A RETRIAL QUEUEING MODEL WITH SCHEDULING ADMISSION CONTROL IN QOS SUPPORTING FOR FDL-EQUIPPED AT OBS CORE NODE

Dang Thanh Chuong[1], Nguyen Hong Quoc[2],
Doan Van Thang[3] and Pham Trung Duc[1]

[1]Faculty of Information Technology, University of Sciences, Hue University, Vietnam
[2]University of Education, Hue University, Vietnam
[3]Faculty of Information Technology, Industrial University of HCM City, Vietnam

ABSTRACT

Optical burst switching (OBS) is being considered as an optical transmission technology that meets the increasing bandwidth demands and requirements of quality of service (QoS) of the next generation all-optical Internet. Due to technological limitations such as optical buffers are not yet available at core nodes, ensuring QoS in OBS networks is difficult to perform flexibly compared to electronic networks (e.g. IP network). In this paper, a combination of the prediction based on burst traffic and fibre delay links (FDL) is proposed to allocate wavelengths properly, ensuring QoS and improving the network performance. Efficiency evaluations based on mathematical analysis and simulation will confirm the advantages of our proposed model.

KEYWORDS

OBS, Quality of Services (QoS), Share-Per- Link (SPL), Fiber Delay Lines (FDL), Retrial Queueing.

1. INTRODUCTION

Appearing since the 1990s, the optical network has undergone several generations of development: From the wavelength routing model which is based on end-to-end light path to the optical packet switching (OPS) model which has been proposed recently. However, due to limitations in optical technology, such as optical buffers (similar to RAM in electronic environments) or switches at nanosecond speeds cannot yet be produced, OPS is not currently possible. A compromise solution is optical Burst switching (OBS), which is a promising communication model for the next generation Internet. In optical burst switching networks (OBS networks), the control part (Burst Header Packet, BHP) is separated from the data part (Data Burst). Specifically, the communication in OBS networks is time-separated in which BHP is created and sent ahead of the corresponding burst in an offset time. This offset-time needs to be large enough for resource reservation and configuration of the switches at intermediate nodes so that these operations must be completed before the burst arrival. The communication in OBS networks is also spatially separated, in which one or several wavelength channels are reserved for BHP transmission, while the remaining channels carry data bursts [1].

In optical burst switching networks, congestion will occur when two or more arriving bursts on the same output port/the wavelength. If the wavelength of an incoming optical burst is occupied at the output port when another optical burst arrives, then the optical burst can switch to another free wavelength (using a wavelength converter). The case of all wavelength channels at one

DOI: 10.5121/ijcnc.2020.12504

Electronic copy available at: https://ssrn.com/abstract=3720431

output port are occupied, the incident optical bursts may use Fibre Delay Lines (FDL) optical delay or deflection routing to resolve contention. One approach to limiting the contention problem that causes bottlenecks at the OBS network core node is scheduling admission control [1]. Scheduling admission control can be performed at both the input and core nodes; however, the use of electronic buffers that control scheduling at the edge node is usually more efficient [1]. In most scheduling admission control techniques, data bursts with a low QoS (priority) layer will have a higher probability of being dropped due to conserve resources for the higher QoS layer when contention occurs. The scheduling of incoming bursts is usually done sequentially, on a first come-first served basis. However, when QoS is taken into consideration, successfully scheduling a burst first (with a low QoS layer) can cause congestion for a subsequent cluster (with a higher QoS layer). Therefore, scheduling with admission control is necessary to conserve more resources for high QoS bursts, while limiting resources for low QoS bursts. Besides, scheduling admission control can also be combined with FDL to further assist in congested resolution. An FDL can allow a certain latency value to propagate the bursts, so the integration of additional FDL into the OBS core node can be considered as a buffer with a limited size. However, different from electronic buffers, in optical networks the bursts cannot wait with an indefinite time value (exceeds the allowed delay for optical network), and then the bursts may be dropped after a waiting period without being served.

Due to the limited delay time and the bulky size of FDLs, the type of fixed FDLs, which only delays bursts in a fixed period, is usually equipped at the OBS core nodes. With this type of FDLs, bursts are only delayed in a fixed time; they will be dropped if they are not served when going out of FDLs. The analysis of retrial queue model with FDLs has been studied in [2, 3] and [4], in which the authors in [4] combine the overflow model based on the MMPP process, which the idea is derived from traditional channel switching and fixed-point iterations. This combination model is to calculate the blocking probability which is presented as a function of the buffer parameters in the system. For the analysis in [3], the queue model $\mathbf{MM}\sum_{k=1}^{K}\boldsymbol{CPP_k}/\mathbf{GE}/\mathbf{c}/\mathbf{L}$ is used, where $\boldsymbol{K}$ bursts arrive independently according to Compound Poisson Process (CPP), which means that the arrival of bursts is in batches; the service time is distributed according to the generalized exponential (GE) distribution and there are $\mathbf{c}$ used wavelengths.

In this paper, we study a retrial queue model to analyze the problem of scheduling admission control considering QoS to limit the issue of dispute, at the same time combine using the FDL optical delay line to handle when the issue of disputes an OBS network core, in which the character of FDL as analyzed above can be considered as properties have/not have the patience of customers in retrial queues. The analytical model is an extension of some proposed models [1, 2] with consideration of the retrial factor, ie considering the possibility that the burst may not be fed to the FDL when it is congested with a probability value of θ or θ_1. A burst is called retrial when it has to pass through one of N_q FDL. A retrial burst will reuse a wavelength channel if it is available at the time the burst that comes out from FDL.In the analysis model here, the QoS threshold value will be adjusted based on the incoming traffic of the bursts at the OBS core node. From the analysis model, we also propose an algorithm that is a combination of the prediction based on arriving burst load and FDLs. The prediction is intended to allocate wavelength channels reasonably, while FDLs are to ensure the quality of service. This is also different from the analysis model in [3] and [4].

The paper consists of three sections. Section 2 and Section 3 will provide an analysis model and a scheduling admission control algorithm of performance analysis at OBS core nodes with different traffics (QoS) will be considered. The results analysis, combined with simulation will be mentioned in Section 4 and the conclusion is given in Section 5.

Electronic copy available at: https://ssrn.com/abstract=3720431

2. THE ANALYSIS MODEL

2.1. The assumptions

As in [1], the Markov model will also be used to perform OBS core node analysis. The model is based on the following assumptions:

- The distribution of traffic to the output ports is the same, so we only need to consider it at one output port.
- The core OBS node with SPL architecture has K input ports and K output ports (Figure 1), one WDM fibre in each port, has W wavelength $\Lambda = \{\lambda_0, \lambda_1, \cdots, \lambda_{W-1}\}$ (Assuming that the wavelength conversion capability is sufficient so there will have W the CWC converter on each output port).
- The model considers the quality of service (QoS) by wavelength resource, that is, the high priority traffic (bursts) will be allocated more resources than the low priority traffic. Accordingly, high QoS bursts are scheduled on any wavelength at an output port (W wavelength) and are scheduled to void filing; while the number of wavelengths that low QoS bursts can use is only W_L ($W_L < W$) and adjusted to increase/decrease periodic depending on the arrival rate of the bursts during the observation period (Δt). In other words, among W wavelengths, high QoS bursts are used exclusively ($W - W_L$) wavelengths, while the remaining W_L wavelengths are shared for both bursts (high QoS burst and low QoS burst when all wavelengths ($W - W_L$) are busy). Unlike some previous studies [1, 2], in this paper, the value of W_L will be a function of the arrival rates of the bursts (as the formula (1)).

