行政院國家科學委員會專題研究計畫成果報告

智慧型寬頻光纖通信網路研究—

子計劃一：高密度波長多工通信網路研究

計畫編號：NSC-89-2215-E-002-012

NSC-89-2215-E-002-061

NSC-90-2215-E-002-012

執行期間:88 年 8 月 1 日至 91 年 7 月 31 日

主　持　人：吳靜雄　教授　　台灣大學電機系

參與人員：翁啟舜、陳坤佐、蔡孟光、古豐瑞、林建宏、

黃柏鈞、楊銘樑、熊致融、張智豪、簡義文、康欣祐、王迺鈞

一、 中文摘要

【第一年】

網際網路使用者日益增加，使得通訊量快速成長；爲了提供使用者更大的服務頻寬，光纖高密度波長多工技術逐漸成爲主要技術，但由於網路營運的需要，如 ATM 及 SDH 仍爲網路之基本架構。故 IP over ATM over SDH/WDM 爲目前網路最常見的架構，主要是 ATM 及 SDH 含有營運所需的 OAM 功能。但 IP over ATM over SDH/WDM 的層次太多，且每一層的功能協定重複，在 host 或 router 處理規約太過繁瑣，因此必須改變其結構。我們提出一個新的 IP over WDM 架構，這個架構使用了標籤交換(label switching)與 WDM 波長資訊的觀念，同時也將 MPLS 流量控制概念與 IP 傳輸與交換平面做有效整合。另一方面我們也分析了兩種可以處理非同步變動長度封包的高速光網路路由器：部份分享暫存器(PSB)與多重路徑路由(MPR)，模擬結果顯示 PSB WDM 光路由器比較適合使用在高速 IP 流量的核心網路(core networks)。

關鍵詞：高密度波長多工，協定，標籤交換，路由器

【第二年】

多協定標籤交換 （Multi Protocol Label Switching, MPLS）與光交換（Optical Switching）的整合已經成爲光通訊網路的趨勢。我們已提出一穩定且具有路由交換及傳輸功能的規約。以此規約爲基礎，我們研究了叢發交換（burst switching）和封包交換（packet switching）所需具備的條件，進而提出了一多平面之光交換器架構。此交換器之架構簡單，而其所能達到的最高流通率（maximum throughput）爲百分之百。同時，其輸入之封包不須在輸入端對齊等待排程，其長度也不需限定爲一定長度，所以可簡化光儲存器的設計。理論分析以及模擬結果顯示，此交換器之效能優異。

現代網路的交換器必須提供多點發送交換（multicast switching）的功能，即允許同一進入的封包同時到達

多個輸出端。此功能可以支援諸如視訊會議、隨選視訊系統和遠距教學等等之網路應用。我們的交換器可以有效地提供多點發送(multicast)。同時，我們研究了一般可變長度封包之多點發送之效能分析，並提出了最佳化的設計原則。

關鍵詞：交換器，多協定標籤交換，光交換，叢發交換，多點發送，可變長度封包，非同步。

【第三年】

光纖通信技術爲固定網路的主流，目前被動光網路被視爲可以真正實現全服務接取網路(FSAN)的方法。現在提出的被動光接取網路以時域多工(TDMA)與非同步傳輸模式(ATM)爲主，然而對於網際網路或乙太網路封包而言，時域多工並不適用；爲了實現真正寬頻接取服務，必需解決寬頻接取網路的問題；本計劃提出一個以光分碼多工應用在被動光網路之實驗網路；研究內容包含同步光分碼多工之編碼研究，非同步光分碼多工之編碼研究，以乙太網路訊框爲基礎結合完美差異碼之被動光網路。我們設計出最好的完美差異碼及非理想交互關係碼。前者可以完全消除其他用戶干擾，是同步光分碼多工系統中效能最好的編碼方式；後者則是在非同步系統中有最多用戶數的編碼方式，這可以有效提高網路容量。我們提出的光分碼多工應用在被動光網路上可以真正達到寬頻接取的服務。

