are suitable for time-constrained traffic such as real-time applications because the upper bound of the burst assembly delay is limited. For a time-insensitive application such as file transmission, to reduce the overhead of control packets and increase OBS transmission efficiency, a threshold-based scheme may be more appropriate. In fig.2 is shown the effect of load on timer-based and threshold-based assembly schemes).

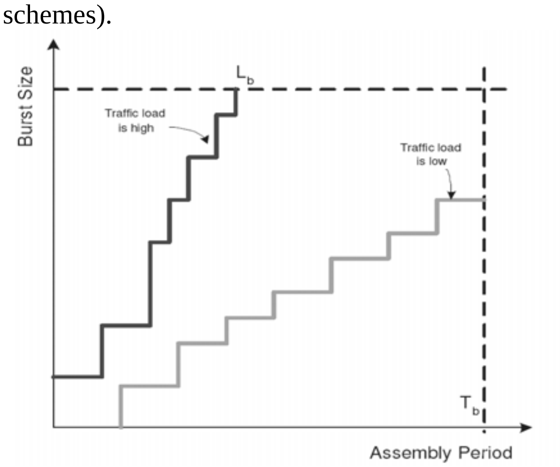

Fig. 2. Load effect in assembly schemes

How to choose the appropriate time-out or threshold value for creating a burst is still an open issue. A smaller assembly granularity leads to a higher number of bursts and a higher number of contentions, but the average number of packets lost per contention is less. Also, it will increase the number of control packets. If the reconfiguration period of optical switching fabric is non-negligible, a smaller assembly granularity will lead to lower network utilization because each switched burst needs a reconfiguration period. On the other hand, a higher assembly granularity will lead to a higher burst assembly delay and the average number of packets lost per contention is larger.

A good solution seems to be a mix between the two basic assembly schemes in order to obtain a hybrid one. For a *hybrid* algorithm, a burst will be generated when either the timer exceeds T_b or the burst size reaches L_b. Chen et al. [6] proposed an adaptive assembly (hybrid) algorithm to match with the TCP congestion control mechanisms that yield a good performance in terms of data loss ratio.

3 Traffic-smoothing using Burst Assembly Schemes

3.1. Analysis of Assembled Traffic

A challenging issue in OBS is how to assemble IP packets into bursts at ingress nodes. As we assume there is no buffer in OBS networks, a burst loss event will occur if multiple bursts from different input ports are destined for the same output port at the same time. The burst arrival process is determined by the traffic characteristics such as the burst inter-arrival time and the burst length distributions, which are dependent on the burst assembly strategy. For the timer-based assembly algorithm, the burst size is unlimited when too many packets arrive suddenly. A larger burst is more likely to be blocked or to block other bursts during transmission. For the threshold-based and hybrid assembly algorithms, a large number of bursts will be generated and injected into the network in a short period when many packets arrive.

The network traffic characteristics have attracted considerable research attention (a representative synthesis is [7]), because the traffic properties greatly influence the network's performance. Let consider in a simple traffic model for a burst assembler $\{A(t), t > 0\}$ be the cumulative amount of the input traffic in bits of the burst assembly function arriving during time interval (0, t], and $\{A^{OBS}(t), t > 0\}$ be the cumulative amount of output traffic of the burst assembly function during time interval (0, t], that is, assembled traffic.

We define $X_n (n \in N)$ as a sampled process of input traffic with the unit time of τ. Therefore, $X_n = \{A(n\tau) - A((n-1)\tau), n \in N\}$.

We assume X_n to be a wide-sense stationary discrete stochastic process, with constant mean $\mu = E[X_n]$. Let $S_{n,m}$ be the sum of m consecutive numbers from X_n; then $S_{n,m} \equiv \sum_{j=(n-1)m}^{nm} X_j$

We define $X_n^{(m)} \equiv \frac{S_{n,m}}{m}$ and use the variance-time function $Var[X_n^{(m)}] = E[(X_n^{(m)} - \mu)^2]$ (1)

to represent the variance of input traffic. Here, $E[X_n^{(m)}] = E[X_n] = \mu$

Similarly, we also define $Y_n(n \in N)$ as a sampled process of assembled traffic with the unit time interval of τ and assume it to be a wide-sense stationary discrete stochastic process. Here, $Y_n = \{A^{OBS}(n\tau) - A^{OBS}((n-1)\tau), n \in N\}$. Then, the variance of assembled traffic can be measured by the variance-time function

$$Var[Y_n^{(m)}] = E[(Y_n^{(m)} - \mu^{OBS})^2] \qquad (2)$$

where $Y_n^{(m)}$ is the mean value of m consecutive numbers from Y_n, and μ^{OBS} is the mean of $Y_n^{(m)}$ and can be expressed by $E[Y_n^{(m)}] = E[Y_n] = \mu^{OBS}$.