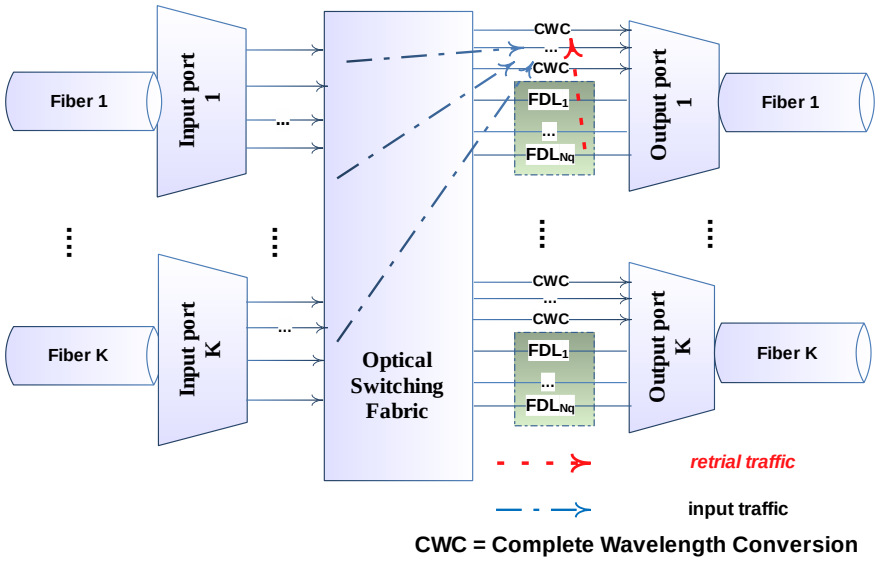


Figure 1. The core OBS node with FDL-SPL architecture

- The burst length is exponentially distributed with the average value $L = 1/\mu$; μ is the average burst serving rate;
- FDL architecture with fixed length F-FDL (Fixed-length FDL) uses N_q FDLs with the same length and therefore all FDL generates the same D delay [2].

Electronic copy available at: https://ssrn.com/abstract=3720431

- In the model we analyzed, a burst to the output port at the time all W wavelengths used was thought to be dropped. A congested low priority burst will always be able to use a buffer (if available) to retry forwarding back to the output port.
- A burst is called retrial when it has to pass through one of N_q FDL. A retrial burst will reuse a wavelength channel if it is available at the time the burst exits from FDL. In this model, we study the queue retrial model with the FDL have to be considered as the retrial of the bursts (Figure 2). This is also different from the analytical model in [1, 5].

2.2. Analysis model with the retrial

2.2.1. The admission control scheduling

This model corresponds to the case of analysis with the traffic is Poisson traffic, corresponding to two service classes with different priorities, corresponding to H (high QoS bursts) and L (low QoS bursts), where class H has a higher priority than class L. Therefore, the model has the form $M/M/\omega/\omega + N_q$ [6] described as in Fig. 2. The high QoS and low QoS bursts coming on the output port with the average rate are γ_H and γ_L, respectively. Here, we set $\gamma = \gamma_H + \gamma_L$ [2] is load flow to the average, so that is $\rho = \rho_H + \rho_L$, where $\rho_H = \gamma_H/\mu$ is the average load of high QoS burst and $\rho_L = \gamma_L/\mu$ is the average load of the low QoS burst.

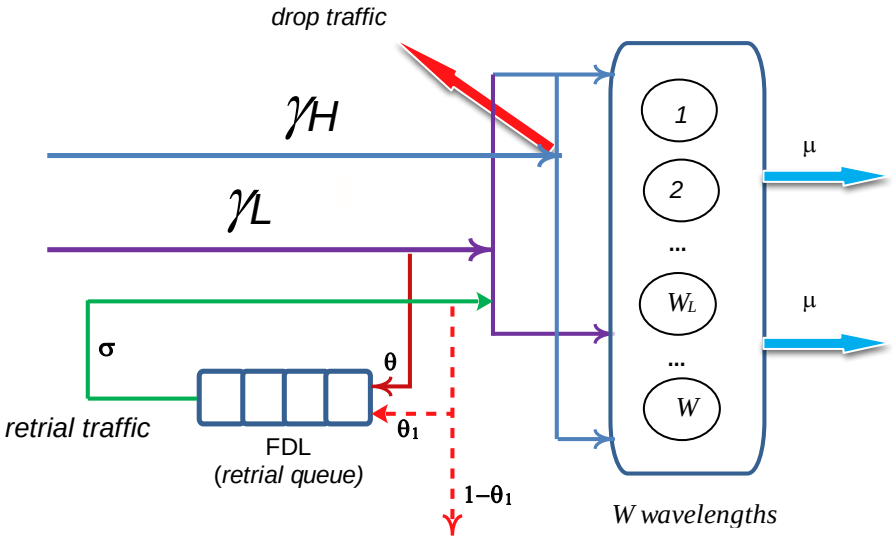


Figure 2. The retrial queueing analysis model with scheduling admission control for core nodes FDL-SPL architecture in QoS supported

- In Figure 2, low QoS bursts are put into the FDL queue with probability θ $(\theta \leq 1)$ when congested, called retrial burst, and is reuse an available wavelength channel at the time the retrial burst exits from FDL. Otherwise, it is returned to one of the FDLs with probability $\theta_1(\theta_1 \leq 1)$ or is dropped with probability $1 - \theta_1$ $((1 - \theta_1)$ is the probability that the burst leaves the system forever) [2].
- The interval time between the retrial bursts is also exponentially distributed and the retrial rate is σ.
- The value of the wavelength channel reserved for the low priority class W_L will be determined by a function according to the value of incoming loads and will vary accordingly with the traffic flow of priority classes 0 and 1 (corresponding to high QoS

Electronic copy available at: https://ssrn.com/abstract=3720431

and low QoS). Accordingly, the need to calculate the burst (high QoS and low QoS) load to perform bandwidth distribution (W_L value) for the low QoS bursts most effectively to avoid wasting resources as well as reducing the probability of congestion. At that time, the number of wavelength channels allocated for low QoS bursts W_L ($W_L < W$) was determined as the upper integer part of the formula (1) as follows [2]:

- $$W_L = \left\lceil W \times \frac{\gamma_L}{\gamma_H + \gamma_L} \right\rceil \quad (1)$$

With loading incoming traffic of high QoS bursts and low QoS bursts are determined based on burst rate, average burst length and current bandwidth per wavelength channel in each observation window Δt and are calculated as follows:

$$\gamma_L = \frac{\gamma_L^A}{\mu_L \times W}; \gamma_H = \frac{\gamma_H^A}{\mu_H \times W} \quad (2)$$

Where $\gamma_L(\gamma_H)$ is the arrival burst rate, $\mu_L(\mu_H)$ is the average burst length of the low QoS bursts (and corresponding high QoS bursts, where we assume that $\mu_L = \mu_H = \mu$) in each observation window, calculated as follows:

$$\gamma_{\mathrm{L}}^{\mathrm{A}} = \frac{\mathrm{number_of_bursts_to_the_QoS_is_low}}{\Delta \mathrm{t}}; \ \gamma_{\mathrm{H}}^{\mathrm{A}} = \frac{\mathrm{number_of\ bursts_to_the_QoS_is_high}}{\Delta \mathrm{t}} \quad (3)$$

$$\mu_{\mathrm{L}} = \frac{\mathrm{total_burst_length_to_QoS_low}}{\mathrm{number_of_bursts_to_the_QoS_is_low}}; \ \mu_{\mathrm{H}} = \frac{\mathrm{total_burst_length_to_QoS_high}}{\mathrm{number_of\ bursts_to_the_QoS_is_high}} \quad (4)$$

With the above-mentioned approach, it makes ensuring priority for high QoS bursts, more flexibly allocating the amount of wavelength that can be used for low QoS bursts. However, to support low QoS traffic when congestion occurs, in this analytical model, we use FDLs additional for low QoS bursts with consideration for factors of re-service (retrial) [2], i.e considering the possibility the burst that may not be brought to FDL when it is congested with a probability value of θ or θ_1. At that time, a burst is called retrial when it has passed through one of N_q FDL. A retrial burst will reuse a wavelength channel if it is available at the time the burst exits from FDL.

