行政院國家科學委員會專題研究計畫　期中進度報告

子計畫一：光纖都會核心網路技術研究(1/3)

計畫類別： 整合型計畫
計畫編號： NSC92-2213-E-009-117-
執行期間： 92 年 08 月 01 日至 93 年 07 月 31 日
執行單位： 國立交通大學資訊工程學系

計畫主持人： 楊啟瑞

報告類型： 精簡報告

處理方式： 本計畫可公開查詢

中　華　民　國 93 年 5 月 25 日

行政院國家科學委員會專題研究計畫成果報告
光纖都會核心網路技術研究

Optical Metro Core Networking Technology

計畫編號 : NSC92-2213-E-009-117
執行期限 : 92 **年** 8 **月** 1 **日至** 93 **年** 7 **月** 31 **日**

主持人 : **楊啟瑞　　國立交通大學資訊工程研究所**

一、中文摘要

全光近屬封包交換系統(OCPS)**已被提出來解全光封包交換系統的限制，藉由使用內頻控制各個** burst**之交換技術並採用訊務控制技術以達成頻寬高使用率及服務品質。在這篇報告中，我們提出一套強化服務品質之訊務控制機制，其使用在**OCPS**網路之入口入由器以提供延遲與遺失等級區分。此機制命名為**(y,t)-Scheduler/Shaper**以顯示其具備雙重目標，此處**y**與**t**分別代表最大的**burst**尺寸與**burst**組合時間。針對延遲方面，**(y,t)-Scheduler**進行封包排程並依據不同延遲關聯權重來組合，以達到不同等級的**99%**延遲界限保證。針對遺失率方面，**(y,t)-Shaper**配置較大的**burst**尺寸給較高遺失優先權的等級。利用整合後的封包模組化為具備批次抵達之**two-state Markov Modulated Bernoulli Process (MMBP)，**我們分析導出**(y,t)-Shaper**的輸出程序。我們亦藉由網路模擬程式來描述** OCPS **與** Just-Enough-Time (JET)-based OBS**在封包遺失率之比較。模擬結果顯示具有**(y,t)-Scheduler/Shaper**之** OCPS**於封包遺失率之表現，比**OBS**要好，特別是高優先權之等級。**

關鍵詞：都會核心網路，全光封包交換系統，全光巨量交換系統，服務品質，訊務排程，訊務流量調節，MMBP**，輸出程序。**

二、英文摘要

Optical Coarse Packet Switching (OCPS) has been proposed to circumvent optical packet switching limitations by using in-band-controlled per-burst switching and advocating traffic control to achieve high bandwidth utilization and Quality-of-Service (QoS). In this report, we present a QoS-enhanced traffic control scheme exerted at ingress nodes, aiming at providing delay and loss class differentiations for OCPS networks. Serving a dual purpose, the scheme is called (y,t)-Scheduler/Shaper, where y and t are the maximum burst size and burst assembly time, respectively. For delay, (y,t)-Scheduler performs packet scheduling and assembly according to delay-associated weights, achieving different classes of 99%-delay-bound guarantees. For loss, (y,t)-Shaper assigns larger burst sizes to higher loss priority classes. We analytically derive the departure process of (y,t)-Shaper with the aggregate packet arrivals modeled as a two-state Markov Modulated Bernoulli Process (MMBP) with batch arrivals. We also conduct network-wide simulations to draw packet loss comparisons between OCPS and Just-Enough-Time (JET)-based OBS. Simulation results demonstrate that OCPS with (y,t)-Scheduler/Shaper outperforms OBS in packet loss probabilities particularly for high priority classes.

Keywords: Metro Core networks, Optical Packet Switching (OPS), Optical Burst Switching (OBS), Quality-of-Service (QoS), Traffic scheduling, Traffic shaping, Markov Modulated Bernoulli Process (MMBP), departure process.

三、緣由與目的

Optical Packet Switching (OPS) technologies [1-3] enable fine-grained channel allocation and have been envisioned as an ultimate solution for data-centric mesh-based metro core networks. Nevertheless, OPS currently faces some technological limitations, such as the lack of optical signal processing and optical buffer technologies, and large switching overhead. In light of this, while some work [1,3] directly confronts the OPS limitations, others attempt to tackle the problem by exploiting different switching paradigms, in which Optical Burst Switching (OBS) [4-9] has received most attention.

OBS [4] was originally designed to efficiently support all-optical bufferless [5,6] networks while circumventing OPS limitations. By adopting per-burst switching, OBS requires IP packets to be first

assembled into bursts at ingress nodes. The most common packet assembly schemes are based on timer [4], packet-count threshold [6], and a combination of both [6,8]. Essentially, major focuses in OBS have been on one-way out-of-band wavelength allocation (e.g., Just-In-Time (JIT) [7], and Just-Enough-Time (JET) [5]), and the support of QoS for networks without buffers [5,6] or with limited Fiber-Delay-Line (FDL)-based buffers [9]. Particularly in the JET-based OBS scheme that is considered most effective, a control packet for each burst payload is first transmitted out-of-band, allowing each switch to perform configuration before the burst arrives. Accordingly, a wavelength is reserved only for the duration of the burst. Without waiting for a positive acknowledgment from the destination node, the burst payload follows its control packet immediately after a predetermined offset time, which is path (hop-count) dependent and theoretically designated as the sum of intra-nodal processing delays.

In the context of supporting QoS in bufferless OBS networks, the work in [5] employs a prioritized extra offset-time method. Namely, a high loss priority class is given a larger extra offset time, allowing the high priority class to make earlier wavelength reservation than lower priority classes. The method effectively provides different grades of loss performance, but at the expense of a drastic increase in the end-to-end delay particularly for high priority classes. Besides, the method undergoes the unfairness and near-far problems [10]. Especially due to the near-far problem, a low priority burst with a longer path to travel may end up with the same or larger offset time than that of a high priority burst, resulting in obstacles to QoS burst truncation in switching nodes. For example, consider a case that there is a high priority burst that arrives after a low priority burst and potentially collides with the low priority burst. If the control packet of the low priority burst has already departed, its length can no longer be updated. In this case, the switching node is left no choice but to truncate the high priority rather than the low priority burst. We refer to this as *restricted* QoS burst truncation. The prioritized burst segmentation approach proposed in [6] adopts the assembly of different priority packets into a burst in the order of decreasing priorities. Should contention occur in switching nodes, the approach supports burst truncation rendering lower-priority packets toward the tail be dropped or deflected with higher probability. The approach achieves low packet loss probability for high priority classes, with the price of excessive complexity paid during burst scheduling in switching nodes.

