Offset Time Configuration in Optical Burst Switching Ring Network

Wei Dai, Guiling Wu, Wenjun Qian, Xinwan Li, *Senior Member, IEEE*, and Jianping Chen

***Abstract*—In this paper, a detailed burst header packet (BHP) processing model is presented to characterize the BHP processing time at the switch control unit of optical burst switching ring network. The distribution function of the total BHP processing delay, which is used to configure offset time, is derived based on the proposed model by using central limit, large deviation, and phase-type distribution theorem-based methods, respectively. A simulation platform that takes the practical processing delay of BHP into consideration is developed to evaluate the efficiency of the proposed model and mechanisms. Different distribution function-based offset time configuration mechanisms are analyzed and compared with the fixed offset time configuration mechanism in terms of burst loss ratio for insufficient offset time and average offset time deviation. The results show the validity of our model.**

***Index Terms*—Control part, distribution function, offset time, optical burst switching.**

I. Introduction

OPTICAL burst switching (OBS) has been proposed as a promising solution for future optical communication networks [1]–[3]. It alleviates the requirement on hardware (such as high-speed optical switch fabric, etc.) by separating the control packet and the related data packet both in time domain and wavelength domain. A burst header packet (BHP) is sent before burst data packet (BDP) by an offset time. The configuration of offset time is a key issue in OBS networks. Too long offset time will result in extra transmission delay, while a shorter offset time may cause the corresponding BDP to arrive before the node reserves an optical path for it and accordingly the BDP has to be dropped.

Theoretically, offset time is equal to the sum of the BHP processing delay at each intermediate OBS node on the OBS path and the setting time of optical fabric [1]. Some studies proposed to configure the offset time by multiplying the hop number along the path with the BHP processing delay at an OBS node, which is considered as a constant value [1], [4].

The BHP processing delay at each OBS node, however, includes both variable queuing delay and variable service time, which is related to the actual BHP traffic load, the processing capacity of the switch control units (SCU), the adopted scheduling algorithm, and so on [5]. If an offset time is configured as a constant value rather than a variable, it may be too long for lightly loaded OBS paths, but not long enough on heavily loaded OBS paths, which causes burst loss. Barakat and Darcie [6] investigated the congestion problem in the control plane of OBS network. In order to solve the problem of burst loss caused by the congestion, Barakat and Darcie[7] and Choi *et al.* [8] proposed to adjust burst assembly parameters, e.g., burst length, to reduce the BHP traffic load. In this case, the assembled bursts will become overlong as the traffic load increases, which will lead to several problems, such as heavier burst collision and longer assembly delay [9]. Configuring an offset time according to the status of networks is a more direct way to solve earlier problems. Hwang *et al.*[10] proposed a dynamic offset time update scheme to guarantee proper offset time according to the backward information from the egress node. Martinez *et al.*[11] proposed a load-adaptive offset time algorithm that takes the variable BHP sojourn time into account through online calculation of probability density function (PDF) of the BHP waiting time. Each node in the OBS network needs to sample individual sojourn time of BHPs, calculate the discrete Fourier transform coefficients, filter and send the coefficients to the ingress nodes. All these studies simply take the BHP processing procedure at the SCU as a single "module" and adopt $\mathrm{M/G/1}$ model to approximate it. The practical BHP processing procedure at the SCU, however, includes several modules and queues. In order to characterize the BHP service time at the SCU that is important for precisely calculating offset time, a BHP processing model, which can describe the inner processing procedure, is needed. At the same time, offset time configuration mechanism including the approaches to obtain the PDF of the total BHP processing time, the feedback information and so on, must be simple enough, so as not to impose too much overhead on the SCU and bandwidth.

In this paper, a detailed BHP processing model at the SCU of optical burst-switched ring (OBSR) network [12]–[15] is built. The distribution function of the total BHP processing delay is then derived by using central limit [16], large deviation [17], and phase-type (PH) distribution theorem-based methods [18], respectively. An OBSR network simulation platform that is able to simulate the practical BHP processing procedure at

Manuscript received August 04, 2008; revised January 23, 2009. First published May 15, 2009; current version published August 26, 2009. This work was supported in part by the National Science Foundation of China (NSFC) under Grant ID90704002 and Grant 60877012, in part by 863 Project under Grant ID2006AA01Z242 and Grant 2007AA01Z275, in part by Dawn Program for Excellent Scholars by the Shanghai Municipal Education Commission, and in part by the Key Disciplinary Development Program of Shanghai under Grant T0102.

W. Dai, G. Wu, X. Li, and J. Chen are with the State Key Laboratory on Advanced Optical Communication Systems and Networks, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: davidsh@163.com; wuguiling@sjtu.edu.cn; lixinwan@sjtu.edu.cn; jpchen62@sjtu.edu.cn).

W. Qian is with China Mobile Group Shanghai Co., Ltd., Shanghai, China (e-mail: qianwj@sh.chinamobile.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2009.2023094

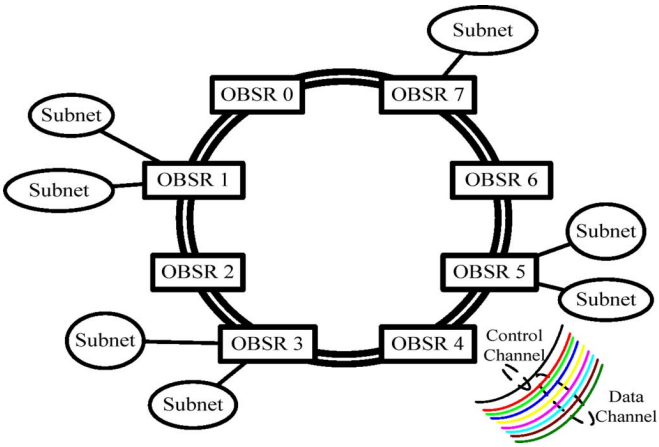

Fig. 1. OBSR network.

the OBSR node is designed. The performance of an OBSR network, which configures offset time according to the distribution function based on the presented model and methods, is evaluated and compared with the fixed offset time configuration mechanism. The paper is organized as follows. Section II builds the BHP processing model at the SCU in the OBSR network. In Section III, an offset time configuration mechanism based on the distribution function is described; the distribution function of the total processing delay is then derived from the built BHP processing model using three approximation methods. Section IV presents an OBSR network simulation platform that considers the practical BHP processing delay in different modules at the SCU. Simulation results of different offset time configuration mechanisms are also presented and compared. Section V is the conclusion.

II. Model for the Control Part in OBSR Node

Fig. 1 shows the architecture of an OBSR network composed of a number of OBSR nodes interconnected by wavelength-division multiplexing fiber links. Each link includes control channels and data channels. In the OBSR network, data packets received from relevant subnets are assembled into BDP based on certain assembly mechanism and sent to the ring according to the destination of the BDP after the BHP with an offset time. BHP from the upstream node is processed at the SCU of the OBSR node. If the current node is the destination, it receives the corresponding BDP and sends the packets to the subnets after packet disassembly. Otherwise, it first reserves optical path for the corresponding BDP according to the information in the BHP and the current status of the network, then forwards the BHP through the control channel to the downstream node.