2.2.2. The state transition diagram

The state transition diagram then is shown in Figure 3. The state of the system is described by two random variables in continuous time, $\{X(t), Y(t): t \geq 0\}$, in here $X(t)$ is the number of occupied wavelength channels and $Y(t)$ is the number of low QoS bursts congested in FDL (retrial queue) at time t. The state-space of the Continuous-time Markov Chain (CTMC) process (denoted by S) is described as follows: $S = \{i, j\}$ with each pair (i, j) defined: $i = 0, 1, 2, \ldots, W; j = 0, 1, 2, \ldots, N_q$. Let π_{ij} be the steady-state probability that the system achieves in the state (i, j).

Electronic copy available at: https://ssrn.com/abstract=3720431

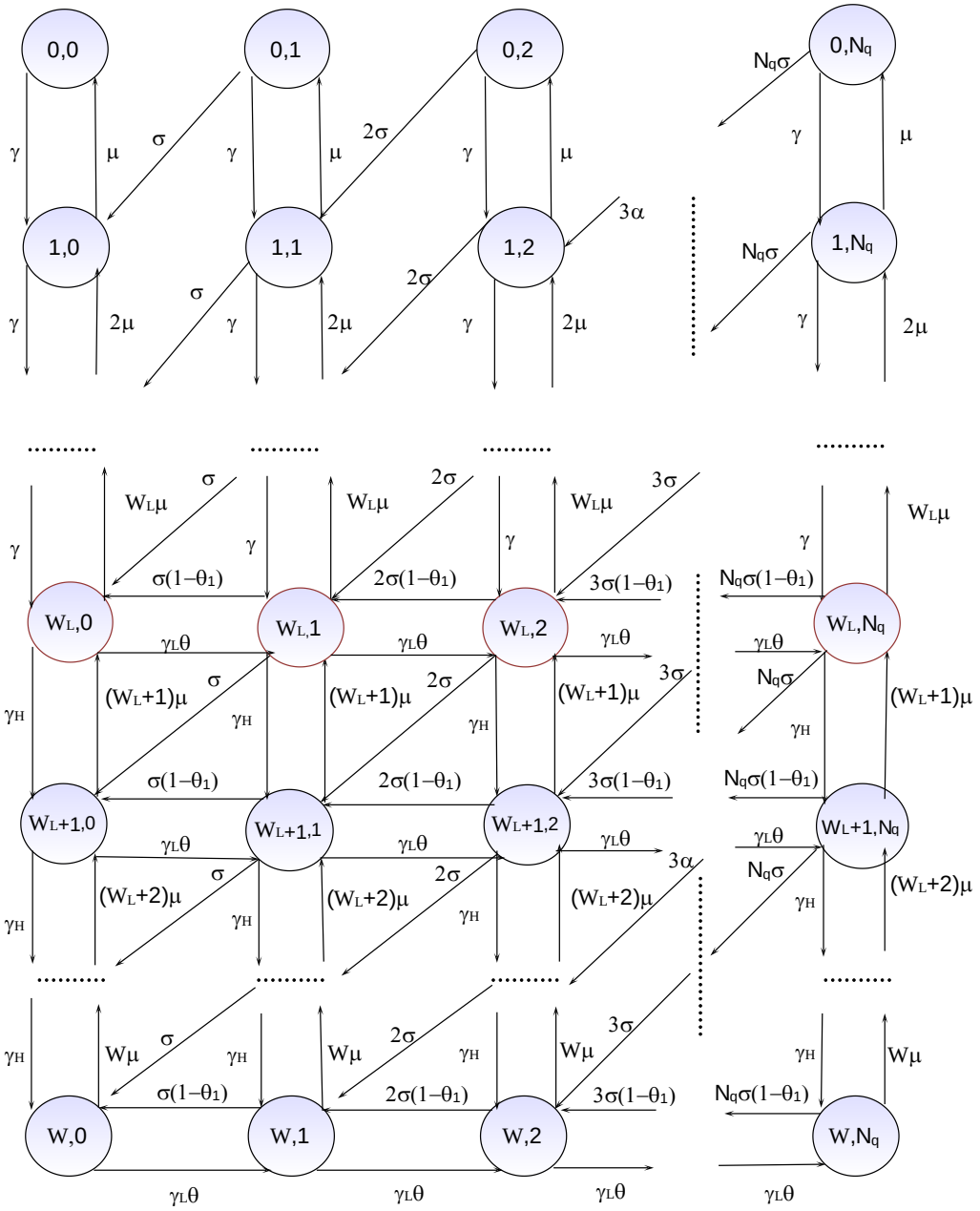


Figure 3. State transition diagram of the analytical model

We construct the transition rate matrix Q based on the transition state diagram shown in Figure 3 and the sub-matrices as follows [2]:

- $A_j(i,k)$: defines the transition rate from the state (i,j) to state (k,j) (with $0 \le i,k \le W; 0 \le j \le N_q$) in cases when the burst arrives or when the burst is served. Service time is exponentially distributed with the parameter μ. The matrix $\mathrm{A_j}$ has dimensions $(W+1)\times(W+1)$ with elements $\mathrm{A_j(i,k)}$. Since j is an independent level of $\mathrm{A_j}$, we can write $\mathrm{A_j} = \mathrm{A}$. The nonzero element of $\mathrm{A_j}$ is $A_j(i,i-1) = i\mu, i = 0,\dots,W$; $A_j(i,i+1) = \gamma, i = 0,\dots,W_L-1$ and $A_j(i,i+1) = \gamma_H, i = W_L,\dots,W$.

Electronic copy available at: https://ssrn.com/abstract=3720431

$$A_j = A = \begin{pmatrix} 0 & \gamma & 0 & \cdots & 0 & 0 & 0 \\ \mu & 0 & \gamma & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (W-1)\mu & 0 & \gamma_H \\ 0 & 0 & 0 & \cdots & 0 & W\mu & 0 \end{pmatrix}, \left(0 \le j \le N_q\right) \quad (5)$$

- $B_j(i,k)$: indicate a jump (up) from state (i,j) to state $(k,j+1)$ (with $0 \le i,k \le W; 0 \le j \le N_q - 1$) due to a request being served from low QoS bursts but all the wavelength channels for them are busy (when $i = W_L$). Matrix B_j (or B since j is dependent level) has dimensions $(W+1) \times (W+1)$ with elements $B_j(i,k)$. The nonzero elements of B_j is $B_j(i,i) = \gamma_L\theta$ with $W_L \le i \le W$.

$$B_j = B = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_L\theta & 0 \\ 0 & 0 & \cdots & 0 & \gamma_L\theta \end{pmatrix}, \left(0 \le j \le N_q - 1\right) \quad (6)$$

- $C_j(i,k)$: indicate a jump (down) from state (i,j) to state $(k,j-1)$ $(0 \le i,k \le W; 1 \le j \le N_q)$ due to a successful server being returned from FDL (orbit). Matrix C_j has dimensions $(W+1) \times (W+1)$ with elements $C_j(i,k)$. The nonzero elements C_j is $C_j(i,i+1) = j\sigma$ (with $0 \le i \le W-1$) and $C_j(i,i) = j\sigma(1-\theta_1)$ with $W_L \le i \le W$.

$$C_j = \begin{pmatrix} 0 & j\sigma & 0 & \cdots & 0 & 0 \\ 0 & 0 & j\sigma & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & j\sigma(1-\theta_1) & j\sigma \\ 0 & 0 & 0 & \cdots & 0 & j\sigma(1-\theta_1) \end{pmatrix}, \left(1 \le j \le N_q\right) \quad (7)$$

From here we also have the birth matrix Q of process S as follows:

$$Q = \begin{pmatrix} Q_1^{(0)} & Q_0^{(0)} & & & \\ Q_2^{(1)} & Q_1^{(1)} & Q_0^{(1)} & & \\ & Q_2^{(2)} & Q_1^{(2)} & \ddots & \\ & & \ddots & \ddots & Q_0^{(N_q-1)} \\ & & & Q_2^{(N_q)} & Q_1^{(N_q)} \end{pmatrix} \quad (8)$$

With

$$\begin{cases} Q_0^{(j)} = B\left(0 \le j \le N_q - 1\right), \\ Q_2^{(j)} = C_j\left(1 \le j \le N_q\right), \\ Q_1^{(0)} = A - D^A - B, \\ Q_1^{(j)} = A - D^A - B - D^{C_j}\left(1 \le j \le N_q - 1\right) \\ Q_1^{(B)} = A - D^A - D^{C_j}. \end{cases} \quad (9)$$

Electronic copy available at: https://ssrn.com/abstract=3720431

2.2.3. Equation system of equilibrium states

Call $v_j = (\pi_{0,j}, \pi_{1,j}, \dots, \pi_{W-1,j}, \pi_{W,j}), (0 \le j \le N_q)$ and $v = (v_0, v_1, \dots, v_{N_q-1}, v_{N_q})$.