關鍵詞：被動光網路，時域多工，非同步傳輸模式，光分碼多工，完美差異碼，非理想交互關係碼

Abstract

【1st year】

Because the Internet traffic increases rapidly, the Internet Protocol (IP) has become very popular. Large transmission bandwidth is required to provide more services. The Dense Wavelength Division Multiplexing (DWDM) technology is growing in optical fiber transmission and can provide a lot of transmission capacity. Currently many companies have the network architectures as IP over ATM

over SDH/WDM. Because these architectures have too many layers, they need a lot of processing power in the host and the router. Thus the architectures must be simplified to make efficient use of resources. A new IP over WDM protocol with routing and transport capabilities is proposed. This protocol utilizes the concept of label switching and wavelength information of Wavelength Division Multiplexing (WDM). The transport adaptation and encapsulation scheme of this protocol is simple and transparent to protocol data unit. It also integrates the Multiprotocol Label Switching (MPLS) traffic engineering control concept as the IP transport and switching control plane. On the other hand, we have fulfilled the performance analysis for two types of asynchronous Wavelength Division Multiplexing (WDM) optical routers that can handle asynchronous and variable length packet switching for high-speed Internet core networks. We compare the network using Partially Shared Buffer (PSB) and using Multiple Path Routing (MPR) by simulation results. We conclude that a PSB type asynchronous WDM optical router using the same number of input/output ports as the re-circulation ports without using void filling algorithm is more suitable for handling IP traffic in the high-speed optical core networks.

Keywords: dense wavelength division multiplexing (DWDM), multiprotocol label switching (MPLS), router, partially shared buffer (PSB), multiple path routing (MPR)

【2nd year 】

Multiple Protocol Label Switch (MPLS) and Optical Switching integration is a trend for future all optical networks. We had proposed a very stable protocol with transport and switching capabilities. Based on this protocol, we study the feasibility of using this protocol on burst switching and packet optical wavelength switching architecture. The architecture may fulfill

the requirements of both burst switching and packet switching networks. The architecture can switch variable length packets asynchronously, i.e., the input packets are not necessary to enter the switch at the inputs ports simultaneously. This property will simplify the buffer design. The result shows that the throughput of this switch is very high.

A modern switch must provide multicasting to allow multiple copies of an incoming packet to be sent to their destinations simultaneously. Such feature is necessary to support video conferencing, video-on-demand, distant learning, etc. An enhanced version of the variable length self-routing packet switch is proposed to perform multicast switching. The optimal design rule and analysis of multicast switching of variable length packets are discussed in this report.

Keywords: switch, MPLS, burst switching, packet switching, optical switching, variable length packet, asynchronously

【3rd year】

Optical fiber is an important medium of the wireline communication systems. Passive optical network (PON) is considered to be a solution to realize full-service access networks (FSAN). It is a point-to-multipoint optical network with no active devices through the signal transmission. The network is easy to maintain and to be upgraded. Combining Ethernet frame with PON (EPON) is a promising candidate for next-generation broadband access networks. PON uses time division multiple access (TDMA) and asynchronous transfer mode (ATM) in upstream and down stream direction. The two schemes have fixed time slot and are not suitable to variable length packets, such as IP packets and Ethernet frames. We proposed an EPON configuration using optical code division multiple access (OCDMA) in its upstream direction. The downstream direction is still broadcasting in nature,

which is the same as the ordinary EPON. We proposed a synchronous OCDMA code: Perfect difference code (PDC). PDC has better performance than the other codes when the multiple user interference (MUI) canceller is used. We also proposed an asynchronous OCDMA code: optical orthogonal codes (OOCs) with nonideal cross correlation. Given the same code weight and code length, the size of the proposed new codes is ten times more than that of ideal optical orthogonal codes. It can thus accommodate more simultaneous users and increases the network utilization. The proposed system can achieve the goal of broadband access networks.