The BHP processing procedure at the SCU can be modeled as shown in Fig. 2. The receiving module is responsible for receiving BHPs from upstream nodes through the control channels, and putting them into the first-in/first-out queue (i.e., forward queue) together with the BHPs generated for local BDPs. The forwarder assigns the BHP to the corresponding scheduler queue or receiving queue. In the OBSR node, each ring corresponds to a scheduler module, which includes a scheduler queue and a scheduler. The scheduler schedules the data channels in

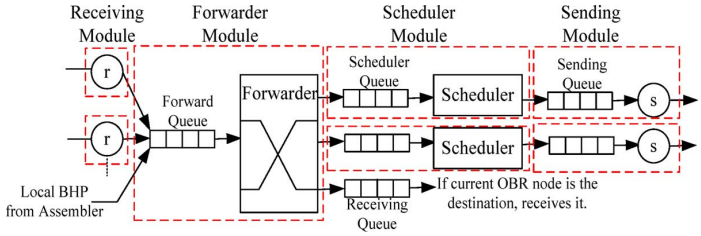

Fig. 2. Model for the SCU of OBSR nodes.

the related ring for the corresponding BDP according to the information in the BHP. After completing the scheduling process, the BHP is put into the sending queue and sent by the sending module.

Based on the earlier model, the total BHP processing delay at each SCU of the OBSR node can be expressed as follows:

$$t_p = t_{\mathrm{r}} + t_{q,\mathrm{f}} + t_{\mathrm{f}} + t_{q,\mathrm{sch}} + t_{\mathrm{sch}} + t_{q,s} \tag{1}$$

where t_p is the total BHP processing delay at the SCU, t_{r} is the time to receive the BHP, i.e., receiving delay, $t_{q,\mathrm{f}}$ is the waiting time in the forwarder queue, t_{f} is the time to forward the BHP, i.e., forwarding delay, $t_{q,\mathrm{sch}}$ is the waiting time in the scheduler queue, t_{sch} is the time to schedule the data channel for the corresponding BDP, i.e., scheduling delay, $t_{q,s}$ is the waiting time in sending queue. Except for t_{r}, all of these delays can be variables (see [19]–[21]) and we treat them as random variables constrained by the corresponding queuing models for generality.

A. Receiving Module

The time to receive a BHP can be expressed as

$$t_{\mathrm{r}} = \frac{L_{\mathrm{BHP}}}{R_{\mathrm{cc}}} \tag{2}$$

where L_{BHP} is the size of the BHP, R_{cc} is the rate of the control channel.

B. Forwarder Module

According to the Kleinrock approximation [22], the complex BHP traffic flowing into the forwarder queue can be approximated as Poisson traffic. Hence, the total BHP arrival rate for the forwarder module can be expressed as

$$\lambda_{\mathrm{f}} = \sum_{i=1}^{k} \lambda_i \tag{3}$$

where λ_i is the arrival rate from import i, and k is the total number of imports.

The most critical part of the forwarding delay is route lookup time, which is quite similar to the conventional electronic router [19], [20]. The route lookup time may be a constant or variable depending on its implementation approach [20], [23], so is the corresponding forwarding delay. When the forwarding delay is a constant under hardware approach, the forwarder module can be modeled as an $\mathrm{M/D/1}$ queue, which is very helpful to configuring offset time in OBS network [20]. For variant forwarding delay, the corresponding forwarding process is typically modeled as $\mathrm{M/G/1}$ or $\mathrm{M/M/1}$ [24], [25], which is related

to the adopted routing searching algorithm, data structure of the forwarding table, the input traffic, and so on. We adopted the $\mathrm{M/M/1}$ model for the forwarder in this paper.

When the system is stable, i.e., $\rho_{\mathrm{f}} = \lambda_{\mathrm{f}}/\mu_{\mathrm{f}} < 1$, where μ_{f} is the average service rate of the forwarder, the average waiting time in the forward queue, $E(t_{q,\mathrm{f}})$, and the average forwarding delay, $E(t_{\mathrm{f}})$, can be expressed as

$$E(t_{q,\mathrm{f}}) = \frac{\rho_{\mathrm{f}}}{(\mu_{\mathrm{f}} - \lambda)} \tag{4}$$

$$E(t_{\mathrm{f}}) = \frac{1}{\mu_{\mathrm{f}}}. \tag{5}$$

C. Scheduler Module

The processing time in the scheduler depends on the adopted scheduling algorithm and implementation approach. We employed LAUC-VF (latest available unscheduled channel with void filling) scheduling algorithm in our analysis [19]. de Pedro *et al.* [21] has studied in detail the relationship between the scheduling delay and the distribution of the available void gaps for such scheduling scheme with binary search algorithm, where the scheduling delay is a variable. Hence, we characterize the scheduling delay as a random variable of exponential distribution. The simulation in Section IV shows that such treatment is valid for low traffic load.

The input rate of scheduler m is $p_m\lambda_{\mathrm{f}}$, where m corresponds to ring m, p_m is the probability that the BHP output from the forwarder module will enter scheduler module m. According to [21], the average scheduling rate (servicing rate) of the scheduler m, μ^m_{sch}, can be expressed as follows:

$$\mu^m_{\mathrm{sch}} = \frac{1}{(N \cdot T_{\mathrm{chk}})} \tag{6}$$

where T_{chk} is the time consumed to check a void gap for the arriving BDP; N is the average number of checked void gaps when scheduling a newly arrived BDP on the data channels.

Assuming the scheduled BDPs are arranged in sequence on each data channel according to the start time of the BDPs, the LAUC-VF scheduler searches each void gap along the sequence until the start time of the new BDP is met. N can be expressed as

$$N = n_{\mathrm{dc}} \cdot \int_0^{t\_\max} f(t_{\mathrm{gap}}) \int_t^{t\_\max} \lambda_{\mathrm{BDP}}(x)dxdt_{\mathrm{gap}} \tag{7}$$

where n_{dc} is the number of data channels; $t\_\max$ is the allowed maximum delay of BDPs at OBSR nodes; $f(t_{\mathrm{gap}})$ is the distribution function of the start time of the new BDP, where t_{gap} is the time gap between the start time of the new BDP and current time t_{c}; The second integral in the expression $(\int_t^{t\_\max} \lambda_{\mathrm{BDP}}(x)dx)$ is the number of scheduled BDPs on the data channels from $t_{\mathrm{c}} + t_{\mathrm{gap}}$ to $t_{\mathrm{c}} + t\_\max$, which represents the number of checked void gaps when the start time of the new BDP is $t_{\mathrm{c}} + t_{\mathrm{gap}}$. $\lambda_{\mathrm{BDP}}(t)$ is the number of the scheduled BDPs per unit time on the data channels at time $t_{\mathrm{c}} + t$. Assuming the output BDP traffic is uniformly scheduled to each data channel on ring m, $\lambda_{\mathrm{BDP}}(t)$ can be expressed as

$$\lambda_{\mathrm{BDP}}(t) = \frac{p_m\lambda}{n_{\mathrm{dc}}} \int_t^{t\_\max} f(y)dy. \tag{8}$$

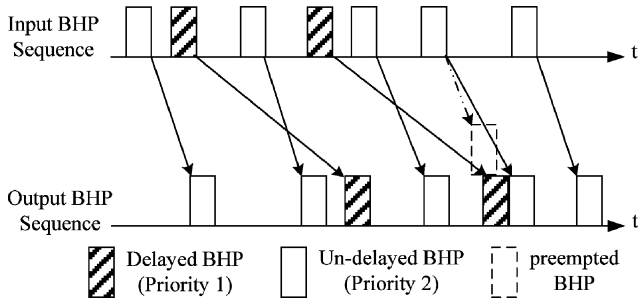

Fig. 3. Example of the input and output BHP sequence in scheduler module.

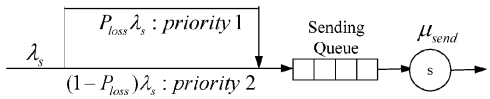

Fig. 4. $\mathrm{M/D/1}$ model with preemptive priority queue for sending module.