Clearly, the system is described by the Quasi-Birth and Death (QBD) process. Then, we have the system of equations of equilibrium written as follows:

$$\begin{cases} v_0 Q_1^{(0)} + v_1 Q_2^{(1)} = \underbrace{(0,0,\dots,0)}_{W+1}, \\ v_{j-1} Q_0^{(j-1)} + v_j Q_1^{(j)} + v_{j+1} Q_2^{(j+1)} = \underbrace{(0,0,\dots,0)}_{W+1}, (1 \le j \le N_q - 1) \\ v_{N_q-1} Q_0^{(N_q-1)} + v_{N_q} Q_1^{(N_q)} = \underbrace{(0,0,\dots,0)}_{W+1} \end{cases} \quad (10)$$

Standardized conditions:

$$\sum_{i=0}^{W} \sum_{j=0}^{N_q} \pi_{ij} = 1 \quad (11)$$

The system of equations of equilibrium was then built:

$$Q \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N_q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (12)$$

Combine standardized equations $\sum_{j=0}^{N_q} \sum_{i=0}^{W} \pi_{i,j} = \sum_{j=0}^{N_q} v_j \, e = 1$. We calculated v_j $(j = \overline{0, N_q})$.

2.2.4. Equation system of equilibrium states

The blocking probability of each high QoS traffic and low QoS traffic in Figure 3 can be calculated as follows [7, 8, 9]:

- The blocking probability of high QoS traffic - PB_{High}: High priority QoS bursts are congested when there is no free wavelength at the output port at the time of arrival (corresponding to states $(0, W)$ to (N_q, W) in Figure 3. At these states, all wavelength channels have been occupied and the waiting space in the retrial queue is only for congested low QoS bursts can be retried the service requests.

$$PB_{High} = \sum_{j=0}^{N_q} \pi_{W,j} \quad (13)$$

- The blocking probability of low QoS traffic - PB_{Low}: Low QoS bursts will be congested when the W_L wavelengths are busy (occupied) at the time the low QoS burst arrives. Therefore, incoming low QoS bursts will not be able to be served on the first request. However, in a model with a retrial queue, low QoS bursts become congested because they cannot be served on the first request, and can request service again after a random

Electronic copy available at: https://ssrn.com/abstract=3720431

time. Therefore, the probability of congestion PB_{Low} can be rewritten with retrial as follows:

$$PB_{Low(retrial)} = \frac{\sigma(1-\theta_1)}{\gamma_L \theta} \sum_{i=W_L+1}^{W} \sum_{j=1}^{N_q} j \cdot \pi_{i,j} \quad (14)$$

With the system of linear equations to the formulas (9) and (10) found above, we can solve to find the equilibrium probability π_{ij}(using the method for calculating matrix Q [8]), from there we can calculate the blocking probability according to the formulas (13) and (14).

3. THE ALGORITHM OF SCHEDULING ADMISSION CONTROL THAT CONSIDERS QOS WITH FDL OPTICAL DELAY

To evaluate the analytical model above, in this paper, we also propose a scheduling admission control algorithm that considers QoS with FDL optical delay based on the analytical model idea in Section 2 above. The algorithm will be installed, simulated with network simulation package Obs0.9a running on the NS2 platform. Simulation results are also compared with the analysis results in Section 3. The algorithm details are presented as follows.

Algorithm 1:

INPUT:

- $I = \{b_1, b_2, \ldots, b_n\}$, $b_i(s_i, e_i, prio_i)$, n is the set of incoming burst during the simulation period;
- W: The number of channels out per link $W = \{1,2,\ldots,w\}$;
- W_L: The number of channels dedicated to the low QoS traffic;
- SB_k set of bursts has been scheduled on channel k ($k \in W$);
- FDL delay, $D = 0.0001ms$; corresponds to the value $N_q = 1$

OUTPUT:

- The set of bursts has been scheduled for low QoS class (I_L);
- The set of bursts has been scheduled for high QoS class QoS (I_H);

METHOD:

(initialize) $\Delta t = 0.0001ms$; $W_L = W$;
FOREACH $b_i \in I$ **DO**
  (initialize) $sc = -1$; $gap_{min} = \infty$;
  IF ($prio_i = 0$)
    number_of bursts_to_the_QoS_is_high = number_of bursts_to_the_QoS_is_high + 1;
    total_burst_length_to_QoS_high = total_burst_length_to_QoS_high + ($e_i - s_i$);
    $sc = BFVF(b_i, W)$;
  ELSE
    number_of_bursts_to_the_QoS_is_low = number_of_bursts_to_the_QoS_is_low + 1;
    total_burst_length_to_QoS_low = total_burst_length_to_QoS_low + ($e_i - s_i$);
    $sc = BFVF(b_i, W_L)$;
  IF ($sc \neq -1$)

Electronic copy available at: https://ssrn.com/abstract=3720431

13 $SCHEDULE(b_i, sc)$
14 **IF** $(prio_i = 0)$
15 $I_L = I_L \cup \{b_i\}$;
16 **ELSE**
17 $I_H = I_H \cup \{b_i\}$;
18 **ELSEIF** $(prio_i = 1)$ //using FDL
19 $s_i = s_i + FDL$;
20 $e_i = e_i + FDL$;
21 $sc = BFVF(b_i, W_L)$;
22 **IF** $(sc \neq -1)$
23 $SCHEDULER(b_i, sc)$
24 $I_L = I_L \cup \{b_i\}$;
25 **ELSE**$DROP(b_i)$
26 **IF** $(timecho \geq \Delta t)$ // recalcution W_L
27 $\gamma_L^A = \frac{\text{number_of_bursts_to_the_QoS_is_low}}{\Delta t}$; $\gamma_{\mathrm{H}}^{\mathrm{A}} = \frac{\text{number_of_bursts_to_the_QoS_is_high}}{\Delta t}$;
28 $\mu_L = \frac{\text{total_burst_length_to_QoS_low}}{\text{number_of_bursts_to_the_QoS_is_low}}$; $\mu_H = \frac{\text{total_burst_length_to_QoS_high}}{\text{number_of bursts_to_the_QoS_is_high}}$;
29 $\gamma_L = \frac{\gamma_L^A}{\mu_L \times W}$; $\gamma_H = \frac{\gamma_H^A}{\mu_H \times W}$;
30 $W_L = \left\lceil W \times \frac{\gamma_L}{\gamma_H + \gamma_L} \right\rceil$;
31 number_of bursts_to_the_QoS_is_high = 0; total_burst_length_to_QoS_high = 0;
32 number_of_bursts_to_the_QoS_is_low = 0; total_burst_length_to_QoS_low = 0;
33 **Return;**

FUCNTION $BFVF(ub, W)$:

INPUT

- $ub(s_{ub}, e_{ub})$: incoming burst not scheduled yet,
- W: the number of channels out on each link $W = \{1,2,\ldots,w\}$;
- SB_k set of bursts have been scheduled on channel k $(k \in W)$;

OUTOUT

- sc: corresponding channel to schedule for the incoming burst.

METHOD:

1 $best_utilisation = \infty;\ sc = -1;$
2 **FOR EACH** $k \in W$ **DO**
3 $e_{0,k} = 0; s_{|SB_k|+1,k} =< high\ value >;$
4 **FOR EACH** $j \in |SB_k|$ do
5 **IF**$(((s_{ub} \geq e_{j,k}) \wedge (s_{j+1,k} \geq e_{ub})) \wedge ((s_{j+1,k} - e_{j,k}) < best_utilisation)))$

Electronic copy available at: https://ssrn.com/abstract=3720431

6 $best_utilisation = s_{j+1,k} - e_{j,k}$;

7 $sc = k$;

8 **RETURN** *sc*;

The computational complexity of the Algorithm 1 is $O(n * log(n))$.