In a landmark paper [8] from 1993, Leland at al. report the discovery of self-similarity in local area network (LAN) traffic, more precisely Ethernet traffic. Self-similarity is a characteristic of traffic in long timescales. The variance of the traffic decreases more slowly, the higher self-similarity of the traffic. The degree of self-similarity can be described by the Hurst parameter H. A larger H indicates a higher self-similarity. In this paper, we focus on the discussion of "burstiness", which is defined as the variance of traffic in bit rate in small timescales. Although the characteristics of assembled traffic have already been widely studied, there is no agreement on how the burst assembly affects the traffic. Mostly the input traffic was generated according to a Pareto distribution or to a Poisson distribution and also it was demonstrated that the variance of the assembled traffic decreased after the assembly.

In this paper, we analyze the assembled traffic through a model based on Fractional Brownian Motion (FBM) proposed by the authors in a previous work [9]. The FBM model is defined by

$$A(t) = \lambda t + \sqrt{\lambda a} Z_H(t); T \in R \qquad (3)$$

where λ is the arrival rate for the packets, a is the variance of the coefficient for the arrival packet, and $Z_H(t)$ is the normalized FBM with zero mean and variance of $Var[Z_H(t)] = |t|^{2H}$, where H is the Hurst parameter and satisfies $H \in [0{:}5; 1)$.

In our simulations we have considered for large timescales ($m > 200$) that the variance-time curve is approximated using the FBM model for the parameter set λ=564 Mb/s, a =2.5x10^6, and H=0,85. In small timescales ($m < 200$) the traffic can be approximated by an FBM model with the parameter set =564 Mb/s, a =9x10^5, and H=0,75.

Since the variance of the aggregate of uncorrelated traffic will equal the sum of the individual source's variance, we only have analyzed the variance of the assembled traffic of one burst source, using a timer-based burst assembly scheme. With a timer-based burst assembly algorithm, all packets in the assembly buffer will be sent out as a burst when the timer reaches the burst assembly period T_b. After a burst is generated, the burst is buffered at the edge node for an offset time before being transmitted to give its BHP enough time to reserve wavelengths along its route. During this offset time period, packets belonging to that queue will continue to arrive. Because the BHP that contains the burst length information has already been sent out, these arriving packets could not be included in the generated burst. When a burst is generated at one buffer, the future arriving packets will be stored at another buffer until the next assembly cycle. We assume that the inter-arrival time of the bursts from thesame burst source is fixed as T_b. Accordingly, there will be no packets left in the assembly buffer at time $kT_b(k \in N)$. Therefore, $A^{OBS}kT_b = A(kT_b)$.

We denote $Q(t)$ as the number of bits that are buffered at the edge router. So $A^{OBS}(t) = A(t) - Q(t)$.

For the timer-based assembly algorithm, the $Q(t)$ bits are at most the packets that arrive during [0, s], where $s \in [0,T_b]$. For simplification, we assume s is uniformly distributed in $[0,T_b]$ and $Q(s)$ is a Gaussian process with mean λ_s and variance $\lambda a s^{2H}$.

We denote $\overline{E_Q}$ as the mean value of $Q(t)$ observed at any sample point. Then, $\overline{E_Q} = \lim_{T\to\infty} \frac{1}{T}\int_0^T Q(t)dt$.

The difference between the original traffic and the assembled traffic can be denoted by

$$\Delta V(t) = Var[A^{OBS}(t)] - Var[A(t)] \qquad (4)$$

To describe the variance of assembled traffic $A^{OBS}(t)$, the variance-time function defined in Eq.(2) becomes

$$Var[Y_n^{(m)}] = E[\frac{A(m\tau)}{m}] + \frac{1}{m^2}\Delta V(m\tau) \qquad (5)$$

From Eq.(5), we can see there is an increase in the variance in short timescales. This indicates that the timer-based assembly algorithm will increase the variance of traffic. As the timescale increases, the difference between the original and assembled traffic becomes negligible because the traffic does not change significantly for large timescales.