The above OBS design complications are the primary motivators behind the design of the OCPS paradigm [11]. While OBS can be viewed as a more efficient variant of OCS; OCPS can be considered as a less stringent variant of OPS. Similar to OBS, OCPS is aimed at supporting all-optical per-burst switched networks, which are labeled-based, QoS-oriented, and either bufferless or with limited FDL-based buffers. Unlike OBS, OCPS adopts in-band control in which the header and payload are together transported via the same wavelength. More specifically, in an OCPS network, IP packets belonging to the same loss class and the same destination are assembled into bursts at ingress routers. A header for a burst payload, which carries forwarding (i.e., label) and QoS (e.g., priority) information, is modulated with the payload based on our newly designed Superimposed Amplitude Shift Keying (SASK) technique [12]. They are time-aligned during modulation via necessary padding added to the header. They are re-aligned in switching nodes should burst truncation occur. Such design eliminates the payload length information from the header, and as will be shown, facilitates restriction-free QoS burst truncation in switching nodes. The entire burst is then forwarded along a pre-established Optical Label Switched Path (OLSP). At each switching node, the header and payload are SASK-based demodulated [12]. While the header is electronically processed, the burst payload remains transported optically in a fixed-length FDL achieving constant delay and data transparency.

The main focus of the report is on QoS-enhanced traffic control exerted during packet burstification at ingress nodes, aiming at providing delay and loss class differentiations for OCPS networks. In our work, we assume optical switches are buffer-less and all wavelengths are shared using wavelength converters. Regarding delay performance, due to the absence of buffering delay in core switches, the end-to-end delay performance is solely determined by the burstification delay. Considering the assembly of packets from flows with different delay requirements, the problem becomes the scheduling of these packets during burstification. At first thought, existing scheduling disciplines reported in [13] are possible candidates. These schemes have placed emphasis on the design of scalable *packet* schedulers achieving fairness and

delay guarantees. All packets follow the exact departure order that is computed according to virtual finishing times being associated with packets. Nevertheless, in the case of burstification, considering tens or hundreds of packets in a burst, the exact position of packets within a burst is no longer relevant. Most existing scheduling schemes thus become economically unviable. Regarding loss performance, rather than exploring reactive contention resolution mechanisms [10], in this work we focus on the design of traffic shaping with QoS provisioning.

In this report, we present a dual-purpose traffic control scheme, called (y,t)-Scheduler/Shaper. Notice that from the packet burstification perspective, it is simply a timer and threshold combined scheme, where y and t are the maximum burst size (packet count) and maximum burst assembly time, respectively. To provide delay class differentiation, for IP packet flows designated with delay-associated weights, (y,t)-Scheduler performs packet scheduling and assembly into bursts based on their weights and a *virtual window* of size y. The Scheduler exerts simple FIFO service within the window and assures weight-proportional service at the window boundary. The scheme, as will be shown, provides different classes of 99% delay bound guarantees.

To provide loss class differentiation, (y,t)-Shaper facilitates traffic shaping with a larger burst size (y) assigned to a higher priority class. To examine the shaping effect on loss performance, we analytically derive the departure process of (y,t)-Shaper. The aggregate packet arrivals are modeled as a two-state Markov Modulated Bernoulli Process (MMBP) with batch arrivals. Analytical results delineate that (y,t)-Shaper yields substantial reduction in the Coefficient of Variation (CoV) of the burst inter-departure time. The greater the burst size, the more reduction in the CoV. Furthermore, we conduct network-wide simulations to draw loss performance comparisons between OCPS and JET-based OBS. Simulation results demonstrate that OCPS with (y,t)-Scheduler/Shaper outperforms OBS in packet loss probabilities particularly for high priority classes.

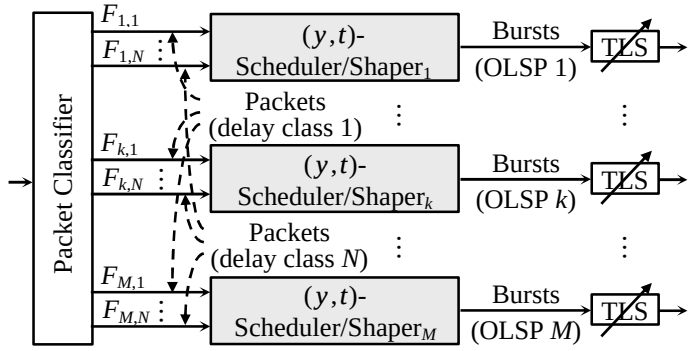
The remainder of this report is organized as follows. In Section 4.1, we introduce the (y,t)-Scheduler/Shaper system architecture. In Section 4.2, we describe the (y,t)-Scheduler design and a 99% delay bound guarantee for each delay class. In Section 4.3, we present a precise departure process analysis for (y,t)-Shaper to analytically delineate the shaping effect on departing traffic characteristics. In Section 4.4, we demonstrate the provision of loss class differentiation, and draw packet loss comparisons between OCPS and JET-based OBS via network-wide simulation results. Finally, concluding remarks are made in Section 4.5.

四、系統架構、結果與討論

4.1. (y,t)-Scheduler/Shaper System Architecture

In any ingress node, incoming packets (see Figure 1) are first classified on the basis of their destination, loss and delay classes. Packets belonging to the same destination and loss class are assembled into a burst. Thus, a burst may contain packets belonging to different delay classes. In the figure, we assume there are *M* destination/loss classes and *N* delay classes in the system. For any one of *M* destination/loss classes, say class *k*, packets of flows belonging to *N* different delay classes are assembled into bursts through (y,t)-Scheduler/Shaper$_k$ according to their pre-assigned delay-associated weights. Departing bursts from any (y,t)-Scheduler/Shaper are optically transmitted, and forwarded via their corresponding, pre-established OLSP.