4. RESULTS AND DISCUSSION

Based on the blocking probability determined in formulas (12) and (13), we make a graphical description (written in Mathematica) that the variation of blocking probability depends on network traffic load (ρ), number of wavelengths output (W), length FDL. The analysis results will also be compared with the simulation according to Algorithm 1 in Section 3 above (using the Obs0.9a package running on the NS2 simulation software on the Dumbell network topology 20 edge nodes, 2 core nodes, with parameters number $W = 12$, bandwidth per channel $1Gb$, incoming bursts according to Poison distribution) to assess the accuracy of the analysis model. The system model with the parameters W, W_L, N_q changes;$\mu = 0.015625$, $\theta = \theta_1 = 0.5$. Analysis results are also compared with simulations in some special cases. Let $\beta = \rho/W$ is traffic load normalized at each output port, β is considered in the range of 0.2 to 0.9 (Erlang) [1]. From Figure 4 (with the parameters $W = 16$, $N_q = 1$, $\rho_L/\rho_H = 2.333$, W_L calculated from formula (1) according to the change in incoming speed (in this case $W_L = 11$), $\mu = 0.015$ (corresponds to the burst length of 64 bytes), $\sigma = 0.8 \cdot \gamma_L$) we see that increasing traffic load by β, then the blocking probability the first time of high QoS and low QoS bursts increased significantly.

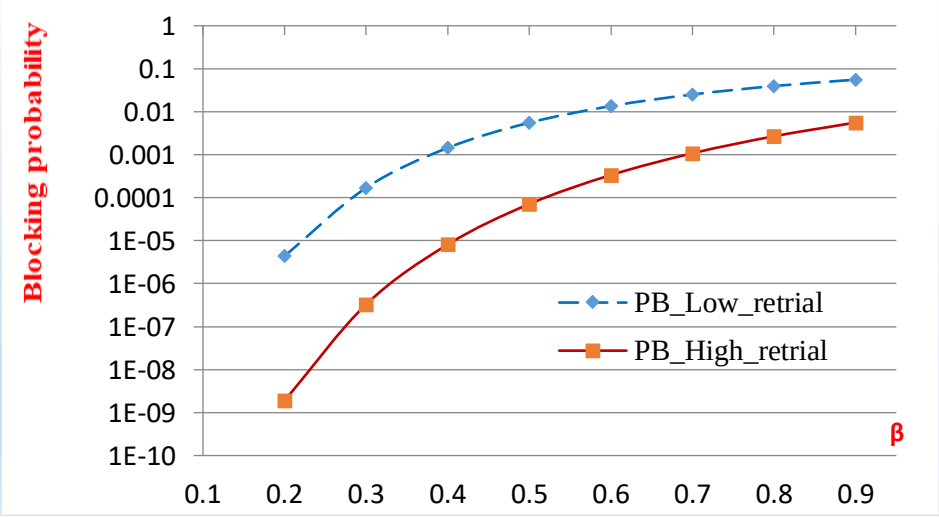


Figure 4. Blocking probability of high QoS bursts and low QoS bursts vs $\boldsymbol{\beta}$

Compared with the result in [1] (consider the case when the second retrial is not considered with $\theta_1 = 0$), we get the result in Figure 5. When retrial is considered with $\theta_1 > 0$, the blocking probability of low QoS bursts is also improved.

Electronic copy available at: https://ssrn.com/abstract=3720431

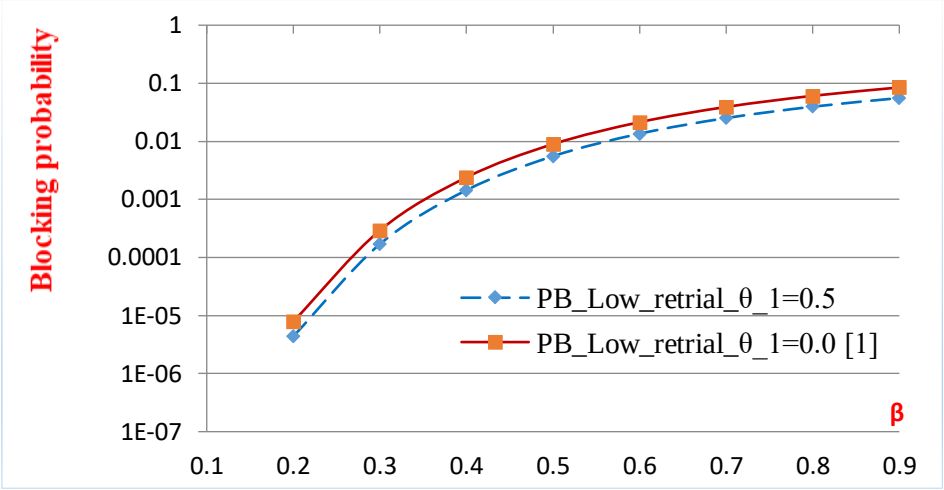


Figure 5. Blocking probability of low QoSbursts with/without retrial for the 2nd time vs $\boldsymbol{\beta}$

Another improvement in this model compared to [1] is that the value of W_L changes with the incoming speed of high and low QoS bursts. This allows adjusting the blocking probability of the bursts when changing loads (changing incoming speed). As shown in Figure 6, when the QoS burst load is high to increase $\rho_H/\rho_L = 2.333$, based on the adjustment of value W_L ($W_L = 5$) value, the blocking probability of high QoS bursts will also be adjusted to not increase too high (as shown in Figure 6, the blocking probability of high QoS bursts will be adjusted from line 2 (which is the line without W_L adjustment) to line 3 (which is the line with $W_L = 5$ adjustment)).

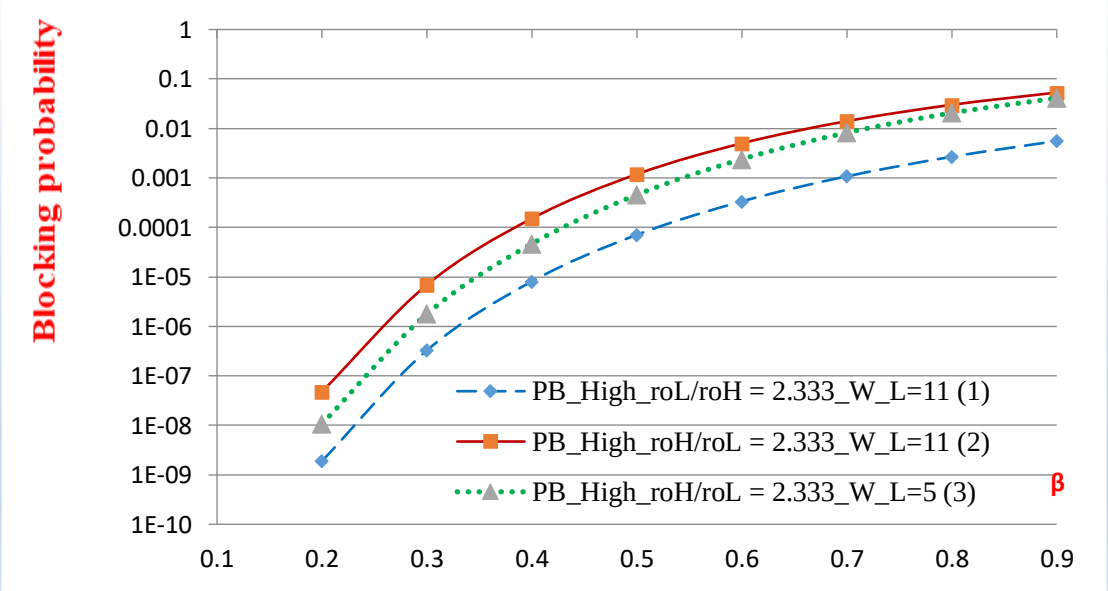


Figure 6. Blocking probability of high QoS bursts when changing load ($\boldsymbol{\rho_L/\rho_H = 2.3}$ and $\boldsymbol{\rho_L/\rho_H = 0.4}$) vs $\boldsymbol{\beta}$

Similarly, when considering the low QoS traffic burst, the adjustment value W_L will be done when the load burst low QoS arrive will increase (as shown in Figure 7, the blocking probability of low QoS bursts will be adjusted from line 1 (which is the line when$W_L = 5$) down to line 2 (is the line with $W_L = 11$ adjustments)).