When the start time of the new BDP is evenly distributed between t_{c} and $t_{\mathrm{c}} + t\_\max$ (i.e., $f(t_{\mathrm{gap}}) = 1/t\_\max$), we have

$$\mu^m_{\mathrm{sch}} = \frac{6}{(p_m \cdot \lambda_{\mathrm{f}} \cdot t\_\max \cdot T_{\mathrm{chk}})}. \tag{9}$$

When the $\mathrm{M/M/1}$ system is stable, i.e.,

$$\rho^m_{\mathrm{sch}} = p_m \cdot \lambda_{\mathrm{f}}/\mu^m_{\mathrm{sch}} < 1,$$

we can obtain the average waiting time in the scheduler queue, $E(t_{q,\mathrm{sch}})$, and the average scheduling delay, $E(t_{\mathrm{sch}})$, as follows:

$$E(t_{q,\mathrm{sch}}) = \frac{\rho^m_{\mathrm{sch}}}{\mu^m_{\mathrm{sch}} - p_m \cdot \lambda_{\mathrm{f}}} = \frac{p_m^3\lambda_{\mathrm{f}}^3 t\_\max^2 T_{\mathrm{chk}}^2}{36 - 6p_m^2\lambda_{\mathrm{f}}^2 t\_\max T_{\mathrm{chk}}} \tag{10}$$

$$E(t_{\mathrm{sch}}) = \frac{1}{\mu^m_{\mathrm{sch}}} = \frac{p_m\lambda_{\mathrm{f}} t\_\max T_{\mathrm{chk}}}{6}. \tag{11}$$

D. Sending Module

If the sequence of BHPs output from the scheduler is the same as the sequence of the input BHPs to the scheduler, the input BHP traffic of the sending module will be Poisson traffic with arrival rate $\lambda_s = p_m\lambda_{\mathrm{f}}$. However, the output BHP sequence of the scheduler may be changed by delaying the corresponding BDP so as to avoid congestion in the data channels, as shown in Fig. 3. To consider this situation, we hereby adopt an $\mathrm{M/D/1}$ model with preemptive priority queue (shown in Fig. 4) to approximate this module, where the delayed BHPs for the congestion have higher priority (priority 1 in Fig. 4) than the undelayed BHPs for the congestion (priority 2 in Fig. 4), since the delayed BHPs can defer the sending time of the undelayed BHPs (see the preemption example in Fig. 3).

Let P_{loss} denotes the possibility that a BHP is delayed to avoid congestion in the data channel, the arrival rate of the delayed BHP traffic and the undelayed BHP traffic can be expressed as

$$\lambda_1 = P_{\mathrm{loss}}\lambda_s, \quad \lambda_2 = (1 - P_{\mathrm{loss}})\lambda_s. \tag{12}$$

According to the M/D/1 model with preemptive priority queue, we have

$$\begin{aligned}\lambda_s t_{s,1-2} &= \lambda_1 t_{s,1} + \lambda_2 t_{s,2} \\ t_{s,1} &= \frac{P_{\text{loss}}\lambda_s}{(2\mu_{\text{send}}(\mu_{\text{send}} - P_{\text{loss}}\lambda_s))} + \frac{1}{\mu_{\text{send}}} \\ t_{s,1-2} &= \frac{(2\mu_{\text{send}} - \lambda_s)}{(2\mu_{\text{send}}(\mu_{\text{send}} - \lambda_s))}\end{aligned} \tag{13}$$

where $t_{s,1-2}$ is the average sojourn delay (including waiting time in the queue and the service time) for all BHPs; $t_{s,1}$ and $t_{s,2}$ is the average sojourn delay for BHPs with priority 1and priority 2, respectively. μ_{send} is the servicing rate of BHP sending module ($\mu_{\text{send}} = R_{\text{cc}}/L_{\text{BHP}}$). Then, we can obtain the average queuing delay of BHPs with priority 2 as follows:

$$\begin{aligned}t_{q,2} &= t_{s,2} - \frac{1}{\mu_{\text{send}}} \\ &= \left(1 + \frac{P_{\text{loss}}}{1 - P_{\text{loss}}}\right)\left(\frac{2\mu_{\text{send}} - \lambda_s}{2\mu_{\text{send}}(\mu_{\text{send}} - \lambda_s)}\right) \\ &\quad - \frac{P_{\text{loss}}}{1 - P_{\text{loss}}}\left(\frac{2\mu_{\text{send}} - P_{\text{loss}}\lambda_s}{2\mu_{\text{send}}(\mu_{\text{send}} - P_{\text{loss}}\lambda_s)}\right) - \frac{1}{\mu_{\text{send}}}.\end{aligned} \tag{14}$$

The priority of a BHP is uncertain at OBSR nodes. Based on the rule of worst-case design, we take $t_{q,2}$ as the average delay of a BHP in the sending queue (i.e., $E(t_{q,s}) = t_{q,2}$), since BHP will experience longer average queuing delay in the sending module when it belongs to priority 2.

III. Distribution Function of Processing Delay

The possibility of burst loss caused by insufficient offset time (the shaded area in Fig. 5) can be expressed as follows:

$$\text{BLR} = P\left(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}}\right) = \int_{t_{\text{total}}}^{\infty} f_{\text{total}}(t)dt \tag{15}$$

where $P(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}})$ is the possibility that the total BHP processing delay would be larger than t_{total}; n is the number of intermediate OBSR nodes on the path; $t_{p,i}$, $(1 \leq i \leq n)$ is the processing delay at intermediate OBSR node i on the path. t_{total} is the time to compensate the total BHP processing delay on the path. Based on (15), t_{total} can be determined according to $P(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}})$ and the allowed BLR. After obtaining t_{total} [11], the offset time can be configured as follows [4]:

$$t_{\text{offset}} = t_{\text{total}} + t_{\text{switch}} = P^{-1}(\text{BLR}) + t_{\text{switch}} \tag{16}$$

where t_{switch} is the time needed to set optical switch matrix.

Generally, $P(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}})$ can be determined by the distribution function of the processing delay at each SCU. However, some approximation must be taken since the total processing delay is the sum of multiple processing delays at the SCU, which are all random variables. Three approximation methods based on central limit, large deviation, and PH distribution theorem, respectively [16]–[18], are adopted to derive $P(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}})$ from the model of a single SCU built in Section II. The static parameters needed by the SCU model like the time consumed to check a void gap and the average service rate of each forwarder can be broadcasted through the existing protocols of OBS network with minor change [19]. Dynamic BHP load information of different links can be obtained by periodically measuring the BHP arrival rate at the connected OBSR nodes [26] and exchanging the information among different nodes through either broadcast or probe-based approaches [27].

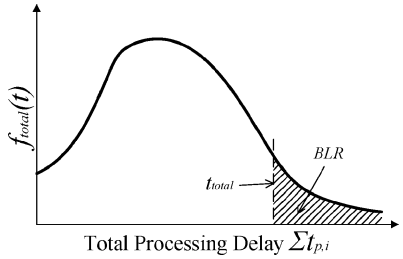

Fig. 5. Distribution function of the total processing delay.