Essentially, (y,t)-Scheduler/Shaper is a dual-purpose scheme. It is a scheduler for packets, abbreviated as **(y,t)-Scheduler**, which performs the scheduling of different delay class packets into back-to-back bursts. On the other hand, it is a shaper for bursts, referred to as **(y,t)-Shaper**, which determines the sizes and departure times of bursts. They are discussed in Sections 4.2 and 4.3, respectively.



Legend:

$F_{d,y}$: Packet flow of destination/loss class *d* and of delay class *y*;

OLSP : Optical Label Switched Path;

TLS : Tunable Laser Source;

Figure 1. (y,t)-Scheduler/Shaper system architecture.

4.2. (y,t)-Scheduler and Delay QoS

In the (y,t)-Scheduler system, each delay class is associated with a pre-determined weight. A higher delay priority class is given a greater weight, which corresponds to a more stringent delay bound requirement. In addition, we assume all packets are of fixed size of one unit.

4.2.1. Scheduling Design

Upon arriving, packets of different classes are sequentially inserted in a sequence of virtual windows. The window size, which is set as the maximum burst size, y, together with the weight (w) of a class, determines the maximum number of packets (i.e., quotas) from this class that can be allocated in a window. For a class, if there are sufficient quotas, its new packets are sequentially placed in the current window in a FIFO manner. Otherwise, its packets are placed in an upward window in accordance to the total accumulated quotas. A burst is formed and departs when the burst size reaches y or the Burst Assembly Timer (BATr) (set as t initially) expires. For convenience, class weights are normalized to the window size. Namely, $\sum w_i = y$ where w_i is the normalized weight of class i. The operation of (y,t)-Scheduler is shown via an example depicted in Figure 2. Notice that, at the end of time 4, there are four packets in the system, which are placed in three consecutive virtual windows. A burst is then generated at the end of time 4. This explains why the "virtual" window is named.

Assumptions:
y = 4, t = 3;
Three classes: X, Y, Z; $w_X : w_Y : w_Z = 2 : 1 : 1$;

Time	Packet Arrival	Virtual-Window Queue	BATr	Burst Departure
1	$Z_1Y_2Y_1X_2X_1$	Y_2 | $Z_1Y_1X_2X_1$	A_a→3	$Z_1Y_1X_2X_1$
2	$Z_2X_4X_3Y_4Y_3$	Y_4 | Y_3 | $Z_2X_4X_3Y_2$	R_d→3	$Z_2X_4X_3Y_2$
3		Y_4 | Y_3	R_d→3	
4	Y_5Z_3	Y_5 | Y_4 | Z_3Y_3	2	$Y_5Y_4Z_3Y_3$
5	Z_4	Z_4	A_a→3	
6	Z_5	Z_5 | Z_4	2	
7		Z_5 | Z_4	1	
8	X_5	Z_5 | X_5Z_4	0	$Z_5X_5Z_4$

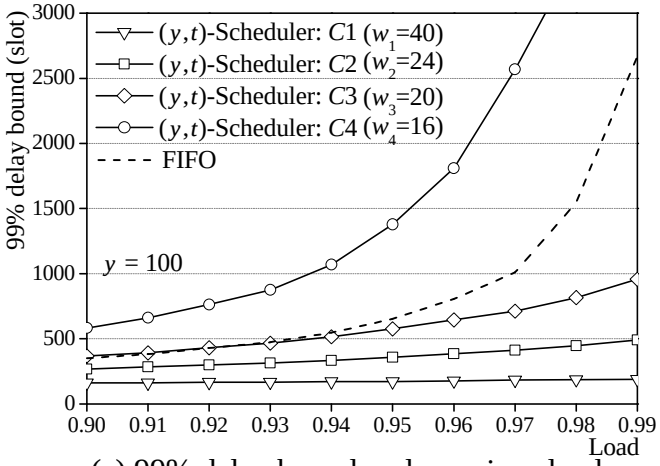
Legend:
C_n: The n_{th} packet of class C;
A_a: Activated by the first packet arrival;
R_d: Reset by burst departure;

Figure 2. (y,t)-Scheduler: an example.

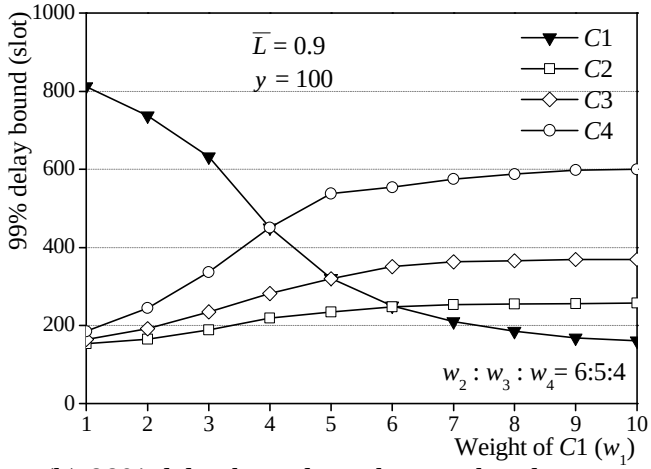
4.2.2. Delay QoS Provision

We carried out event-based simulations in which the 99% delay bound (in units of slots) were measured. In the simulations, we have four delay classes (C1-C4), with the weights set as 10, 6, 5, and 4 (or 40, 24, 20, and 16, normalized with respect to y = 100). The system is served by a wavelength in a capacity of one 60-byte packet per slot time. Each of these four classes generate an equal amount of traffic based on a two-state (H and L) MMBP. In the MMBP, the probability of switching from state H (L) to L (H) is equal to $a = 0.225$ ($b = 0.025$), and the probability of having one packet arrival during state H (L) is equal to $\bar{L}$ ($\bar{L}/6$), under an offered load, $\bar{L}$. Accordingly, the burstiness of traffic is B = 4. To draw a comparison, a FIFO system was also experimented. Simulations are terminated after reaching 95% confidence interval. Simulation results are plotted in Figure 3.