Keywords: Passive optical network (PON), Perfect difference code (PDC), multiple user interference (MUI), optical orthogonal codes (OOCs), cross correlation

二、 緣由與目的

【第一年】

The explosive growth of Internet traffic demands more bandwidth on the core networks. While IP backbone systems are being realized using Synchronous Optical Network / Synchronous Digital Hierarchy (SONET/SDH) and Asynchronous Transfer Mode (ATM) technologies, existing transport networks will not yield the bandwidth-effective transporting and routing schemes for IP traffic. The integration of IP and Dense WDM (DWDM) has been the most popular research and development area for next generation Internet. These new emerging technologies, i.e., Generalized Multiprotocol Label Switching (GMPLS) and the related signaling protocols play an important role of the control plane that could integrate IP and DWDM capability. They aim at providing an efficient and flexible end-to-end

lightpath for IP centric and non-IP centric services. MPLS is rapidly emerging as an Internet Engineering Task Force (IETF) standard. MPLS intends to enhance the speed, scalability, and service provisioning capabilities in the Internet. MPLS uses the technique of packet forwarding based on labels, to enable the implementation of a simpler high-performance packet-forwarding engine. MPLS is best viewed as a 2.5th layer protocol, integrating layer 2 and layer 3 functions. With this point of view, optical label switching has been deemed as the most probable technique for the future all optical networks.

However, at the present time, due to the gap to success of optical elements, optical labeling is not so realizable. Other techniques using optical sub-carriers to carry the labels may suffer the optical transparent problem in all optical networks. For this sake, we need to reconsider 1) the properties of optical transmission, 2) the properties of IP traffic, and 3) future requirements for IP backbone traffic. What we need are as follows: 1) an IP over WDM transport mechanism that can handle physical framing, data link delineation and part of layer 3 routing functions at the same time. This mechanism also needs to speed up the processing, while keeping all possible optical networks transparent 2) As for IP traffic properties, we need additional traffic engineering models as the control plane for IP networks. In other words, we need a high-speed asynchronous transport mechanism integrating with label switching techniques and the advantages of control plane functions. This mechanism must be able to migrate to future all optical networks smoothly.

Although GMPLS could support packet, Time Division Multiplexing (TDM), wavelength, waveband and fiber switching, it still treats DWDM as a big bandwidth pipeline and acts more like cross-connect provisioning mechanism. This may not be the most efficient and flexible solution that could be achieved

directly by combining the advantage of DWDM with all optical packet switching capabilities to yield WDM optical router for IP centric services. However, at the present time, due to the availability of optical elements, all-optical packet switching and processing are not so realizable. Almost all-optical packet switching using optical-to-electrical-to-optical (O-E-O) techniques for the header processing is more practical. At the same time, the optical buffer can only be emulated by using fiber delay lines with fixed delay unit. On the other hand, considering the present Internet network scenario and IP properties, the transfer mode based on variable length packets and asynchronous operation is more effective.

【第二年】

Recently, the traffic of backbone and access networks increase dramatically. The transmission capacity of optical networks doubles every nine months. However, the switching and routing capacity increases slowly compared with the transmission capacity. Many switching or access schemes are proposed [1]-[4]. Among them, integration of MPLS and wavelength division multiplexing technologies is a promising choice.

Combination of electronic and optimal routing for the hybrid Internet Protocol (IP) networks is inevitable. A WDM optical packet switch consists of input interface unit, the switch fabric, the output interface unit and the control unit. In a conventional switch, the input interface unit processes the packet header information extraction, packet delineation and alignment. The switching fabric is the most important element. The output interface unit regenerates the optical signals. Based on the information in the packet header, the control unit controls the switch. Because of the synchronization requirements, the packet size must be fixed. Currently, the

packet header can be out-of-band signaling including sub-carrier multiplexing and dual wavelength, or in-band-signaling.

In this report, we propose a self-routing variable-length packet switch, which can provide unicast and multicast switching. The performance of the switch operating in unicast or multicast switching is verified both by simulation and queueing analysis. The first part of this report focuses on the unicast switching, and the second part discusses the multicast switching.

【第三年】

Recently, fiber-optic code-division multiple-access (CDMA) systems have attracted much attention. Because synchronous optical CDMA (SOCDMA) systems are more efficient to utilize the bandwidth than any asynchronous optical CDMA systems, many studies of SOCDMA systems have been reported in the literature. There are several possible ways to achieve slot synchronization. One way is that we add a synchronizer node located at the star coupler. This node broadcasts a pulse called the sync pulse periodically. Using the periodic pulses, the synchronization may be achieved among all the transmitters. One of the major concerns of designing a direct-detection SOCDMA system is the multiple-user interference (MUI) because the performance of this system is usually interference limited. That is, there is an asymptotic error floor no matter how much power the receiver received. Unlike in the electric domain, the signature sequence in a direct-detection SOCDMA consists of unipolar (0, 1), sequences. As a result, there are no strict orthogonal codes under the constraint that the total number of codes is equal to the code length. We study the "perfect difference sets" and propose a new SOCDMA system using these sets as the signature code sequences instead of modified prime sequences. The perfect

difference codes have the following interesting characteristics:

1) Any two distinct codes are cyclic-shifted with each other.