A. Central Limit Theorem-Based Method

When n is large enough, the central limit theorem can be used. Then we have [16]

$$\begin{aligned}&P\left(\sum_{i=1}^{n} t_{p,i} \geq t_{\text{total}}\right) \\ &= \int_{t_{\text{total}}}^{\infty} \frac{1}{\sqrt{2\pi \sum_{i=1}^{n} D(t_{p,i})}} \exp\left(-\frac{\left(t - \sum_{i=1}^{n} E(t_{p,i})\right)^2}{2\sum_{i=1}^{n} D(t_{p,i})}\right) d \\ &\quad \cdot \sum_{i=1}^{n} t_{p,i}\end{aligned} \tag{17}$$

where $t_{p,i}$ is the processing delay at OBSR node i; $E(t_{p,i})$ is the average of $t_{p,i}$; $D(t_{p,i})$ is the variance of $t_{p,i}$. Based on (1), (2), (4), (5), (10), (11), and (14), $E(t_{p,i})$ can be obtained as follows:

$$\begin{aligned}E(t_{p,i}) &= \frac{1}{\mu_{\text{f}} - \lambda_{\text{f}}} + \frac{1}{\mu_{\text{sch}} - p_m\lambda_{\text{f}}} \\ &\quad + \left(\frac{1}{1 - P_{\text{loss}}}\right)\left(\frac{2\mu_{\text{send}} - p_m\lambda_{\text{f}}}{2\mu_{\text{send}}(\mu_{\text{send}} - p_m\lambda_{\text{f}})}\right) \\ &\quad - \left(\frac{P_{\text{loss}}}{1 - P_{\text{loss}}}\right)\left(\frac{2\mu_{\text{send}} - P_{\text{loss}}p_m\lambda_{\text{f}}}{2\mu_{\text{send}}(\mu_{\text{send}} - P_{\text{loss}}p_m\lambda_{\text{f}})}\right).\end{aligned} \tag{18}$$

The variance of $t_{p,i}$ can be calculated as follows:

$$\begin{aligned}D(t_{p,i}) &= D(t_{\text{r}}) + D(t_{q,\text{f}} + t_{\text{f}}) + D(t_{q,\text{sch}} + t_{\text{sch}}) + D(t_{q,s}) \\ &= \frac{1}{(\mu_{\text{f}} - \lambda_{\text{f}})^2} + \frac{1}{(\mu_{\text{sch}} - \lambda_s)^2} + D(t_{q,s}).\end{aligned} \tag{19}$$

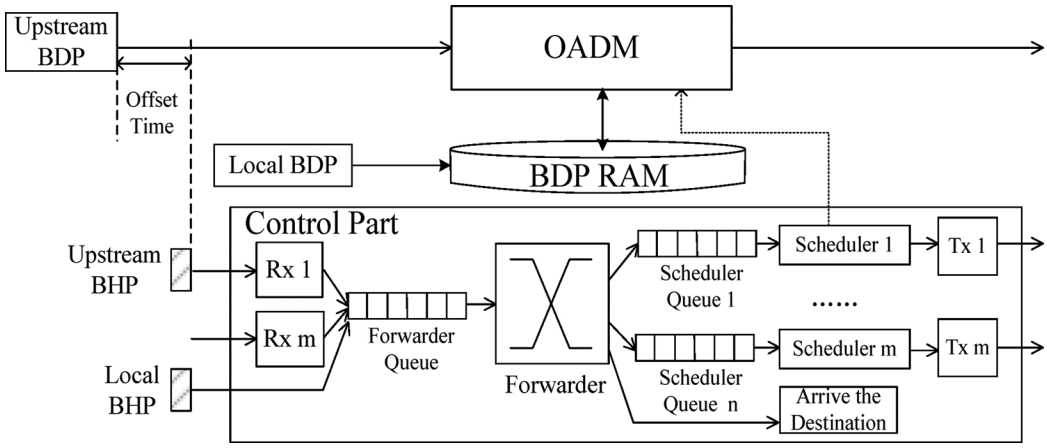

Fig. 6. OBSR node architecture for network simulation.

For normal operating OBS networks, the control channel should not become a bottleneck. The delay of a BHP in the sending queue is relatively low compared with that of the forwarder module and the scheduler module. The variance of $t_{q,s}$ compared with that of other random variables is also less significant. At the same time, the sending module is more similar to the M/D/1 model when P_{loss} is low. So $D(t_{q,s})$ can be approximately obtained by using the M/D/1 model

$$D(t_{q,s}) = \left(\frac{\sqrt{\frac{\rho_s}{3} - \frac{\rho_s^2}{12}}}{(\mu_{\text{send}}(1-\rho_s))} \right)^2 \tag{20}$$

where $\rho_s = \lambda_s/\mu_{\text{send}}$.

B. Large Deviation-Based Method

Since burst loss for insufficient offset time at the SCU of the OBSR network is rare event (i.e., BLR is very small), large deviation theory can be used to approximate the corresponding distribution function [17]

$$P\left(\sum_{i=1}^{n} t_{p,i} > t_{p,\text{total}}\right) \approx \frac{e^{-F(\theta^*)}}{\left(\sqrt{2\pi}\theta^* \cdot \sigma(\theta^*)\right)}$$

$$\sigma^2(\theta^*) = \sum_{i=1}^{n} \frac{\left(M_i''(\theta^*) \cdot M_i(\theta^*) - (M_i'(\theta^*))^2\right)}{M_i^2(\theta^*)}$$

$$F(\theta^*) = \theta^* \cdot t_{\text{total}} - \sum_{i=1}^{n} \log M_i(\theta^*), \quad \theta^* = \sum_{i=1}^{n} \frac{M_i'(\theta^*)}{M_i(\theta^*)} \tag{21}$$

where $M_i(\theta)$ is the moment generating function of the processing delay at OBSR node i. It can be obtained based on the model in Section II as $M_i(\theta) = M_{\text{f},i}(\theta) + M_{\text{sch},i}(\theta) + M_{s,i}(\theta)$, where $M_{\text{f},i}(\theta)$, $M_{\text{sch},i}(\theta)$, and $M_{s,i}(\theta)$ are the moment generating functions for the forwarder module, scheduler module and sending module, respectively, and they can be expressed based on M/M/1 and M/D/1 model as follows:

$$M_{\text{f},i}(\theta) = \frac{\mu_{\text{f}}}{(\theta + \mu_{\text{f}} - \lambda_{\text{f}})} \tag{22}$$

$$M_{\text{sch},i}(\theta) = \frac{\mu_{\text{sch}}}{(\theta + \mu_{\text{sch}} - p_m\lambda_{\text{f}})} \tag{23}$$

$$M_{s,i}(\theta) = \frac{\theta(1-\rho_s)B^*(\theta)}{(p_m\lambda_{\text{f}}B^*(\theta) + \theta - p_m\lambda_{\text{f}})} \tag{24}$$

where $\rho_s = p_m\lambda_{\text{f}}/\mu_{\text{send}}$, $B^*(\theta) = e^{-\theta/\mu_{\text{send}}}$.

$$P\left(\sum_{i=1}^{n} t_{p,i} > t_{\text{total}}\right) = \overrightarrow{\alpha} e^{T(t_{\text{total}} - n \cdot t_{\text{r}})} \overrightarrow{1}$$

$$\overrightarrow{\alpha} = \underbrace{(1,0,\ldots,0)}_{2n}, \quad \overrightarrow{1} = \underbrace{(1,1,\ldots,1)'}_{2n}$$

$$T = \begin{bmatrix} -(\mu_{\text{f},1} - \lambda_{\text{f},1}) & \mu_{\text{f},1} - \lambda_{\text{f},1} & 0 & \cdots & 0 \\ 0 & -(\mu_{\text{sch},1} - \lambda_{s,1}) & \mu_{\text{sch},1} - \lambda_{s,1} & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \cdots & 0 & -(\mu_{\text{f},n} - \lambda_{\text{f},n}) & \mu_{\text{f},n} - \lambda_{\text{f},n} \\ 0 & \cdots & \cdots & 0 & -(\mu_{\text{sch},n} - \lambda_{s,n}) \end{bmatrix} \tag{25}$$