We observe from Figure 3(a) that 99% delay bound of all classes increase with the offered load. Superior to the FIFO system that undergoes long delay/bound at high loads, (y,t)-Scheduler invariably



(a) 99% delay bound under various loads



(b) 99% delay bound via the weight adjustment

Figure 3. Delay QoS provision.

assures low delay/bound for high priority classes (e.g., $C1$ and $C2$) at a cost of increased delay/bound for low priority classes (e.g., $C4$). In Figure 3(b), we illustrate how the weight of a class can be adjusted to meet its delay bound requirements. For example, to meet a 99% delay bound guarantee of 200 slots for class $C1$, the weight of $C1$ must be greater than 7, given the weights of three other classes of 6, 5, and 4, respectively.

4.3. (y,t)-Shaper and Departure Process Analysis

In (y,t)-Shaper, a burst of size y is generated and transmitted if the total number of packets reaches y before the burst assembly time exceeds t. Otherwise, a burst of size less than y is generated when BATr expires. The BATr is initialized as the t value when it is *activated* or *reset*. The BATr is activated when the system is changed from being idle to busy due to new packet arrivals. The BATr is immediately *reset* when a burst departs leaving behind a non-empty queue.

4.3.1. Departure Process Analysis

In a (y,t)-Shaper system, bursts are served (transported) by one wavelength and forwarded via the same OLSP. In the analysis, we consider (y,t)-Shaper a discrete-time single-server queueing system, MMBP/G/1, in which a time slot is equal to the transmission of a fixed-length packet. The aggregate packet arrivals are assumed to follow a two-state (H and L) MMBP that allows batch arrivals at each state. The MMBP is characterized by four parameters (a, b, l_H, l_L), where a (b) is the probability of changing from state H (L) to L (H) in a slot, and l_H (l_L) represents the probability of having a batch arrival at state H (L). For ease of description, the state change probability is denoted as $P_{i,j}$, $i, j \in \{H, L\}$. Namely, $P_{H,L} = 1 - P_{H,H} = a$, and $P_{L,H} = 1 - P_{L,L} = b$. The batch sizes at state H and L possess distributions $b_H(m)$ and $b_L(m)$, with mean sizes $\bar{b}_H$ and $\bar{b}_L$ respectively. Let $\bar{L}$ represent the mean arrival rate (packets/slot) (i.e., the load), and B the burstiness of the arrival process, we thus have

$$B = \frac{l_H \bar{b}_H}{\bar{L}} = \frac{l_H \bar{b}_H}{\frac{b}{a+b} \cdot l_H \bar{b}_H + \frac{a}{a+b} \cdot l_L \bar{b}_L} \qquad (1)$$

There are five possible events that sequentially occur in a slot as follows: (1) arrival process state change, (2) begin-of-burst departure, (3) packet arrivals, (4) end-of-burst departure, and (5) BATr activation/reset. The departure process distribution consists of two parts: burst inter-departure time ($\tilde{t}$), and burst size ($\tilde{s}$) distributions. The burst inter-departure time takes values which are integer multiples of a slot. It is defined as the interval from the end of a previous burst to the beginning of the following burst. Our goal is to find the joint distribution of $\tilde{t}$ and $\tilde{s}$ i.e., $P_{\tilde{t},\tilde{s}}(t,s), t \geq 0, s \leq y$. To approach it, we first obtain the queue length distribution seen by departing bursts, based on an imbedded Markov chain analysis placing the imbedded points at burst departure instants.

Define random variable $\tilde{q}_{y_k}^k$ to be the number of packets left behind by the k_{th} departing burst, say at time slot t_k, under the condition that the arrival process is in state y_k (=H or L) at t_k. Let random variable $\tilde{u}_{z|y}$ represent the number of packets that arrive during the burst inter-departure interval, under the condition that the arrival process changes from state y prior to the beginning of the interval, to state z at the end of the interval. Moreover, let random variable $\tilde{v}_{z|y}^n$ denote the number of packets that arrive during the transmission time of an n-packet burst, namely n slots, under the condition that the arrival process changes from state y prior to the beginning of the time interval, to state z at the end of the interval. Accordingly, we find that

$$\tilde{q}_{y_{k+1}}^{k+1} = (\tilde{q}_{y_k}^k + \tilde{u}_{z_{k+1}|y_k} - y)^+ + \tilde{v}_{y_{k+1}|z_{k+1}}^{\min\{\tilde{q}_{y_k}^k + \tilde{u}_{z_{k+1}|y_k}, y\}}, \qquad (2)$$

where $y_k, y_{k+1}, z_{k+1} \in \{H, L\}$, and $a^+ = \max\{a, 0\}$. In Equation (2), since BATr is reset or activated after the k_{th} burst departure time, and $\tilde{u}_{z_{k+1}|y_k}$ and $\tilde{v}_{y_{k+1}|z_{k+1}}^{\min\{\tilde{q}_{y_k}^k + \tilde{u}_{z_{k+1}|y_k}, y\}}$ are independent of any events that occur prior to time index k, $\{\tilde{q}_{y_k}^k, y_k \in \{H, L\}, k \geq 1\}$ is hence an imbedded Markov chain. Based on Equation (2), we can derive the limiting distributions of the queue length seen by departing bursts. We first derive the distribution for the number of packets that arrive in any given interval. Let $c_{r_t|r_0}^t(m)$ denote the probability that m packets have arrived in an interval of t slots, under the condition the arrival process changes from state r_0 prior to the beginning of the interval, to state r_t at the end of the interval. For $t = 0$, we immediately have $c_{r_0|r_0}^0(0) = 1$. For $t \geq 1$, $c_{r_t|r_0}^t(m)$ can be recursively computed as

$$c^t_{r_t|r_0}(m)=\sum_{x\in\{H,L\}} P_{x,r_t}\cdot[c^{t-1}_{x|r_0}(m)(1-l_{r_t})+\sum_{n=1}^{m} c^{t-1}_{x|r_0}(m-n)\,l_{r_t}b_{r_t}(n)], \quad (3)$$