3.2. Description of the proposed Burst Assembly Algorithm

The simulation results on the model described in Section 3.1 show that the timer-based assembly algorithm could not reduce the variance of the real traffic, but do increase it in small timescales. Defining the traffic burstiness as the variance of the bitrate in small timescales, one can observe that a larger burstiness indicates that the traffic is burstier and more likely to exceed the capacity of the network and it results in burst loss events. So, a larger burstiness implies a higher burst loss ratio in bufferless OBS networks. One way to reduce the burst loss ratio is to control the burst sources at the edge nodes and thereby inject the bursts more smoothly into the network. In this section is discussed a novel burst assembly algorithm with traffic smoothing functions, to reduce the burstiness of the assembled traffic.

A simple way to reduce the burstiness is a peak rate restriction. Conceptually, the number of bursts simultaneously arriving at an input port is most likely to reach a maximum value when the traffic is at a peak rate. Reducing the number of overlapping bursts on a link is for each ingress node to restrict the assembled traffic to a specified rate. In a timer-based assembly scheme, because the bursts are generated periodically, the traffic rate can be restricted by restricting the burst length to a threshold. Based on this idea, we propose a scheme, called *adaptive timer-based,* to reduce the burstiness of traffic.

For this scheme we suppose each edge router has G queues to sort the arriving packets. Let the timer of queue $Q[i]$ be denoted by $T[i]$ and the length of $Q[i]$ be denoted by $L[i]$. The threshold for generating a burst is $L_{th}[i]$. When the value of the queue length $L[i]$ is smaller than $L_{th}[i]$, all packets in $Q[i]$ will be assembled into a burst. Otherwise, a burst is generated with the length of $L_{th}[i]$ and the other packets are left in $Q[i]$.

The scheme is thus implemented using the following algorithm:

Step 1. When a packet with a length of b arrives at $Q[i]$, then if $Q[i]$ is empty, start timer $T[i]$, $L[i]= b;$ else push packet into $Q[i]$, with $L[i] = L[i] + b$

Step 2. When $T[i] = T_b$ if $(L[i] > L_{th}[i])$, $L_b= L_{th}[i]$, $L[i] = L[i]-L_{th}[i]$, $T[i]=0$ and restart timer $T[i]$; else $L_b = L[i]$, $L[i] = 0$, $T[i]=0$ and stop timer $T[i]$.

Step 3. Generate a burst with length L_b and send it into the OBS network

Figure 3 shows a comparison of the timer-based (a) and adaptive timer-based (b) assembly algorithms. For the timer-based assembly algorithm, all packets in the assembly buffer will be multiplexed to a burst every assembly period. This makes for various burst sizes because the number of packets that arrive in each assembly period varies. On the other hand, the burst size in the adaptive timer-based assembly algorithm is restricted by the threshold $L_{th}[i]$. After an $L_{th}[i]$ length of packets are assembled into a burst, the other packets will be left in the assembly queue for a future assembly process. So, the adaptive timer-based scheme can avoid a sudden increase in the burst size and makes the burst sent out more smooth than the timer-based assembly algorithm does.