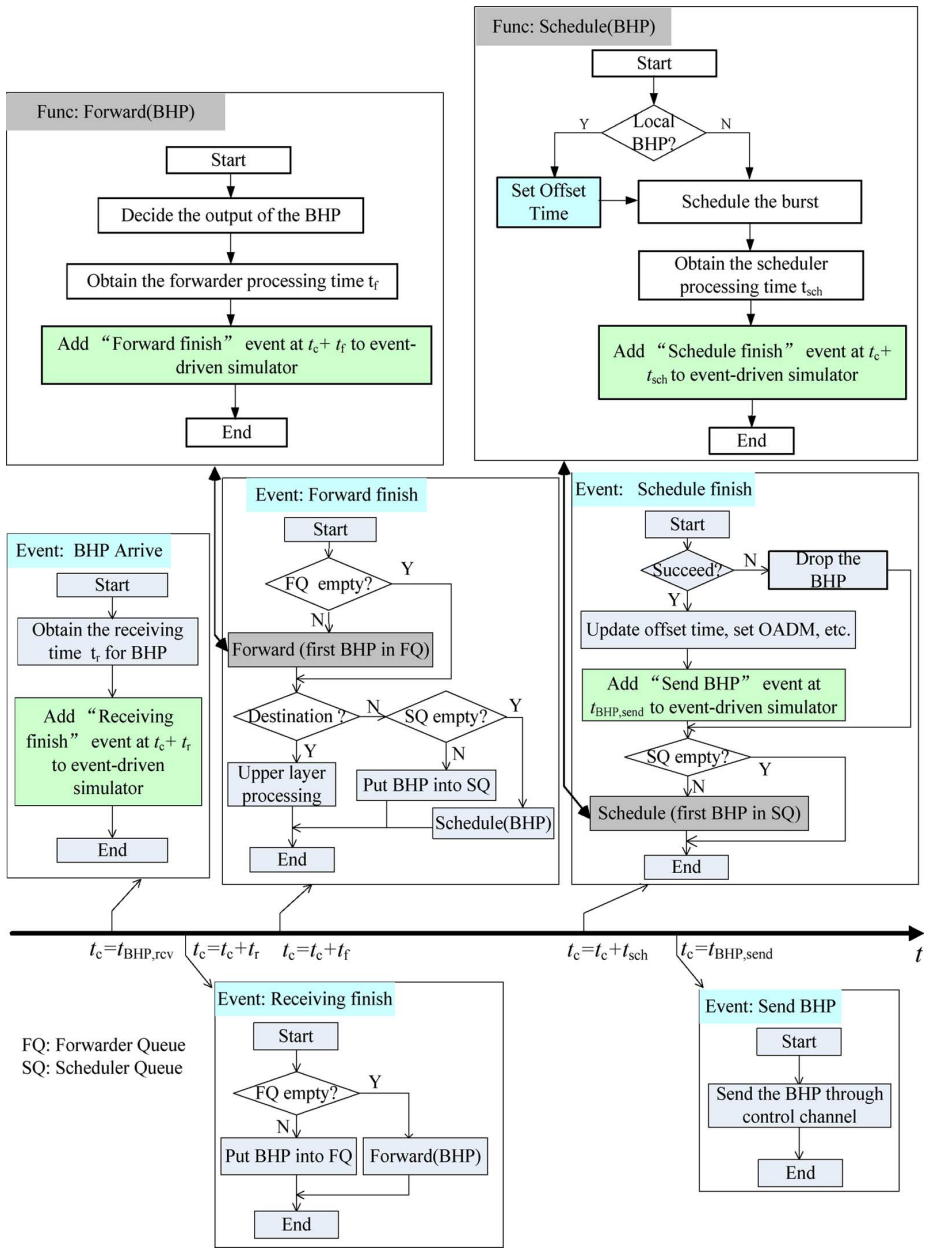

Fig. 7. BHP processing events and the corresponding time sequence in the OBSR simulation platform.

C. PH Distribution-Based Method

Since BHP delay in the sending queue is relatively small compared with that of the forwarder module and the scheduler module as mentioned earlier, $t_{q,s}$ can be neglected. So PH distribution can be adopted when the processing delay in both forwarder module and scheduler module is of exponential distribution. The distribution function of the total BHP processing delay can thus be expressed as shown in (25) at the bottom of the previous page [18], where $\mu_{\mathrm{f},i}$, $\mu_{\mathrm{sch},i}$, $\lambda_{\mathrm{f},i}$, $\lambda_{s,i}$ stands for μ_{f}, μ_{sch}, λ_{f}, λ_s at node i on the OBS path, respectively.

IV. Simulation Results and Discussion

In order to evaluate the offset time configuration mechanisms proposed in Section III, an OBSR network simulation platform is designed based on network simulator 2 [28]. It can simulate the practical BHP processing delay at the SCU. Fig. 6 shows the designed OBSR node architecture. BDP RAM is used to store locally generated BDPs and the BDPs terminated at the node. The control part, which includes receiving module, forwarder queue, forwarder, scheduler queue, scheduler, and sending module, is used to process BHPs and set the optical add/drop multiplexer.

Fig. 7 shows the main events and their processing sequence in the OBSR simulation platform. The BHP processing delay and the queuing delay are simulated by setting the start time of the finishing event of the corresponding BHP processing module. In order to focus on modules unique to OBSR network, the forwarder module in the simulation, which has been studied extensively, is the same as the model in Section III. All other modules in the simulation follow practical BHP processing procedure.

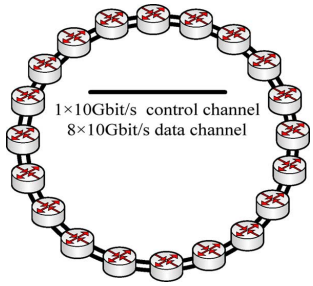

Fig. 8. OBSR network topology for simulation.

The scheduling delay set in the finishing event of schedulers is determined by the number of searched void gaps (i.e., N) and the cost per search (i.e., $T_{\rm chk}$) according to [21]. The number of searched void gaps is obtained from real simulated scheduling process. In the event of sending module, the output BHP sequence is also affected by the corresponding simulated BDP scheduling results. Each queuing delay is related to the practical input traffic flows of the network and the according processing procedure of the modules.

The simulation scenario is shown in Fig. 8, which is a bidirectional ring with 20 OBSR nodes. Shortest path first algorithm is adopted for ring selection. The link in each ring contains one control channel and eight data channels with rate of 10 Gbit/s. Each OBSR node sends Poisson traffic to all other nodes. The burst is assembled by the time and size mixed-based algorithm [29], where the maximum time is 1 ms and the maximum size is 40 kByte. The size of BHP is 64 Byte. LAUV-VF scheduling algorithm without fiber delay lines is adopted; the allowed maximum delay of BDPs (i.e., $t\_\max$) in the electronic RAM to avoid collision at the ingress edge nodes is 0.1 ms; the time consumed to check a void gap (i.e., $T_{\rm chk}$) is 20 ns; and the average service rate of each forwarder (i.e., $\mu_{\rm f}$) is 20 Mpps.

In order to check the validity of the $\rm M/M/1$ assumption on the scheduler module, the distribution of the scheduling delay obtained by simulation and the assumed exponential distribution derived from the $\rm M/M/1$ model are compared under different traffic load in Fig. 9. We can observe that when the traffic load is light, the assumed exponential distribution of the scheduling delay is quite close to the simulation results. As the traffic load becomes heavier, the distribution of the scheduling delay obtained by simulation is more close to Poisson distribution than exponential distribution. The deviation between our assumption on the schedulers and the simulation results will slightly affect the accuracy of the accordingly configured offset time for heavy input traffic load as will be explained later.