where $r_0, r_t \in \{H, L\}$. With the "$()^+$" sign removed, Equation (2) can be expanded into three cases, as

$$\tilde{q}^{k+1}_{y_{k+1}} = \begin{cases} \tilde{q}^k_{y_k} - y + \tilde{v}^y_{y_{k+1}|y_k} & \text{, if } \tilde{q}^k_{y_k} \ge y \\ \tilde{q}^k_{y_k} + \tilde{u}_{z_{k+1}|y_k} - y + \tilde{v}^y_{y_{k+1}|z_{k+1}} & \text{, if } \tilde{q}^k_{y_k} < y,\ \tilde{q}^k_{y_k} + \tilde{u}_{z_{k+1}|y_k} \ge y \\ \tilde{v}^{\tilde{q}^k_{y_k}+\tilde{u}_{z_{k+1}|y_k}}_{y_{k+1}|z_{k+1}} & \text{, if } \tilde{q}^k_{y_k} + \tilde{u}_{z_{k+1}|y_k} < y \end{cases} \quad (4)$$

We now compute the queue length distribution by first conditioning on the value of $\tilde{q}^k_{y_k}$ and separating case one from cases two and three in Equation (4), as

$$P[\tilde{q}^{k+1}_{y_{k+1}} = d] = \sum_{q=y}^{y+d} \sum_{y_k\in\{H,L\}} F_1 \cdot P[\tilde{q}^k_{y_k} = q] + \sum_{q=0}^{y-1} \sum_{y_k\in\{H,L\}} F_2 \cdot P[\tilde{q}^k_{y_k} = q], \quad (5)$$

where

$$F_1 \equiv P[\tilde{q}^k_{y_k} - y + \tilde{v}^y_{y_{k+1}|y_k} = d \,|\, \tilde{q}^k_{y_k} = q] = P[\tilde{v}^y_{y_{k+1}|y_k} = d - q + y] = c^y_{y_{k+1}|y_k}(d-q+y), \quad (6)$$

and

$$\begin{aligned} F_2 &\equiv P[(\tilde{q}^k_{y_k} + \tilde{u}_{z_{k+1}|y_k} - y)^+ + \tilde{v}^{\min\{\tilde{q}^k_{y_k}+\tilde{u}_{z_{k+1}|y_k}, y\}}_{y_{k+1}|z_{k+1}} = d \,|\, \tilde{q}^k_{y_k} = q] \\ &= \sum_{u=0}^{y-q-1} P[\tilde{v}^{q+u}_{y_{k+1}|z_{k+1}} = d] \cdot P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q] \\ &\quad + \sum_{u=y-q}^{d+(y-q)} P[\tilde{v}^y_{y_{k+1}|z_{k+1}} = d - (q+u-y)] \cdot P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q] \quad (7) \\ &= \sum_{u=0}^{y-q-1} c^y_{y_{k+1}|z_{k+1}}(d) \cdot P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q] \\ &\quad + \sum_{u=y-q}^{d+(y-q)} c^y_{y_{k+1}|z_{k+1}}(d-q-u+y) \cdot P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q] \end{aligned}$$

To proceed, we need to solve $P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q]$ in Equation (7). It can be resolved by considering five cases depending on different ranges of *u* and *q* values as given in Equation (8) below:

$$P[\tilde{u}_{z_{k+1}|y_k} = u \,|\, \tilde{q}^k_{y_k} = q] = \begin{cases} 1 & \text{, if } q \ge y,\ u=1, z_{k+1}=y_k \\ \sum_{t=1}^{t} \sum_{m=0}^{y-q-1} \sum_{x\in\{H,L\}} c^{t-1}_{x|y_k}(m) P_{x,z_{k+1}}\, l_{z_{k+1}} b_{z_{k+1}}(u-m) & \text{, if } 0<q<y,\ u \ge y-q \\ c^t_{z_{k+1}|y_k}(u) & \text{, if } 0<q<y,\ u<y-q \\ \sum_{r\in\{H,L\}} \sum_{m=1}^{u} \{P[\tilde{u}_{z_{k+1}|r} = u-m \,|\, \tilde{q}^k_r = m] \cdot [\sum_{r=1}^{\infty} \sum_{r\in\{H,L\}} c^{t-1}_{x|y_k}(0) P_{x,r}\, l_r b_r(m)]\} & \text{, if } q=0 \\ 0 & \text{, otherwise} \end{cases} \quad (8)$$

With Equations (3), and (5)-(8), the limiting queue length distribution can be given by $P[\tilde{q}_y = d] = \lim_{k\to\infty} P[\tilde{q}^k_y = d]$, $y \in \{H, L\}$. The departure process distribution, $P_{\tilde{t},\tilde{s}}(t,s)$ can be considered, depending on different *t* and *s* values, as

Case I: $t = 0$

$$P_{\tilde{t},\tilde{s}}(t,s) = \begin{cases} \sum_{y\in\{H,L\}} P[\tilde{q}_y \ge y] & \text{, if } s = y \\ 0 & \text{, if } s < y \end{cases}. \quad (9)$$

Case II: $0 < t < t$

$$P_{\tilde{t},\tilde{s}}(t,s) = \begin{cases} \sum_{\substack{q+m<y \\ y,i,j\in\{H,L\}}} c^{t-1}_{i|y}(m) P_{i,j}\, l_j \sum_{n\ge y-q-m} b_j(n) P[\tilde{q}_y = q] & \text{, if } s = y \\ 0 & \text{, if } s < y \end{cases} \quad (10)$$

Case III: $t = t$

$$P_{\tilde{t},\tilde{s}}(t,s) = \begin{cases} \sum_{\substack{q+m<y \\ y,i,j\in\{H,L\}}} c^{t-1}_{i|y}(m) P_{i,j}\, l_j \sum_{n\ge y-q-m} b_j(n) P[\tilde{q}_y = q] & \text{, if } s = y \\ \sum_{\substack{0<q<y \\ y,i\in\{H,L\}}} c^t_{i|y}(s-q) P[\tilde{q}_y = q] & \text{, if } s < y \end{cases} \quad (11)$$

Case IV: $t > t$

$$P_{\tilde{t},\tilde{s}}(t,s) = \begin{cases} \sum_{\substack{m<y \\ y,i,j,h\in\{H,L\}}} c^{t-t-1}_{i|y}(0) c^t_{j|i}(m) P_{j,h}\, l_h \sum_{n\ge y-m} b_h(n) \cdot P[\tilde{q}_y = 0] & \text{, if } s = y \\ \sum_{y,i,j,h\in\{H,L\}} c^{t-t-1}_{i|y}(0) P_{i,j}\, l_j \sum_{n\ge y-m} b_j(n) c^t_{h|j}(s-n) \cdot P[\tilde{q}_y = 0] & \text{, if } s = y \end{cases} \quad (12)$$

Combining Equations (9)-(12), we achieve the joint-form departure process distribution.