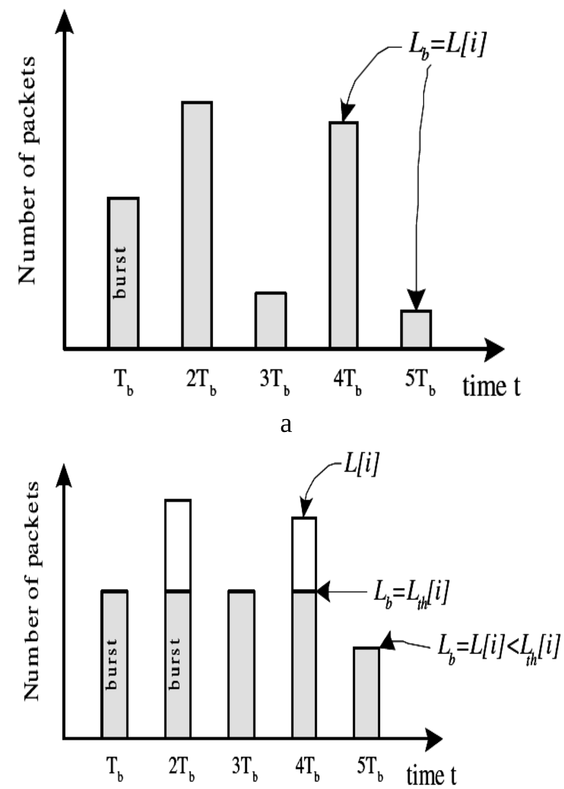

Fig. 3. Comparison of timer-based and adaptive timer-based algorithms

Take into consideration the choice of $L_{th}[i]$ to take advantage of the effects of the peak rate restriction. It is clear that the restriction of on the peak rate should be bigger than the average rate and the threshold $L_{th}[i]$ must exceed the average burst length $\overline{L_M}(i)$, otherwise the traffic would be blocked at the edge nodes. As $L_{th}[i]$ is close to $\overline{L_M}(i)$ ($\alpha \to 0$), the transmission is almost the same as in a CBR transmission. However, this will result in enormous backlogs at the edge routers. The choice of α is a tradeoff between the effects of the peak rate restriction and the edge buffering delay. When too many packets suddenly arrive, if we still assemble packets into bursts using a small threshold $L_{th}[i]$, the packets will suffer a large edge buffering delay.

4 PROPORTIONAL DIFFSERV SCHEME

4.1. QoS DiffServ Models in OBS

The differentiated services enable Internet users to provide diverse quality of service (QoS). Two categories of methods for realizing differentiated services have been proposed [10]: one is non-proportional method and another is proportional method. In the non-proportional differentiated services model using an extra-offset-time-based QoS scheme, different offset times are assigned to different priority classes without any buffer in the WDM layer. By providing a larger offset time, a higher priority burst is more likely to have wavelength reserved for it because of its early reservation. In OBS, a BHP is processed and sent to the next hop without waiting for the arrival of its DB. After that, if its DB is truncated or dropped, the BHP is unaware of these changes and cannot update its carried information (e.g., burst length). Any attempt to preempt the reserved resources of lower priority classes by higher priority classes is therefore awkward and inefficient. However, the difference of the burst loss ratios of each class is unstable because it depends on the traffic load. Although we can change the extra offset times difference to modify the difference of burst loss ratios, a quantitative solution cannot be found in the approach. To the contrary we call the proportional differentiated model in an OBS network if the burst loss ratio of one service class is proportional to those of other classes regardless of the traffic load. Hence, in the proportional differentiated model, the burst loss ratio of one class is "predictable" if we know that of another class and is also "controllable" because the network provider can adjust the class differentiation parameters to adjust the burst loss ratios of each class [11]. It can be expected that introducing a proportional differentiated model into an OBS network would be favorable to both network operators and users.

4.2. Novel DBA Algorithm for QoS DiffServ

Delayed Burst Assignment (DBA) is a service differentiation technique in which bursts of lower priority are processed after a delay to guarantee that bursts of higher priority are more likely to have wavelengths reserved for them. A service differentiation technique of this kind was proposed in [12]. It provides differentiated services by maintaining a BHP queue for each class and ensures that BHPs of higher priority are processed before BHPs of lower priority. We define BHP queueing delay as the period a BHP spent waiting in the BHP queue. With this algorithm, the BHP queueing delay for the lower priority class is uncontrollable when the higher priority traffic is heavy, and the scheme might deteriorate into a classless one when the traffic load is low (no BHP in queue when the inter-arrival time of BHPs is much longer than the processing time of a BHP). Our new proposed DBA scheme is simpler and cheaper, because no optical buffer but electronic buffer is needed.

DBA divides BHPs into two types: type 1 and type 2. Type 1 has priority over type 2. The bursts of both types have the same offset time. The scheme works as follows:

1. When a BHP of type 1 arrives, it is processed normally and is sent to the next node. When a BHP of type 2 arrives, it is queued in a BHP queue for a waiting time period T_{wait}. Because BHPs are processed electronically, we can delay them in the Random Access Memory (RAM). During this waiting period, the BHPs of type 1 are processed, resulting in a reducing burst loss ratio for this type.