Electronic copy available at: https://ssrn.com/abstract=3720431

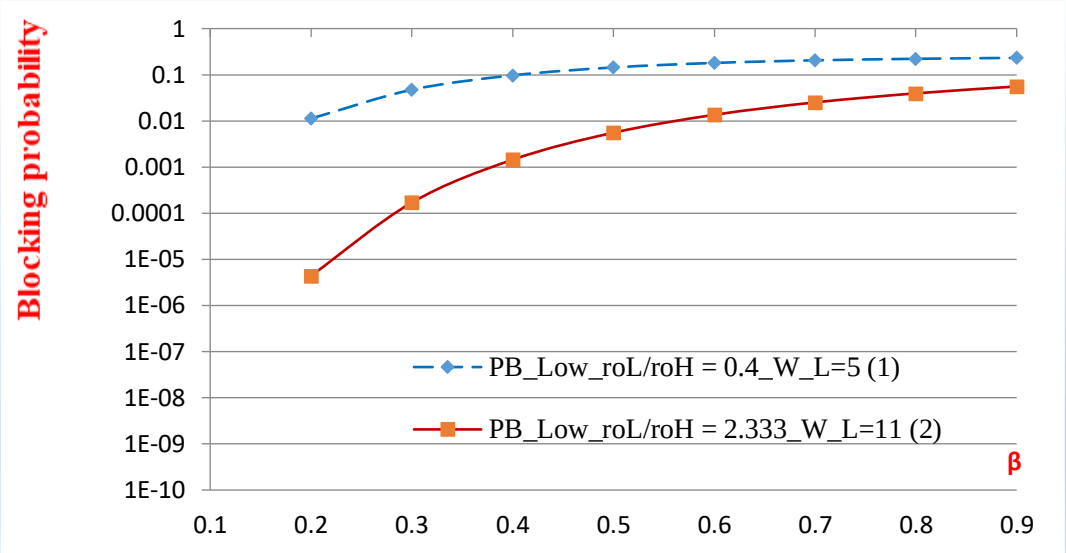


Figure 7. The blocking probability of low QoS bursts when changing loads ($\boldsymbol{\rho_L/\rho_H = 0.4}$ and $\boldsymbol{\rho_L/\rho_H = 2.3}$) vs $\boldsymbol{\beta}$

With the same parameters as shown in Figure 7 but considered with the model's overall blocking probability (blocking probability of both flows) is also shown in Figure 8.

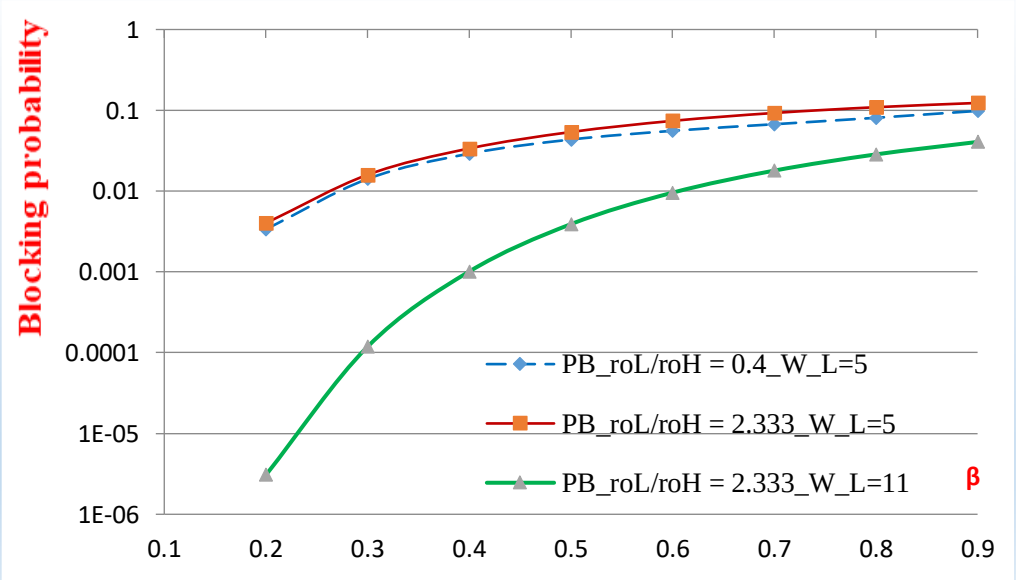


Figure 8. The burst blocking probability when changing loads ($\boldsymbol{\rho_L/\rho_H = 0.4}$ and $\boldsymbol{\rho_L/\rho_H = 2.3}$) vs $\boldsymbol{\beta}$

Figures 9 and 10 show the blocking probability of the low QoS bursts and the general blocking probability (respectively) when adjusting the values of θ_1, $\rho_L/\rho_H = 2.3$, it is clear that when increasing probability θ_1, the blocking probability of bursts low QoS, as well as the overall blocking probability is significantly reduced.

Electronic copy available at: https://ssrn.com/abstract=3720431

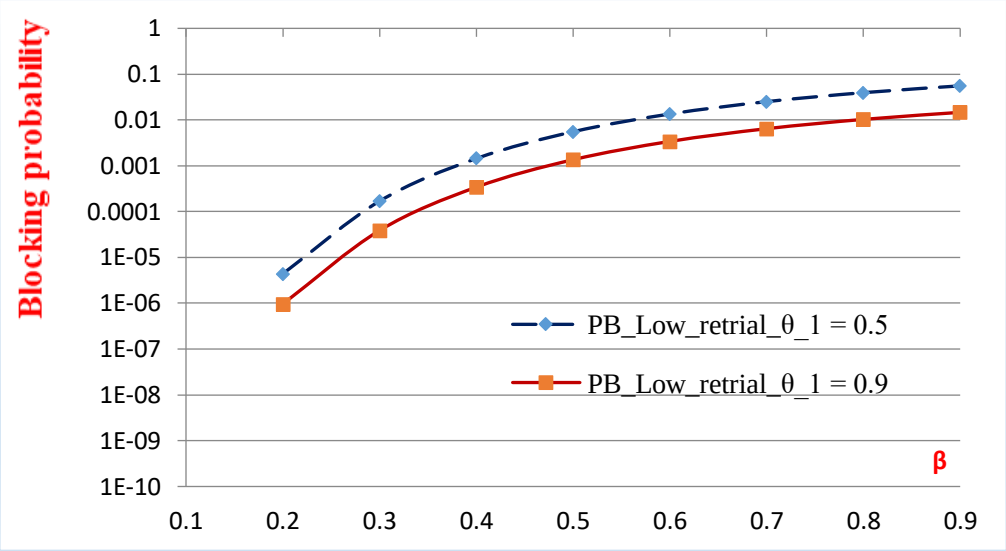


Figure 9. The blocking probability of low QoS bursts when changing $\boldsymbol{\theta_1}$ vs $\boldsymbol{\beta}$