4.3.2. Numerical Results

Analytic and simulation results of the departure process distribution are plotted in Figure 4. In the MMBP, we adopt $a = 0.225$, $b = 0.025$, $l_H = 0.36$ and $l_L = 0.0933$ at load 0.6; and $l_H = 0.48$ and $l_L = 0.1244$ at load 0.8. The batch size was uniformly distributed between 1 and 9. Accordingly, the burstiness of traffic is $B = 3$ under both loads. All analytical results are in profound agreement with simulation results.

To examine the effectiveness of shaping, we further compute the Coefficient of Variation (CoV) for the inter-departure time and burst size, under three *y* values (y = 1, 10, and 100) and various MMBP

arrivals (B = 1, 3, and 5). Notice that the setting of y = 1 corresponds to a FIFO system with no shaping. Numerical results are plotted in Figure 6., As shown in Figure 6, as expected, the CoV of the inter-departure time increases with the offered load. Crucially, under any MMBP arrival, we discover that the CoV declines significantly with larger y values, yielding substantial reduction in burst loss probability. This fact will be again revealed in the network-wide simulation results presented in the next section.

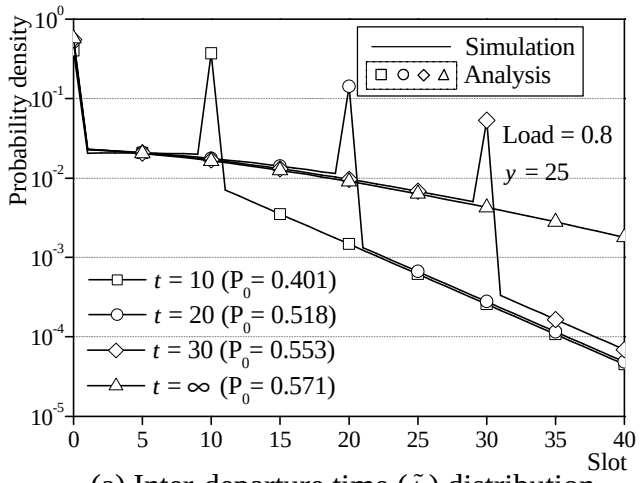
4.4. Loss QoS Provision and Comparison

In this section, we demonstrate the performance of (y,t)-Shaper from two aspects: loss QoS provisioning for OCPS networks, and loss QoS performance comparison between the OCPS and the JET-based OBS [5] networks. We have simulated an entire optical network with QoS burst truncation and full wavelength conversion capabilities equipped in each switching node. The network we used in the experiment is the ARPANET network [14] with 24 nodes and 48 links, in which 14 nodes are randomly selected as edge nodes. OLSP routing is subject to load balance of the network. Each link has up to 100 wavelengths, transmitting at 1 Gb/s, or one 60-byte packet per slot of duration 0.48μs. In simulations, we generate packets according to the MMBP with $a = 0.225$ $b = 0.025$ and the batch size in both H and L states being uniformly distributed between 1 and 9 ($\tilde{b}_H = \tilde{b}_L = 5$).

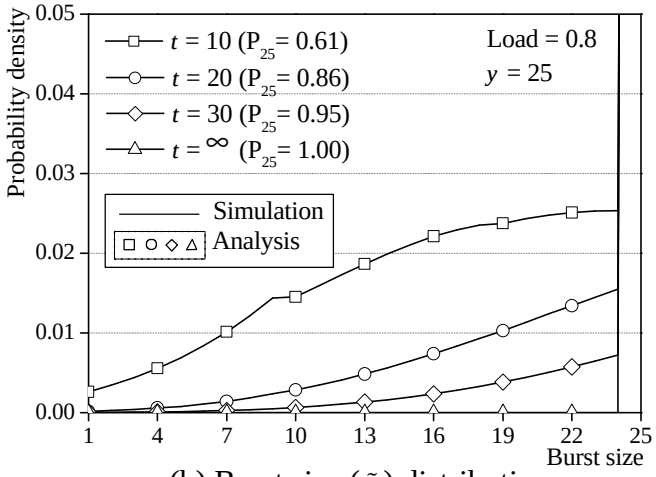
4.4.1. Loss QoS Provisioning

Within OCPS networks, each switching node performs QoS burst truncation in the absence of free wavelengths. Specifically, an arriving high priority burst that finds no free wavelength will preempt a burst that is of *lower* priority (than the arriving burst's priority), and that has the *least* amount of data left unsent. Namely, the preemption is made on a "least-harm" basis.

In simulations, we employ three traffic classes- H, M, and L, in the order of decreasing loss priorities. Each of these three classes generates an equal amount



(a) Inter-departure time ($\tilde{t}$) distribution



(b) Burst size ($\tilde{s}$) distribution

Figure 4. Departure process distributions.