2. When T_{wait} has passed, the BHPs of type 2 are processed and the wavelengths that have not been reserved by type 1 bursts are reserved. T_{wait} should be included in the extra offset time. If it is not included, the BHP's residual offset time will be less than the processing time of the BHP for its remaining route after it has been buffered at intermediate nodes. In this case, the corresponding DB will be dropped by the core node because the BHP could not be processed before the DB's arrival. When T_{wait} is close to zero, the whole system will deteriorate into a classless one because there is not enough time for processing type 1 BHPs. A large T_{wait}, on the other hand, will cause a large extra offset time and a large end-to-end delay. As we assume there is no optical buffer in core nodes, the end-to-end delay is the sum between end-to-end propagation delay and the end-to-end BHP queueing delay. Here, the last term is defined as the sum of the BHP queueing delays at core nodes during transmission, which should be included in the extra offset time. Let N_{max} be the maximum hop number and T_{exoff} be the maximum required extra offset time in the OBS network. Then $T_{maxoff} = T_{wait} \times N_{max}$ (i.e., the worst case whereby lower priority BHPs buffered for T_{wait} at each core node). To prevent the end-to-end delay from becoming too large, we introduce the parameter N_r ($N_r \leq N_{max}$), the maximum number of times for a BHP to be queued during transmission. When the number of times a BHP has been queued

reaches N_r, the burst will be dropped immediately if it can not find a suitable wavelength in its candidate wavelength set. The required extra offset time is thus limited to: $T_{maxoff} = T_{wait} \times N_r$.

4.3. Absolute QoS DiffServ model

In the previous section we have introduced a scheme for a *relative QoS* model, in which the QoS of one class is defined relatively in comparison to other classes (i.e., a higher priority burst is guaranteed to experience lower loss probability than a lower priority burst). However, no upper bound on the loss probability is guaranteed for the higher priority burst. There are many types of traffic that require strict QoS guarantees (for example, a data transfer operation cannot bear packet loss ratio exceeding a certain threshold). The *absolute QoS* model provides a worst-case QoS guarantee to applications. This kind of hard guarantee is essential to support applications with delay and loss ratio constraints, such as multimedia applications. Zhang et al. [13] proposed an early dropping scheme to drop lower priority bursts to assure higher priority bursts have more probability in reserving wavelength to meet the absolute constraint of higher priority class. However, this absolute QoS model makes no differentiation among the classes when the traffic load is low.

Although it has been accepted that proportional differentiated services with absolute constraints is important [14], [15], there is no scheme in the literature to provide proportional differentiated services with absolute constraints in OBS. Our model defines a system that supports both absolute constraints and proportional constraints, assuring that absolute QoS constraints have higher priority over proportional QoS constraints. When there are conflicts between constraints, the constraints with lower priorities will be relaxed. Besides the parameters already introduced, we add a new parameter Pb_{imax}, the maximum burst loss ratio at each node for class *i*. The scheme is thus realized as the following algorithm:

Step 1. When a burst of class *i* arrives, use DBA algorithm to schedule burst. If schedule is successful, forward it to next hop and create a new entry with the $(N + i)$-th bit set 1; else discard burst and create a new entry with the *i*-th and the $(N + i)$-th bit set 1

Step 2. Push the new entry into the FIFO and pop the oldest entry and compute Pb_i. Then if $Pb_{i-1} > Pb_{max}$ put $a_i = a_{i-1}$; else put $a_i = a_{i+1}$ where a_i denotes the wavelength number.

In the following, we define an integrated scheme for proportional differentiated services using the DBA algorithm presented in section 3.2 and the *absolute QoS* model presented in section 4.2 to support joint QoS. The core node first use DBA scheme to process BHP and assign wavelength for its burst. If the arriving BHP could not find an available wavelength for its burst, the BHP will be buffered for T_{wait}. So it has an opportunity to reschedule its burst to the wavelengths in its rescheduling wavelength set in which the wavelengths have been assigned to the higher priority class but haven't been reserved yet. The rescheduling wavelength set is dynamically adjusted by the following algorithm in which the absolute constraints have priority over proportional constraints:

Step 1. A BHP arrives and we test if the DBA scheme schedules its burst successfully. If yes, then reservation is done; otherwise get to step 2.

Step 2. Test if the residual offset time is large enough. If yes, BHP is put in queue for T_{wait}, and then we get to step 2; otherwise reservation failed.

Step 3. Test if the rescheduling of the burst on its rescheduling wavelength set was successfully made. If yes, then reservation is done; otherwise reservation failed.