2) The cross correlation between any two distinct codes is exactly one.

However, the modified prime sequence codes have the cross correlation either one or zero. Although these codes are not strictly orthogonal, we may say that they are quasi-orthogonal. Using the first property, we propose a simple transmitter structure, in which the code sequence can be cyclic-shifted to get another transmitter. Using the second property, we also can easily design the receiver to remove the effect of the MUI. However, from the viewpoint of asynchronous systems, the perfect difference codes are basically identical because they are cyclically shifted with one another. In other words, we have only one code for asynchronous systems. Therefore, we must modify the perfect difference codes to get a larger code size. To do so, we consider the case that the correlation is relaxed to two. When the two codes are not aligned—that is, they are cyclically shifted with each other—the cross correlation between the two codes is exactly one. In such a situation, it does not violate the constraint of correlation. However, when there is no cyclic shifting between them, the value of the cross correlation is equal to the code weight, which is far from the constraint. To overcome this, we can drop some marks appropriately from marks of each code, such that the modified code weight is reduced and the code length is still equal to original size. Moreover, the cross-correlation property should satisfy the constraint.

We further discuss the implementation of OCDMA system in high-speed access networks. EPON is promising a solution of broadband full service access networks and is illustrated in Fig. 1.

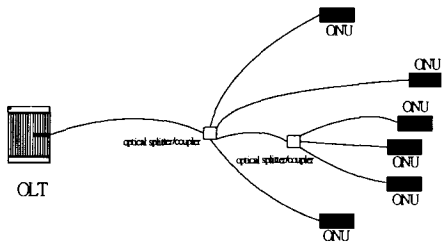


Fig. 1. An EPON configuration

The downstream direction is a point-to-multipoint transmission and is broadcasting in nature. Downstream packets are sent from the optical line terminal (OLT) and received by optical network units (ONU) according to the specific address. In the upstream direction, each ONU transmits packets to the OLT and may collide with other packets if more than one packet arrive at the OLT simultaneously. To solve this problem, three multiple access schemes can be used: time division multiple access (TDMA), wavelength division multiple access (WDMA), and code division multiple access (CDMA). In TDMA, each ONU is allocated a time slot telling when to send data and for how long. This allocation can be static or dynamic. In WDMA, each ONU transmits data at different wavelengths. The WDM components such as tunable laser and tunable optical receiver are much more expensive than that of TDMA and CDMA. The operation and maintainence are also more complicated. Thus makes WDMA not suitable for FSAN.

We propose an EPON configuration using synchronous OCDMA (SOCDMA) in its upstream direction. The downstream direction is still broadcasting in nature, which is the same as the ordinary EPON. Bidirectional transmission is performed over a single fiber by wavelength multiplexing of 1.5 μm wavelength in the downstream direction and 1.3 μm wavelength in the upstream direction. In the downstream direction, the packets are broadcasted to all ONUs from the OLT. Only the frames with the desired ONU's Ethernet address are extracted by that ONU and the others are discarded. In the upstream direction, Ethernet frames from the ONU are spread by a spreading circuit, which uses PDC as the code signature. The OLT decodes the incoming OCDMA signals by using the despreading receivers. Upstream packets

of different addresses can be received simultaneously without collision. The EPON configuration with maximum range of at least 20km can split at most 32 subscribers. The synchronization strategy is worked out as the following. In the proposed network, the OLT has a very high precision clock. When there is an Ethernet packet to be sent, the OLT just transmits the packet. Otherwise the OLT transmits the preamble signal. Thus the OLT plays the role of master clock and transmits clock information to the downstream ONU. Initially, when an ONU is powered, the local clock of the ONU locks the down stream clock signal transmitted from the OLT. In order to synchronize all the ONUs, the OLT polls the inactive ONUs sequentially and periodically. The polled ONU responses immediately and the OLT will estimate the round trip delay and send synchronization information back to the ONU. The ONU synchronizes the local clock and sends training packet to the OLT. The procedure will be repeated until satisfactory synchronization is obtained.