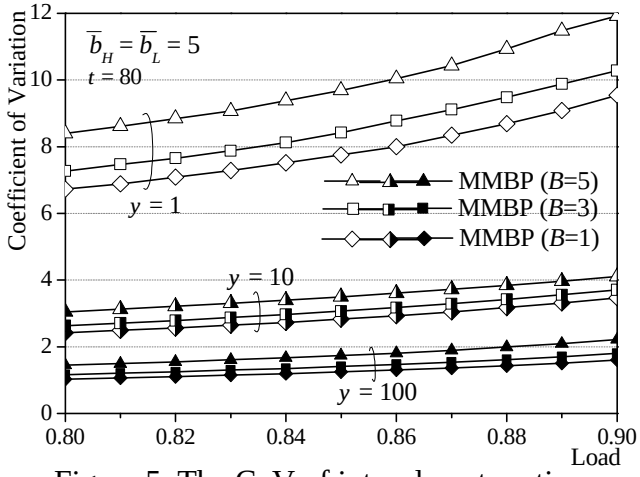


Figure 5. The CoV of inter-departure time.

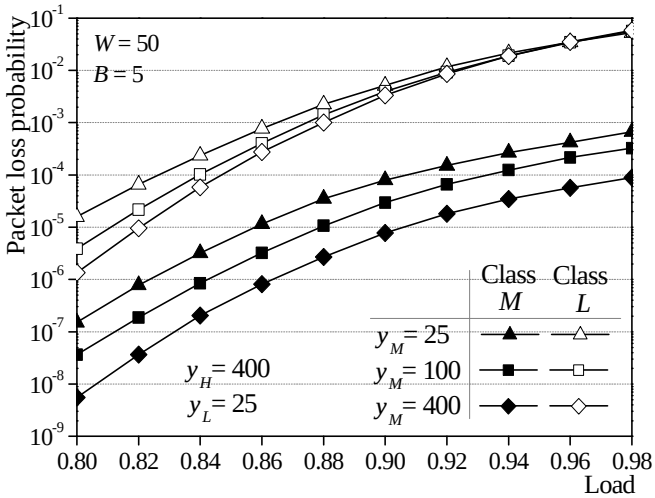


Figure 6. Packet loss under various Class M's burst sizes.

of MMBP traffic into the network. Notice that, the packet loss probability for Class *H* is too low to be measured within affordable time periods. Though, it is sufficient to show the packet loss behavior for both Classes *M* and *L*. Simulation results are shown in Figure 6.

In Figure 6, we plot the packet loss probabilities of Classes *M* and *L* as functions of the load and burst size of Class *M*. We discover a win-win phenomenon from the figure that, by increasing the burst size of Class *M*, the packet loss probabilities for both Classes *M* and *L* decline noticeably. This is because since Class *M* experiences better loss performance due to the use of a larger burst size (better shaping effect), Class *M* makes less preemption toward Class *L* traffic.

4.4.2. OCPS and OBS Performance Comparison

We carried out simulations on the same 24-node ARPANET network in which three traffic classes (Classes *H*, *M*, and *L*) were adopted. In simulations, each ingress node generates a total of 39 connections (3 classes for each of 13 destination nodes) that follow different load-balancing OLSPs. For ease of comparison, we use the same burst size for all three classes during burstification, namely $y_H = y_M = y_L$.

For OCPS networks, we conduct QoS burst truncation in switching nodes on priority plus least-harm-preemption bases. For OBS networks, the offset time assigned to a burst is the total control packet processing time (path-dependent) plus the extra delay $x \cdot T$, where T is the maximum burst transmission time (e.g., 48μs for y=100), and x is (6,3,0), (4,2,0), and (2,1,0) for Classes *H*, *M*, and *L*, respectively. In addition, the header processing time (d) at each switching node is assumed fixed. Finally, we employ restricted QoS burst truncation during contention for the OBS network. Specifically, truncation of bursts is also accomplished on priority plus least-harm- preemption bases, but restricted to those bursts whose control packets have not yet departed from the switch.

In Figure 7, we draw comparisons of packet loss probabilities of all three traffic classes between the OCPS and three variants of OBS networks using three extra-delay settings, respectively, under three cases set by two burst sizes (y= 25, 100) and two header processing times (d= 9.6μs, 48μs). Significantly, we discover that, compared to OCPS as shown in Figures 7(a), OBS undergoes several orders of magnitude deterioration in packet loss performance for Class *H* traffic particularly under a smaller burst size, i.e., $y_H = y_M = y_L$=25. Among the three OBS variants, OBS(2,1,0) using the smallest extra offset time difference (=T) invariably suffers from the poorest packet loss probability. Such performance degradation is caused by the near-far problem that exacerbates under a smaller burst size, a larger header processing time, and/or a smaller extra offset time difference. Under any of the conditions, the offset time of a Class-*H* burst is more likely to be smaller than that of a Class-*M* or Class-*L* burst, resulting in failing to make earlier wavelength reservation for the burst. As the burst size increases and the processing time decreases, as shown in Figures 7(b), and (c), the near-far problem is relaxed, yielding noticeable performance improvement for Class *H* in OBS networks. As opposed to OBS, the in-band-controlled-based OCPS networks are shown to provide invariably superior packet loss probability for Class *H* traffic.

4.5. Conclusions

In this report, we have proposed a dual-purpose, QoS-enhanced traffic control scheme, (y,t)-Scheduler/Shaper, for OCPS mesh-based metro core networks. Providing delay class differentiation, (y,t)-Scheduler assures each weight-based delay class a 99% delay bound guarantee. (y,t)-Shaper provides loss class differentiation by means of assigning larger burst sizes to higher priority classes. Through a precise departure process analysis of an MMBP/G/1 system, we have delineated that (y,t)-Shaper effectively reduces the CoV of the burst inter-departure time, resulting a substantial reduction in packet loss probability. Simulation results demonstrated that, due to the near-far problem, OBS undergoes several orders of magnitude increase in packet loss probability for Class *H* traffic particularly under a smaller burst size. As opposed to OBS, the in-band-controlled-based OCPS network was shown to provide invariably superior packet loss performance for a high priority traffic class, enabling effective facilitation of loss class differentiation.

五、計畫成果自評

A complete version of the report has been accepted (Mar. 2004) and to be appeared in *IEEE J. Select. Areas Commun*.

六、參考文獻

[1] F. Callegati, G. Corazza, and C. Raffaelli, "Exploitation of DWDM for Optical Packet Switching with Quality of Service Guarantees," *IEEE J. Select. Areas Commun.*, vol. 20, no. 1, Jan. 2002, pp. 190-201.