To prevent the end-to-end delay from being too large caused by the BHP buffering at immediate nodes, we use the parameter N_r, representing the maximum buffering times for each BHP during transmission. The required extra offset time is thus limited also to: $T_{maxoff} = T_{wait} \times N_r$.

5. Simulation results

To check the efficiency of our schemes we have simulated a multiple hop network with a ring topology on a dedicated platform. We use OPNET as a simulation tool to study the performance of our schemes and compare them with existing dropping schemes especially when the offset times are varied during transmission. The shortest-path-first routing method is used to establish a route between each pair of edge nodes E_i(i=1 to 16), and the maximum hop distance is 10. Bursts are generated at each edge node E_i. We assume that the burst inter-arrival time follows an exponential distribution and the burst size follows a normal distribution. Note that these assumptions are the same as the ones in [16]. Therefore in our simulations we have use the same ring topology and environment to test the performance of the integrated scheme when performing joint QoS with absolute constraints. The average burst size is 50 Kbytes. All bursts are assumed to have the same initial offset time (the default value is 5ms, which is small enough even for real-time applications). For a core node C_i(i=1 to 16), we assume that each output link consists of 16 wavelengths with a transmission rate of 1 Gbps per wavelength. The basic processing time for BHP at each core node is set to be 0.1 ms. To investigate the service differentiation, we consider four classes, a load distribution of $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$, and proportional parameters of s_1=1, s_2=2, s_3=4, and s_4=8.

For simplicity, in the simulations the traffic load refers to the average traffic load on the links connecting two core nodes. Regarding the performance of the advanced timer-based assembly algorithm, we set M = 50. Figure 4 shows the impact of parameter α on the burst loss ratios for different traffic loads. It shows that the burst loss ratio decreases as α decreases. On the other hand, the average edge buffering delay increases as α decreases, as shown in Fig.5.

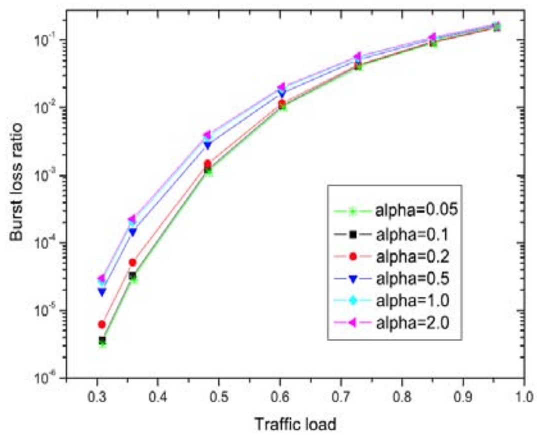

Fig. 4. Impact of α on burst loss ratio

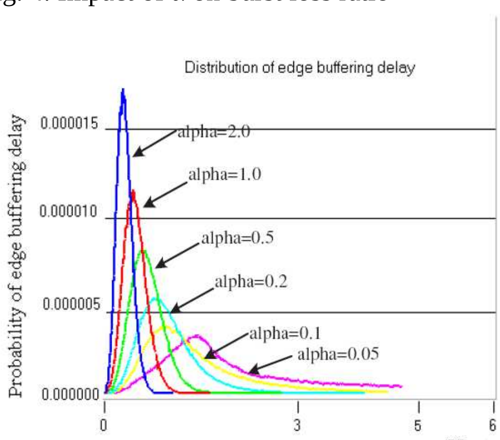

Fig. 5. Impact of α on edge buffering delay

How to choose α is a tradeoff between the burst loss ratio and the average edge buffering delay. A suitable α should be able to achieve a low burst loss ratio while not increasing the edge buffering delay too much. From Figs.4 and 5, we can see that when α is smaller than 0.1, the burst loss ratio decreases slowly and is almost the same as the one for α = 0.1 while the average edge buffering delay obviously increases.

Regarding the performance of the DBA scheme, Figs.6 and 7 show the impact of N_r and T_{wait} on the average burst loss ratios for a low traffic load (0.3) and for a high traffic load (0.89), respectively. In the figures 6 and 7, the curve "*aL*" denotes the burst loss ratio when T_{wait} is set to *a* times the average burst length (*L*). The initial offset time is set to 30 ms. The results indicate that as N_r and T_{wait} increase, the burst loss ratio decreases. However, when T_{wait} exceeds the average burst length duration and N_r exceeds 4, the burst loss ratios decrease slowly and eventually become almost the same. Thus, in the simulations we set 0.4 ms (*average burst length duration=average burst length (50 Kbytes) /bandwidth (1 Gbps)*) and 4 as the default values of T_{wait} and N_r.