Fig. 10 shows the burst loss ratio for insufficient offset time [Fig. 10(a)] and average offset time deviation between the configured offset time and the actually spent offset time of the successfully transmitted BHPs [Fig. 10(b)], as a function of the hop number along an OBS path, when the traffic load is 6.25% and the allowed BLR for insufficient offset time is 1E-3 and 1E-4, respectively. For the purpose of comparison, results of the fixed offset time configuration mechanism are also presented, where the fixed processing time at each node are representatively taken as 2 and 1.5 times of the average processing time.

From Fig. 10, we can find that burst loss ratio for insufficient offset time under the fixed configuration mechanisms varies significantly with the hop number. As the hop number increases from 1 to 8, the corresponding burst loss ratio for insufficient offset time decreases from 0.1 to 1E-4. For the distribution function-based mechanisms that are derived from the proposed model, the burst loss ratio under different hop numbers varies in a smaller range, which can improve the fairness of OBSR network and average offset time deviation. The results can be explained as follows. For the fixed configuration mechanism, the configured offset time will deviate further from the average total processing time as the hop number increases. The gap between the configured offset time and the average total processing time for the path with n hops can be expressed as $n\times$(fixed processing time – average processing time). For example, when the fixed processing time is two times of the average processing time, the gap for 1 hop and 6 hops is one and six times of the average processing time, respectively. The larger the gap is, the less possible that the total processing time will exceed the configured offset time. Fig. 11 shows the probability distribution of the total processing time on the path [Fig. 11(a)] and the possibility that the total processing time will exceed certain configured offset time [Fig. 11(b)] in the same scenario as Fig. 10. From the figure, we can observe that the possibility to exceed the offset time configured by the fixed configuration mechanism (two times) under 6 hops is much smaller than the possibility under 1 and 3 hops. So, the burst loss ratio for insufficient offset time will decrease with the hop number under the fixed configuration mechanism. On the other hand, the distribution function-based mechanism directly calculates the offset time according to the allowed BLR for insufficient offset time and the cumulative probability function derived from the proposed model under different hop numbers. So, the corresponding burst loss ratio is restricted by the allowed BLR under different hop numbers.

At the same time, we can find from Fig. 10(a) that the large deviation method-based and PH distribution method-based mechanisms meet the allowed BLR better than the central limit method-based mechanism under different hop numbers, except that the burst loss ratio for insufficient offset time decreases slightly with the hop number. It is because that the tail probability of the scheduling delay in the assumed $\rm M/M/1$ model is larger than that of the simulated scheduling delay, whose maximum value is only 0.22 μs, see Fig. 9(a). The gap of the tail probability between the simulated scheduling delay and the assumed scheduling delay will become even larger as the hop number increases, which will make the derived threshold of BHP total processing time (i.e., $t_{\rm total}$) larger than the expected threshold. Then the burst loss ratio for insufficient offset time will decrease, since the gap between the configured offset time and the expected threshold increases with the hop number. Central limit method-based mechanism does not meet the allowed BLR, especially under small hop numbers. It is reasonable since central limit theorem is only effective for a

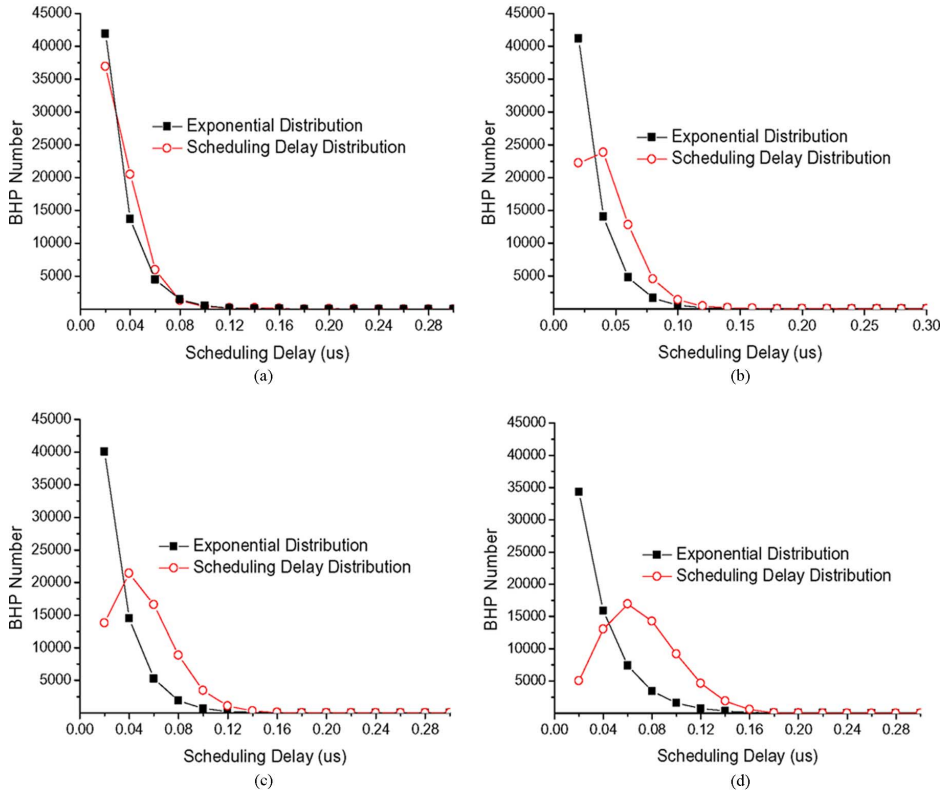

Fig. 9. Distribution of the simulated scheduling delay and the exponential distribution derived from the $\mathrm{M}/\mathrm{M}/1$ model when the traffic load is (a) 6.25%, (b) 12.5%, (c) 18.75%, and (d) 31.25%.

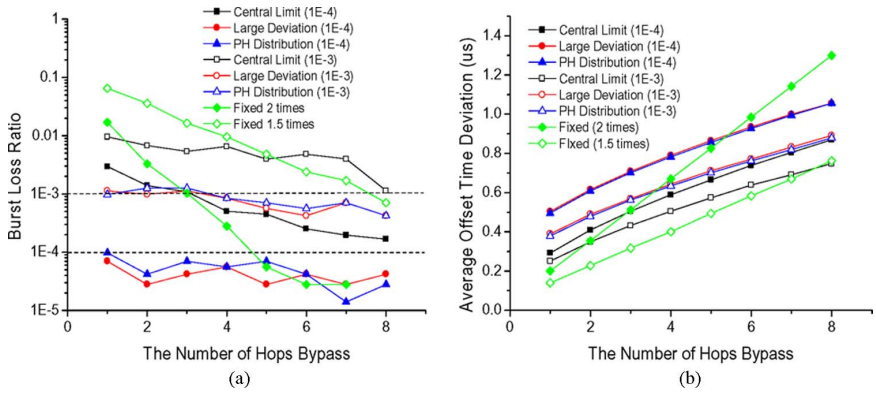

Fig. 10. Dependence of (a) burst loss ratio for insufficient offset time and (b) offset time deviation on the hop number along an OBSR path when the traffic load is 6.25%.

larger number of the hops, and the hop number along the path in Fig. 8 is relatively small, indicating insufficient independent random variables for the total processing time. We can also find that the offset time deviation will increase with the hop number along the path [see Fig. 10(b)]. It is because the variance of the total processing time on the path increases greatly as the hop number on the path increases [see Fig. 11(a)].