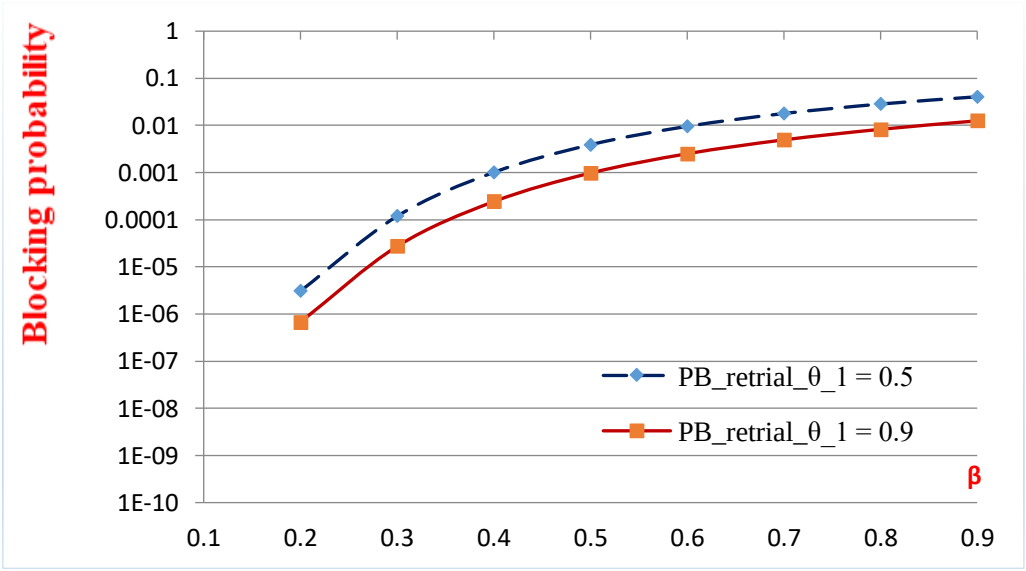


Figure 10. The overall blocking probability of the bursts (low QoS and high QoS) when changing $\boldsymbol{\theta_1}$

We consider the same as above, the blocking probability of high QoS bursts in Figure 11 is also changed (with increasing, especially at high loads) although not significantly (this shows that the value of θ_1 does not significantly affect high QoS bursts).

Electronic copy available at: https://ssrn.com/abstract=3720431

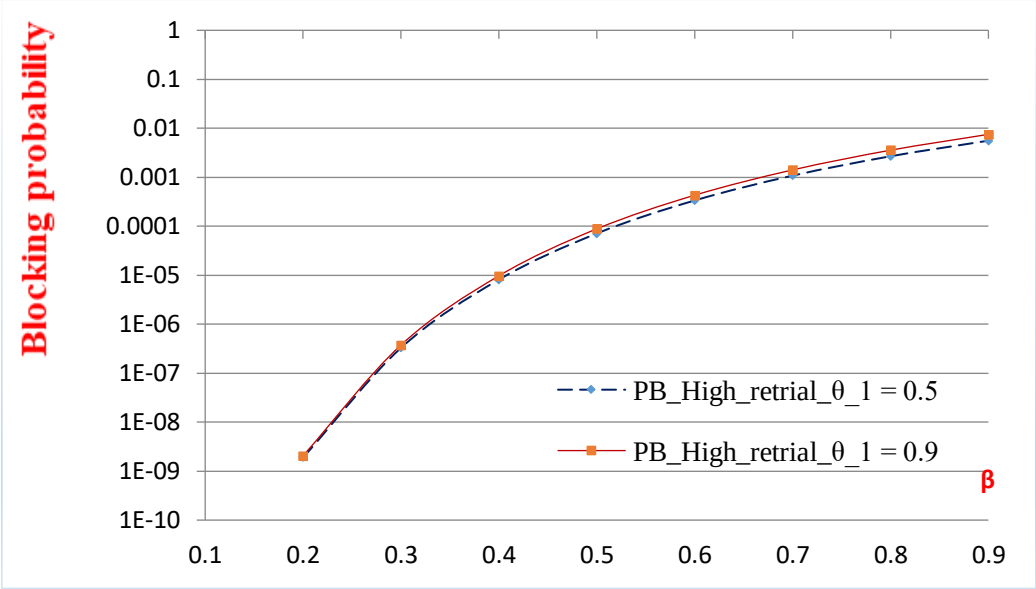


Figure 11. The blocking probability of high QoS bursts when changing$\boldsymbol{\theta_1}$ vs $\boldsymbol{\beta}$

Thus, with the above analysis results, the model presented by the paper meets the requirements of the analysis to ensure that the blocking probability of high QoS bursts is always lower than low QoS bursts based on allocating wavelength channel numbers for low QoS bursts (W_L) and allowable probabilities for low QoS bursts (θ and θ_1). The results of the analysis to ensure reliability and accuracy. In addition, the results of the analysis are also compared with the simulation results (using the OBS.09a simulation package running on NS2 with Algorithm 1) with parameters consistent with the analysis model: $W = 12$, $N_q = 1$, $\rho_H/\rho_L = 2.333$, W_L from formula (1) according to the change in incoming speed (in this case $W_L = 4$), $\mu = 0.015625$ (burst length is 64 byte), $\sigma = 0.65 \cdot \gamma_L$), $\theta = 0.8$; $\theta_1 = 0$. The results shown in Figure 12 show the validity of simulation and analysis results, especially at high loads. However, in some low load cases, the number of bursts arriving in each time slot is small so the prediction has not been effective.

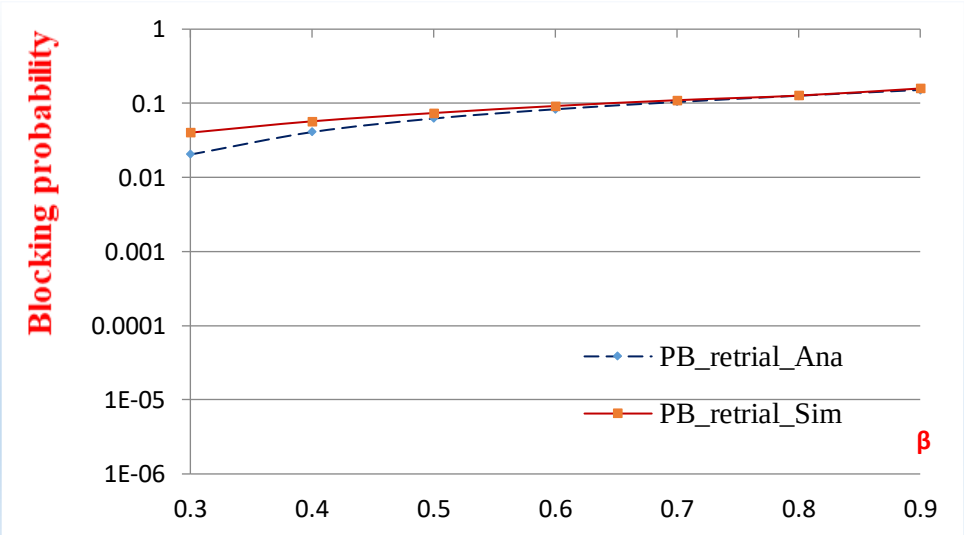


Figure 12. The overall blocking probability compared between analytical and simulation with the retrial vs $\boldsymbol{\beta}$

Electronic copy available at: https://ssrn.com/abstract=3720431

The results in Figure 12 show high accuracy at high loads, but with low loads, there is also a difference but also within permitted limits. This difference is also perfectly theoretical, whereby, when the load is low, the number of bursts coming in each time slot is small, so the prediction may not be effective (in terms of the simulation). Similarly, in Figure 13, with a comparison of the overall blocking probability of the two high and low QoS classes in the absence of retrial between analysis results and simulation results. Figures 12 and 13 confirm the correctness of our proposed model.