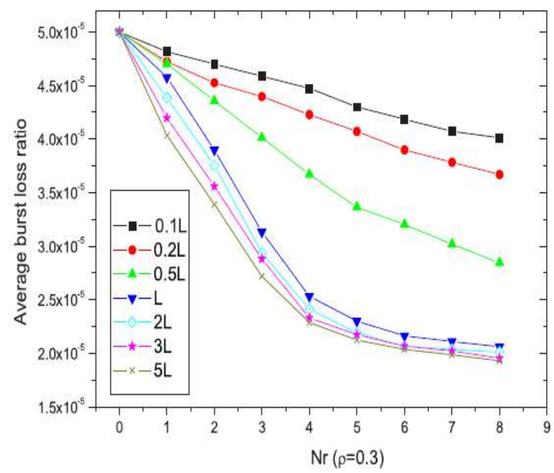

Fig. 6. Average burst loss ratios versus N_r with traffic load 0.3

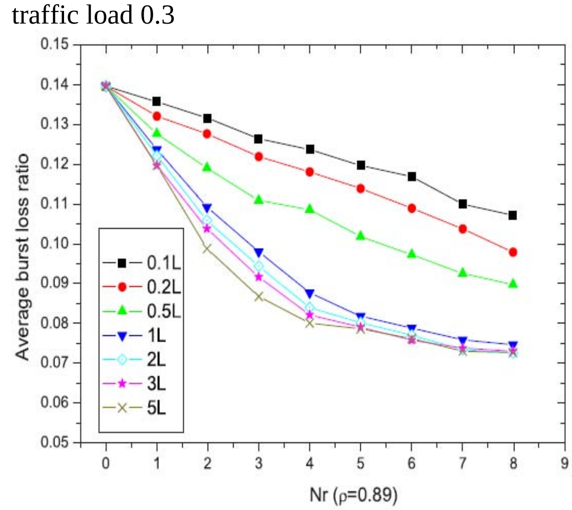

Fig. 7. Average burst loss ratios versus N_r with traffic load 0.89

Regarding the performance of the integrated scheme presented in section 4.3, the results obtained by simulation show that the proportions of different classes are close to the ratios of predefined parameters and are independent of the traffic load for the integrated scheme. Therefore, the integrated scheme achieves proportional differentiated services for multi-class traffic. Figure 8 shows the end-to-end BHP queueing delay normalized to the average burst length duration for each class during transmission.

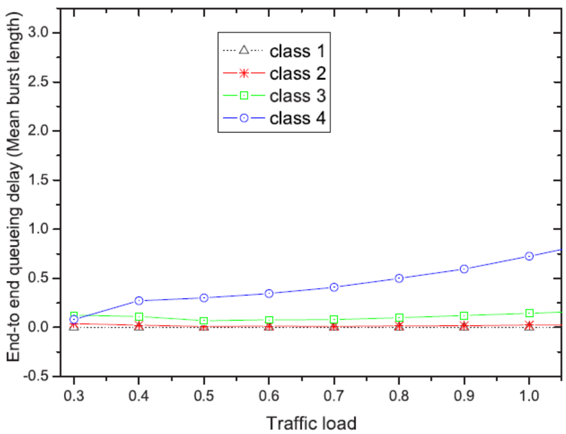

Fig.8 End-to-end BHP queueing delay

We can see that the lower priority bursts have larger queueing delay than do the higher priority bursts. For example, the queueing delay is several microseconds for class 2, tens of microseconds for class 3, and hundreds of microseconds for class 4. The simulation results show that although the integrated scheme improves the burst loss performance at the expense of increasing extra offset time, the increase of end-to-end delay is very small (at most hundreds of microseconds) and would be negligible for real applications.