Fig. 12 shows the burst loss ratio for insufficient offset time [Fig. 12(a)] and average offset time deviation [Fig. 12(b)] as a function of the hop number when the traffic load is 31.25%. Similar results as Fig. 10 can be observed from the figure. At the same time, the burst loss ratio for insufficient offset time under all three approximation methods increases a little with the hop number. It can be explained as follows. From Fig. 9(d), we can observe that under heavy input traffic load, the simulated scheduling delay distribution cannot be perfectly approximated by the assumed exponential distribution. The average value of the simulated scheduling delay is larger than that of assumed exponential distribution; the variances of both distributions are very close to each other. As the hop number increases, the average value of the simulated scheduling delay will become even larger than that of the scheduling delay from the assumed model, while the variances of both distributions are still close to each other. So the simulated scheduling delay is more likely to become larger than the expected scheduling delay from the assumed model as the hop number increases, and the burst loss ratio for insufficient offset time will also gradually increase accordingly.

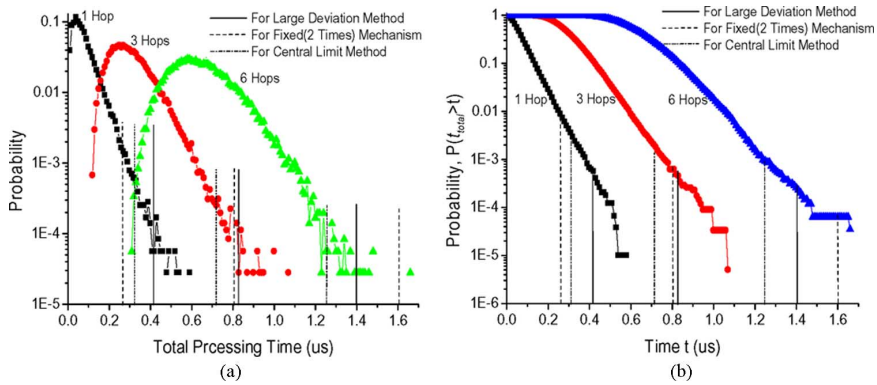

Fig. 11. Probability distribution of (a) the total processing delay on the path and (b) the probability that the total processing time is large than time t when the traffic load is 6.25%.

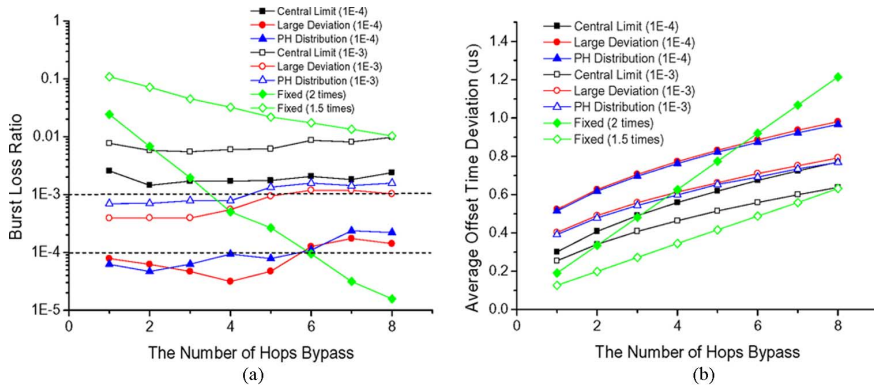

Fig. 12. Dependence of (a) burst loss ratio for insufficient offset time and (b) offset time deviation on the hop number along an OBSR path when the traffic load is 31.25%.

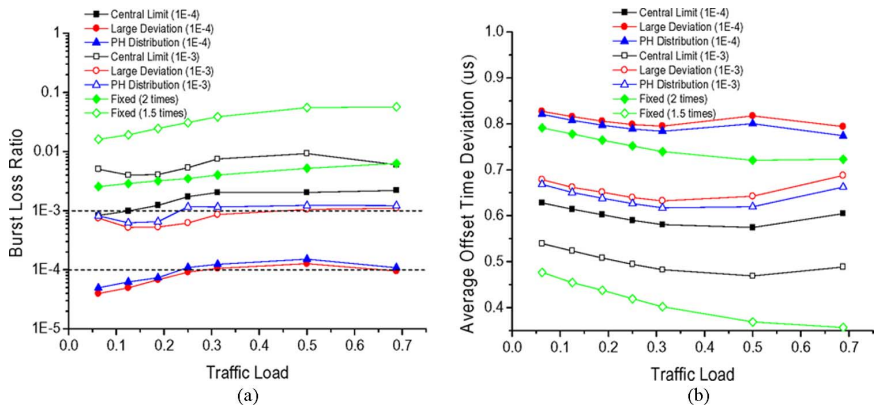

Fig. 13. Dependence of (a) overall burst loss ratio for insufficient offset time and (b) overall offset time deviation in the OBSR network on the traffic load.

Fig. 13 shows the dependence of overall burst loss ratio for insufficient offset time [Fig. 13(a)] and overall offset time deviation [Fig. 13(b)] for all BHPs in the OBSR network on the input traffic load. The figure shows that large deviation method and PH distribution method-based mechanisms can meet the requirement of the allowed BLR under different traffic load. Although their burst loss ratio for insufficient offset time increases a little at first with the traffic load due to the simplification and assumption mentioned earlier, it gradually stabilizes on the allowed BLR when the traffic load is large enough. Besides, offset time deviation almost does not change with the traffic load under the proposed distribution function-based mechanisms [see Fig. 13(b)] but decreases slightly with the traffic load under the fixed configuration mechanism. It is because the burst loss ratio for insufficient offset time is limited by the allowed BLR under the proposed mechanisms but not under the fixed configuration mechanism. We can also find from Fig. 13 that fixed configuration mechanism (two times) is less efficient because it has both higher burst loss ratio and offset time deviation than central limit method (1E-4), large deviation method (1E-3), and PH distribution method-based mechanisms (1E-3). At the same time, fixed configuration mechanism (1.5 times) is also not desirable since its burst loss ratio is too high to be tolerable.

V. Conclusion

In this paper, we develop a BHP processing model for the SCU of an OBSR network. The receiving module, forwarder module, scheduler module, and sending module are analyzed in order to characterize the BHP processing time. Three methods based on central limit, large deviation, and PH distribution theorem, respectively, are adopted to derive the distribution function of the total BHP processing time, so that offset time can be configured according to the derived distribution function and the allowed BLR for insufficient offset time. A simulation platform that takes the practical BHP processing time into account is deigned to evaluate the offset time configuration mechanisms.

The simulation results show that the fixed configuration mechanism tends to configure insufficient offset times for BHPs bypassing smaller number of hops, and too long offset times for BHPs bypassing larger number of hops, which lead to a great difference on burst loss ratio for insufficient offset time under different hop numbers. Offset time configuration mechanisms based on the proposed model and the distribution functions can configure appropriate offset times, which make the burst loss ratio for insufficient offset time more close to the allowed BLR with proper burst transmission delay under different hop numbers, especially when large deviation method and PH distribution method-based mechanisms are adopted.

References

[1] C. Qiao and M. Yoo, "Optical burst switching (OBS)-A new paradigm for an optical internet," *J. High Speed Netw.*, vol. 8, no. 1, pp. 69–84, 1999.

[2] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: A viable solution for terabit IP backbone," *IEEE Netw.*, vol. 14, no. 6, pp. 48–53, Nov./Dec. 2000.

[3] X. W. Li, J. P. Chen, G. L. Wu, H. Wang, and A. Ye, "An experimental study of an optical burst switching network based on wavelength-selective optical switches," *IEEE Commun. Mag.*, vol. 43, no. 5, pp. S3–10, May 2005.