[2] L. Xu, H. Perros, and G. Rouskas, "The Perspective of Optical Packet Switching in IP Dominant Backbone and Metropolitan Networks," *IEEE Comm. Mag.*, vol. 39, no. 3, March 2001, pp. 136-141.

[3] D. Hunter, *et al.*, "SLOB: A switch With Large Optical Buffers for Packet Switching," *Journal of Lightwave Technology*, vol. 16, no. 10, Oct. 1998, pp. 1725-1736.

[4] T. Battestilli, and H. Perros, "An Introduction to Optical Burst Switching," *IEEE Comm. Mag.*, vol. 41, no. 8, Aug. 2003, pp. S10-S15.

[5] M. Yoo, C. Qiao, and S. Dixit, "Optical Burst Switching for Service Differentiation in the Next Generation Optical Internet," *IEEE Comm. Mag.*, vol. 39, no. 2, Feb. 2001, pp. 98-104.

[6] V. Vokkarane, and J. Jue, "Prioritized Burst Segmentation and Composite Burst-Assembly Techniques for QoS Support in Optical Burst-Switched Networks," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, Sep. 2003, pp. 1198-1209.

[7] J. Wei, and R. McFarland, "Just-In-Time Signaling for WDM Optical Burst Switching Networks," *Journal of Lightwave Technology*, vol. 18, no. 12, Dec. 2000, pp. 2019-2037.

[8] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, Oct. 2000, pp. 1838-1851.

[9] M. Yoo, C. Qiao, and S. Dixit, "QoS Performance of Optical Burst Switching in IP-over-WDM Networks," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, Oct. 2000, pp. 2062-2071.

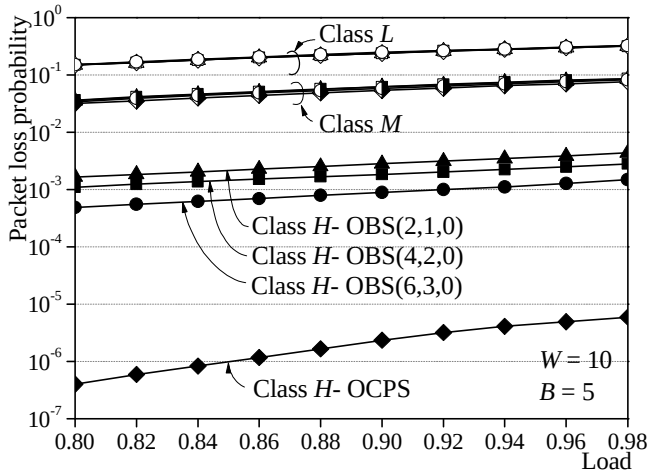
[10] L. Yang, Y. Jiang, and S. Jiang, "A Probabilistic Preemptive Scheme for Providing Service Differentiation in OBS Networks," in *Proc. IEEE GLOBECOM*, 2003.

[11] M. Yuang, J. Shih, and P. Tien, "Traffic Shaping for IP-over-WDM Networks based on Optical Coarse Packet Switching Paradigm," in *Proc. European Conference on Optical Communication (ECOC)*, 2003.

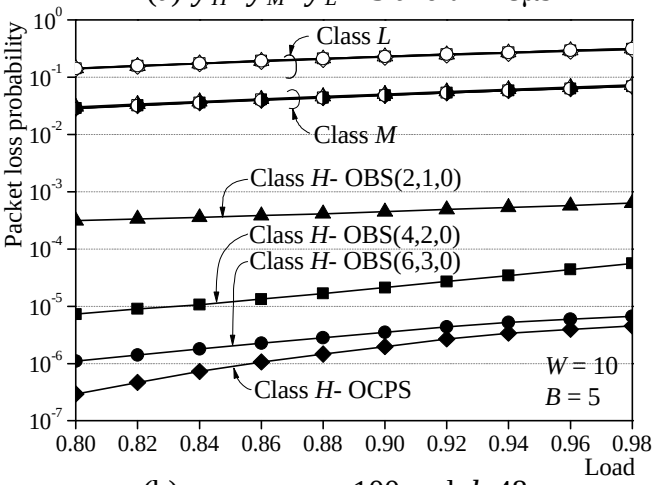
[12] Y. Lin, M. Yuang, S. Lee, and W. Way, "Using Superimposed ASK Label in a 10 Gbps Multi-Hop All-Optical Label Swapping System," to appear in *Journal of Lightwave Technology*, 2004; and also appear in *Proc. IEEE OFC*, 2004.

[13] A. Parekh, and R. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Trans. on Networking*, vol. 1, no. 3, June 1993, pp. 344-357.

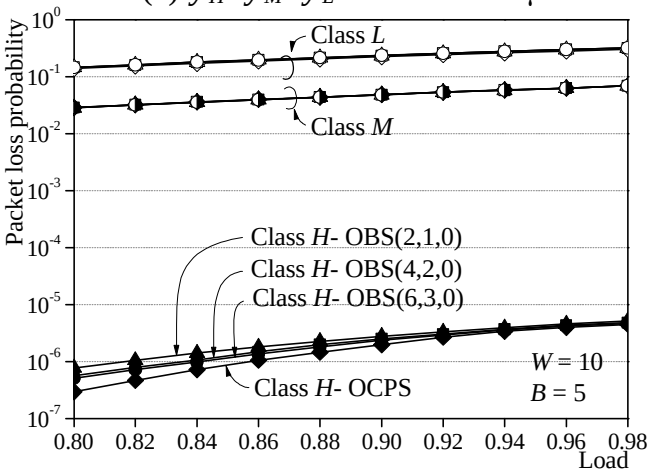
[14] H. Harai, *et al.*, "Performance analysis of wavelength assignment policies in all-optical networks with limited-range wavelength conversion," *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, Sep. 1998, pp. 1051-1060.



(a) $y_H=y_M=y_L=25$ and d =48μs



(b) $y_H=y_M=y_L=100$ and d=48μs



(c) $y_H=y_M=y_L=100$ and d =9.6μs

Figure 7. OCPS and OBS loss performance comparison.