6 Conclusions

In this paper we described how to provide proportional differentiated services in an OBS network. First, we proposed a delayed burst assignment scheme based on relative QoS model. Then we used a Fractional Brownian motion (FBM) model to analyze the effect of OBS assembly mechanisms on the traffic burstiness. Our simulation results showed that the timer-based assembly algorithm could not reduce the burstiness of real traffic. Therefore, we proposed a novel adaptive burst assembly algorithm with traffic smoothing functions. Simulation results showed that the advanced timer-based assembly was best, because it was not only able to reduce the burst loss ratio, but could also control the edge buffering delay. Through simulation, we also found that when the BHP waiting time period and the maximum number of times for a BHP to be queued exceed some values, the performance is improved smoothly. These results could prevent the end-to-end delay from becoming too large. As an advantage over the existing priority schemes, our integrated scheme does not need any complex burst segmentation or wavelength preemption support, so it have a simple implementation. Moreover, it provides controllable and predictable proportional differentiated services for each class.

For further work we will proceed to test the discussed algorithm over other network topologies. We will try also if the algorithms which lead to the improvement of TCP performance, can be used for UDP flows too. As an unreliable transport layer protocol, UDP has neither congestion control nor error recovery mechanism. Unlike TCP flows, UDP senders don't require the correct receipt of all transmitted packets. However, as the Internet applications become diverse, reliability requirement also becomes varied. So, if we can transmit some special packets, the performance will be improved distinctly.

ACKNOWLEDGEMENTS

This work was partially supported by the Romanian Ministry of Education and Research under Grant No. 121-IDEI/2007.

References:

[1] H.J. Chao and X. Guo, *Quality of Service Control in High-Speed Networks*, Wiley-Interscience, 2001

[2] H.Zang, J.P.Jue and B.Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," SPIE Optical Networks Magazine, vol.1, no.1, January 2000

[3] Tabbane, N., Tabbane, S., Mehaoua,M., Stationary Stochastic models for Forecasting QOS in Ad Hoc Networks for real-time Service support, *WSEAS Transactions on Communications*, Issue 2, vol.3, pp. 405-410, April 2004

[4] Y.J.Xiong, M.Vandenhoute and H.C.Cankaya, "Control architecture in optical burst-switched WDM networks," IEEE Journal on Selected Areas in Comm., vol.18, no.10, pp.1838–1851, 2000

[5] P. Du, QoS Control and Performance Improvement Methods for Optical Burst Switching Networks, PhD thesis, The Graduate University for Advanced Studies (SOKENDAI), 2007

[6] Y.Chen, C.Qiao and X.Yu, "Optical Burst Switching (OBS): A New Area in Optical Networking Research," IEEE Network Magazine. vol.18, pp.16–23, May 2004

[7] K. Park and W. Willinger (editors), *Self-similar network traffic and performance evaluation*, Wiley, 2000

[8] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, *On the self-similar nature of Ethernet Traffic*, Proceedings of ACM SIGCOMM'93, pp.183-193, 1993

[9] Dobrescu, R. Taralunga, S., Mocanu, S., Parallel Internet Traffic Simulator with Self-Similar Scale-Free Network Models, *WSEAS Transactions on Advances in Engineering Education* Volume 5, Issue 2, Volume 5, February 2008, pp. 61-68

[10] Yang, W-H, Wen, C-C., Optical Label Switching Method based on Optical CDM Code Processing, *WSEAS Transactions on Communications*, Issue 5, vol.6, pp. 701-706, May 2007

[11] Dovrolis, C., Ramanathan,P., Dynamic Class Selection: From Relative Differentiation to Absolute QoS," *Proceedings of the IEEE Int. Conf. Network Protocols*, 2001

[12] Kaheel, A., Khattab, T. Mohamed, A., Alnuweiri, H., Quality-of-Service Mechanisms in IP-over-WDM Networks, *IEEE Communications Magazine*, vol.40, no.12, December 2002.

[13] Zhang, Q., Vokkarane, V.M., Chen, B., Jue, J.P., Early drop and wavelength grouping schemes for providing absolute QoS differentiation in optical burstswitched networks, *Proceedings of IEEE GLOBECOM 2003*, vol.5, pp.2694–2698, December 2003

[14] Chen, Y., Hamdi, M., Tsang, H.K., Qiao, D., Proportional QoS provision: a uniform and practical solution, *Proceedings of IEEE ICC 2002*, vol.4, pp.2363–2367, May 2002

[15] Zhang, H., Jue, J.P., Mukherjee, B., A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks, *SPIE Optical Networks Magazine*, vol.1, no.1, January 2000

[16] Yu, X., Chen, Y., Qiao, C., A Study of traffic statistics of assembled burst traffic in optical burst switched networks, *Proceedings of SPIE Optical Networking and Communication Conference 2002*, pp. 149–159, 2002