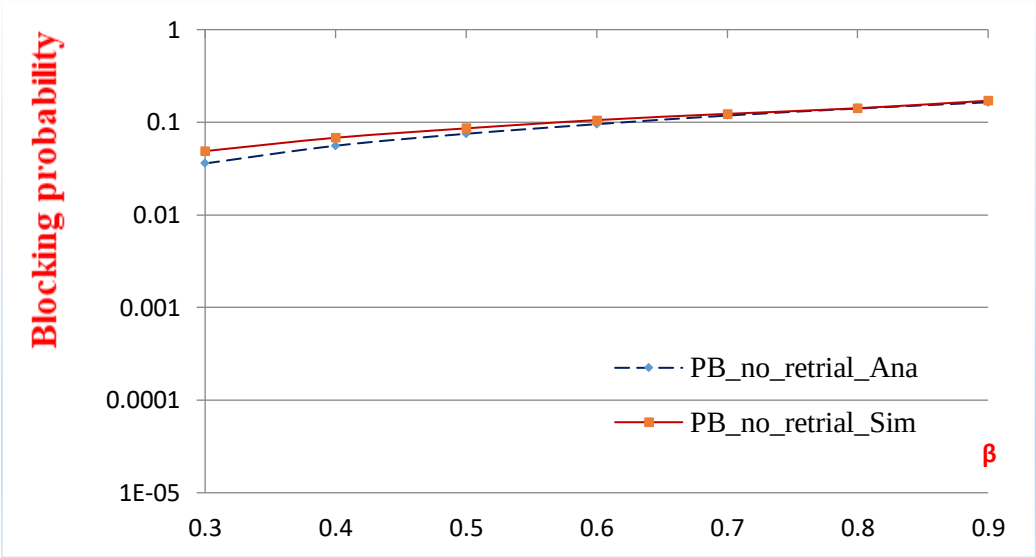


Figure 13. The overall blocking probability compared between analytical and simulation without considering a retrial vs $\boldsymbol{\beta}$

As the results were analyzed on the chart in Figures 7, 8 and 9, we found that when the traffic load bursts changes, W_L value will also be adjusted to change appropriately. In this paper, W_L changes are also done through simulation as shown in Figure 14.

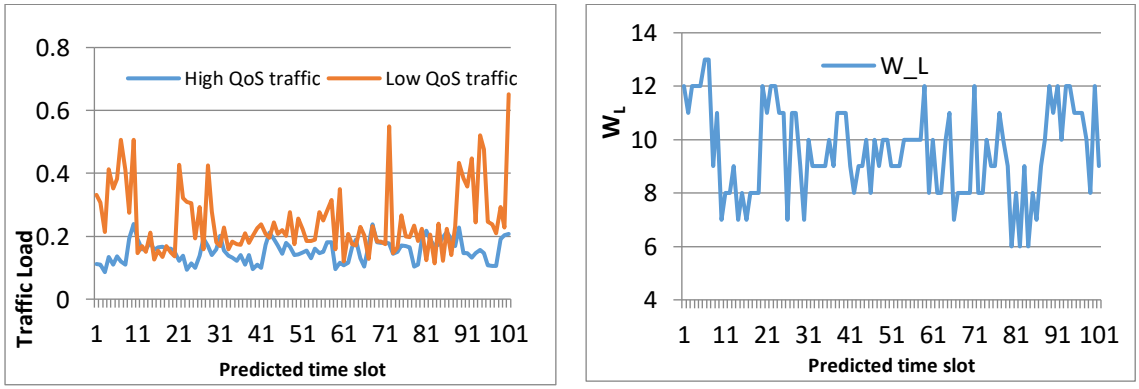


Figure 14. The results for statistics of incoming traffic load and number of reserved $\boldsymbol{W_L}$ channels for the low priority class

The results are shown in Figure 14 show that when the traffic load between the high and low priority class changes, Algorithm 1 will adjust the appropriate W_L value to meet the status of incoming flows until the next time, for the purpose of increasing bandwidth utilization and reducing burst loss. This has been shown in Figures. 6, 7 and 8 (adjustment in value W_L varies

Electronic copy available at: https://ssrn.com/abstract=3720431

when the traffic load of the low QoS bursts increases), which shows that the blocking probability is significantly improved when a reasonable increase of the valueW_L (to $W_L = 11$) is compared with the fixed case W_L ($W_L = 5$). This is also an improvement of the propose model compared to the model in [2].

5. CONCLUSIONS

The paper proposes a performance analysis and evaluation model at the OBS core architecture SPL - feed-forward with the retrial queue model. Different from previous studies, the proposed model considers the retrial factor for FDL optical delay, i.e. considering the possibility that the burst may not be brought to FDL when it is congested with a probability value of θ or θ_1. Also, we proposed an algorithm scheduling admission control with QoS consideration for the analysis model. Theoretical analysis results, as well as simulations, show the correctness of our proposed model.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank everyone, just everyone! This work was supported by the Strong Research Group Program of Hue University.

REFERENCES

[1] Chuong Dang Thanh, Loi Vu Duy, Nhat Vo Viet Minh (2011). A Performance Analysis of Deflection Routing and FDLs with Wavelength-Based QoS in OBS Networks. *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, ISBN: 978-1-4577-1827-4, pp 72-78 (IEEE).

[2] Chuong Dang Thanh, Duc Pham Trung, Thang Doan Van (2018).A Retrial Queueing model with FDL at OBS core node. Network Protocols and Algorithms, ISSN 1943-3581, Vol. 10, No. 3, 1-17

[3] Tien Van Do, Ram Chakka (2010). A new performability model for queueing and FDL-related burst loss in optical switching nodes. *Computer Communications 33* (2010) S146–S151.

[4] Akar Nail, and Sohraby Khosrow (2011). Retrial Queueing Models of Multi-wavelength FDL Feedback Optical Buffers. *IEEE Trans. Communications* 59(10): 2832-2840

[5] Conor McArdle, Daniele Tafani and Liam P. Barry (2011). Analysis of a Buffered Optical Switch with General Interarrival Times. *Journal of Networks*, vol. 6, no. 4, April 2011.

[6] Venkatesh, C. Siva Ram Murthy. An Analytical Approach to Optical Burst Switched Networks, *Springer* ISBN 978-1-4419-1509-2, Chennai, India, August 2009.

[7] Yoshinori Ozaki and Hideaki Takagi (2010). Analysis of Mixed Loss-Delay M/M/m/K Queueing Systems with State-Dependent Arrival Rates. *Advances in Queueing Theory and Network Applications*, W. Yue et al. (eds.), Springer Science+Business Media LLC, pp. 181-194.

[8] Tien Van Do, Ram Chakka. An efficient method to compute the rate matrix for retrial queues with large number of servers. *Applied Mathematics Letters* 23, (2010) 638-643

[9] Daniele Tafani, (2012). Analytic Modelling and Resource Dimensioning of Optical Burst Switched Networks. Doctor of Philosophy, School of Electronic Engineering, Faculty of Engineering and Computing, Dublin City University.

Electronic copy available at: https://ssrn.com/abstract=3720431

AUTHORS

Dang Thanh Chuong obtained his doctorate in Mathematical Foundation for Computers and Computing Systems in 2014 from the Institute of Information Technology, Vietnam Academy of Science and Technology (VAST). He has published over 20 research papers. His research interests are in the fields of all-optical networks with emphasis on packet/burst-based switching, Contention Resolution, and Quality of Service; Queueing Theory and Retrial Queue. Email: dtchuong@hueuni.edu.vn



Nguyen Hong Quoc received his Ph.D. degree in computer science from the Faculty of Sciences, Hue University, Vietnam, in 2018. His research interests are in the fields of all-optical networks with emphasis on packet/burst-based switching, scheduling, and quality of service. Email: nhquoc@hueuni.edu.vn



Van Thang Doan obtained his doctorate in Mathematical Foundation for Computers and Computing Systems in 2014 from the Institute Of Information Technology, Vietnam Academy of Science and Technology (VAST). His main research interests include the fuzzy database, hedge algebras, and detection, prevention, avoidance of cloud computing, and distributed systems. He serves as a technical committee program member, track chair, session chair, and reviewer of many international conferences and journals. Email: doanvanthang@iuh.edu.vn



Pham Trung Duc In 2010, he graduated with a Bachelor of Information Technology from the University of Sciences, Hue University. In 2012, he received a Master's degree in Computer Science from Hue University of Sciences. Currently, he is a PhD student at Hue University of Sciences, Hue University (from December 2016 to the present).
Research fields: OBS network, QoS differentiation, scheduling admission control, QoS improvement, QoS provisioning. Email: phamtrungduc@hueuni.edu.vn



Electronic copy available at: https://ssrn.com/abstract=3720431