[4] J. Teng and G. Rouskas, "A comparison of the JIT, JET and horizon wavelength reservation schemes on a single OBS node," in *First Int. Workshop Opt. Burst Switching*, USA, 2003.

[5] M. Kfinkowskil, D. Careglio, and J. So-Pareta, "Modeling of control part in OBS Networks," in *9th Int. Conf. Transparent Opt. Netw. (ICTON)*, Rome, Italy, Jul. 1–5, 2007.

[6] N. Barakat and T. E. Darcie, "Control-plane congestion in OBS networks," in *3rd Int. Conf. Broadband Commun., Netw. Syst. (BROADNETS)*, San Jose, CA, Oct. 2006.

[7] N. Barakat and T. E. Darcie, "The control-plane stability constraint in optical burst switching networks," *IEEE Commun. Lett.*, vol. 11, pp. 267–269, Mar. 2007.

[8] J. Y. Choi, J. S. Choi, and M. Kang, "Dimensioning burst assembly process in optical burst switched networks," *IEICE Trans. Commun.*, vol. E88-B, no. 10, pp. 3855–3863, 2005.

[9] Y. Chi, H. Junbin, L. Zhengbin, and X. Anshi, "A novel burst assembly algorithm for OBS networks based on data-length time-lag product," in *Proc. 2005 Asia-Pacific Conf. Commun.*, Oct. 2005, pp. 319–323.

[10] I.-Y. Hwang, S. Lee, and H.-S. Park, "Impact of burst control packet congestion on burst loss rate in optical burst switched networks," in *Proc. Int. Conf. Inform. Netw. (ICOIN)*, Jan. 2006, pp. 369–378.

[11] A. E. Martinez, J. Aracil, and J. E. Lopez de Vergara, "Optimizing offset times in optical burst switching networks with variable burst control packets sojourn times," *Opt. Switching Netw.*, vol. 4, no. 3–4, pp. 189–199, 2007.

[12] J. Segarra, V. Sales, and J. Prat, "An all-optical access-metro interface for hybrid WDM/TDM PON based on OBS," *IEEE/OSA J. Lightw. Technol.*, vol. 25, no. 4, pp. 1002–1016, Apr. 2007.

[13] S. Sheeshia, Y. Chen, V. Anand, and C. Qiao, "Performance comparison of OBS and SONET in metropolitan ring networks," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1474–1482, Oct. 2004.

[14] J. Kim, J. Cho, S. Das, D. Gutierrez, M. Jain, C.-F. Su, R. Rabbat, T. Hamadaand, and L. G. Kazovsky, "Optical burst transport: A technology for the WDM metro ring networks," *IEEE/OSA J. Lightw. Technol.*, vol. 25, no. 1, pp. 93–102, Jan. 2007.

[15] J. M. Finochietto, J. Aracil, A. Ferreiro, G. Fernandez-Palacios, and J. P. Gonzalez de Dios, "Migration strategies toward all optical metropolitan access rings," *IEEE/OSA J. Lightw. Technol.*, vol. 25, no. 8, pp. 1918–1930, Aug. 2007.

[16] S. Kaczmarek and M. Narloch, "Methods for evaluation packet delay distribution of flows using expedited forwarding PHB," in *Proc. 2nd Polish-German Teletraffic Symp. (PGTS 2002)*, Sep. , pp. 85–94.

[17] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Boston, MA: Jones & Bartlet, 1992.

[18] C. A. O'Cinneide, "Characterization of phase-type distributions," *Stochastic Models*, vol. 6, no. 1, pp. 1–57, 1990.

[19] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Control architecture in optical burst switched WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1838–1851, Oct. 2000.

[20] P. Mehrotra and P. D. Franzon, "Novel hardware architecture for fast address lookups," *IEEE Commun. Mag.*, vol. 40, no. 11, pp. 66–71, Nov. 2002.

[21] L. de Pedro, J. Aracil, J. A. Hernandez, and J. L. Garcia-Dorado, "Analysis of the processing and sojourn times of burst control packets in optical burst switches," in *Int. Conf. Opt. Netw. Design Model. (ONDM)*, Mar. 2008, pp. 1–3.

[22] D. Bertse and R. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 1992.

[23] M. Degermark, A. Brodnik, S. Carlsson, and S. Pink, "Small forwarding tables for fast routing lookups," in *Proc. ACM SIGCOMM 1997*, pp. 3–14.

[24] N. Gulpinar, P. Harrison, B. Rustem, and L.-F. Pau, "An optimization model for a two-node router network," in *Proc. IEEE Comput. Society's 12th Annu. Int. Symp. Model., Anal., Simul. Comput. Telecommun. Syst. (MASCOTS 2004)*, Oct. , pp. 147–156.

[25] L. Zhiyong, X. Ke, and W. Jianping, "A novel model to analyze the performance of routing lookup algorithms," in *Proc. Int. Conf. Commun. Technol. (ICCT 2003)*, Apr. , vol. 1, pp. 508–513.

[26] S. K. Tan, G. Mohan, and K. C. Chua, "Link scheduling state information based offset management for fairness improvement in WDM optical burst switching networks," *Comput. Netw.*, vol. 45, no. 6, pp. 819–834, 2004.

[27] G. Thodime, V. Vokkarane, and J. Jue, "Dynamic congestion-based load balanced routing in optical burst switched networks," in *Proc. IEEE GLOBECOM*, Dec. 2003, vol. 5, pp. 2694–2698.

[28] Network Simulator NS-2 [Online]. Available: http://www.isi.edu/nsnam/ns

[29] X. Cao, J. Li, Y. Chen, and C. Qiao, "Assembling TCP/IP packets in optical burst switched networks," in *Proc. IEEE GLOBECOM*, Taipei, Taiwan, Nov. 2002, pp. 2808–2812.

Wei Dai received the B.S. and M.S. degrees from Shanghai Jiao Tong University, Shanghai, China, in 2006 and 2009, respectively.

His recent research interests include optical networking and optical burst switching.

Guiling Wu received the B.S. degree from Ha Er Bing Institute of Technology, Harbin, China, in 1995, and the M.S. and Ph.D. degrees from Huazhong University of Science and Technology, Wuhan, China, in 1998 and 2001, respectively.

Currently, he is an Associate Professor at Shanghai Jiao Tong University, Shanghai, China. His recent research interests include optical networking and high-speed optical signal processing.

Wenjun Qian received the B.S. and M.S. degrees from Shanghai Jiao Tong University, Shanghai, China, in 2004 and 2007, respectively.

Currently, she is working at China Mobile Group Shanghai Co., Ltd., Shanghai, China.

Xinwan Li received the B.S. degree from Soochow University, Suzhou, China, in 1990, M.S. degree from Shanghai University, Shanghai, China, in 1993, Ph.D. degree from Shanghai Jiao Tong University, Shanghai, in 2005.

From 1997 to 1998 he was a Research Assistant at Essex University. In 2001, he joined OPCOM Inc. as an Engineer. Currently, he is a Professor at Shanghai Jiao Tong University. His recent research interests include optical switching technologies and advanced optical fiber components.

Dr. Li is a senior member of IEEE LEOS and ComSoc.

Jianping Chen received the B.S. degree from Zhejiang University, Hangzhou, China, in 1983, and the M.S. and Ph.D. degrees from Shanghai Jiao Tong University, Shanghai, China, in 1986 and 1992, respectively.

Currently, he is a Professor at Shanghai Jiao Tong University. His recent research interests include photonic devices and subsystems, optical networking, and sensing optics.