三、 研究結果與討論

【第一年】

A. An IP over WDM Protocol with Routing and Transport Capabilities

The concept of the proposed IP over WDM protocol is not merely an IP over DWDM adaptation and encapsulation scheme, but also the concept of utilizing MPLS Label Distribution Protocol (LDP/CR-LDP) as the DWDM network signaling protocol. It also integrates MPLS traffic engineering control for optical network elements. However, each transport protocol does have its own framing and encapsulation method. In this section, we will first introduce the transmission format of the newly proposed IP over WDM protocol.

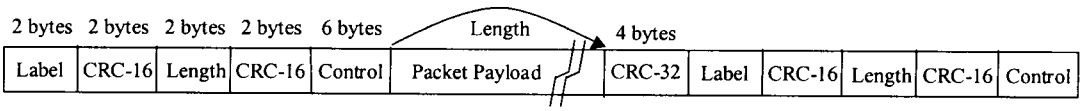


Fig. 2 Physical and Data Link Layer Frame Delineation of the Proposed IP over WDM Protocol

The transmission format of the proposed protocol is shown in Fig. 2. The frame encapsulation is done by using <Label> <CRC-16> <Length> <CRC-16> <Control> and followed by the protocol data unit of variable length and a <CRC-32> field as the trailer for payload CRC checking. As shown in Fig. 2, the "Label" field has two bytes and is used as the label switching function. The CRC-16 circuit with generating polynomial: $g(x) = x^{16} + x^{12} + x^5 + 1$, is used to check if the "Label" field is correct or not during searching process of re-framing. These bytes with "Label" field together form the framing indication of physical layer. The "Length" field together with the second CRC-16 bytes is used to indicate the length of the protocol data unit, and to be the data link layer frame delineation. Once the framer arrives at the "Sync" state, the CRC-16 check circuit is triggered as the single-bit error correction circuit (so does at the "Post Sync" state). So, the framer just need one CRC-16 and one logic control circuit which records the status of the framer and controls the operation of CRC-16 circuit to be multiple errors detection or single-bit error correction. This simplifies the hardware of the framer a lot. "Control" field is used to indicate the type of protocol data unit, scrambler synchronization status and maintenance information. In order to avoid the inefficiency of using preamble techniques as the bit synchronization method, there must have the inter-packet fills to maintain the channel continuity. Those idle inter-packet fills contain only the necessary <Label>, <CRC-16>, <Length> and <CRC-16> fields. Maintenance packets like those to

deliver OAM messages will have additional <Control> and <CRC-32> fields. These OAM packets will periodically be inserted into optical channel with higher priority than the normal data packets.

We propose 16 bits for labels in the proposed IP over WDM protocols. To accommodate all the addressing range, i.e., 2^{20}, and incorporate the technologies of wavelength routing, we adopt 16 different wavelength, $\lambda_0, \lambda_1, \cdots, \lambda_{15}$, as a group of MLPS path addresses. In that way, we can fully utilize the wavelength information while keeping electrical labels as small as possible. In this case, it is possible that we could integrate the MPLS technologies with wavelength routing. These similarities between wavelength and label have been used for combining MPLS traffic engineering control with optical cross-connects (OXCs). Additional need for a General Switch Management Protocol (GSMP) or GSMP-like protocol to interface the optical router gear may also be considered.

I. Analysis of Framing and Encapsulation Performance

To find the frame in "Sync" state at the first time that the system is powered up or at the time of re-framing after loss of frame detected, the framer is in "Hunt" state. The framer uses byte-by-byte search to find the first valid CRC-16 sequence of label field. The framing and delineation state transition diagram of the proposed IP over WDM protocol is shown in Fig. 3. The framing state transition diagram consists of four distinct states: "Hunt", "Presync", "Sync" and "Post Sync" states.

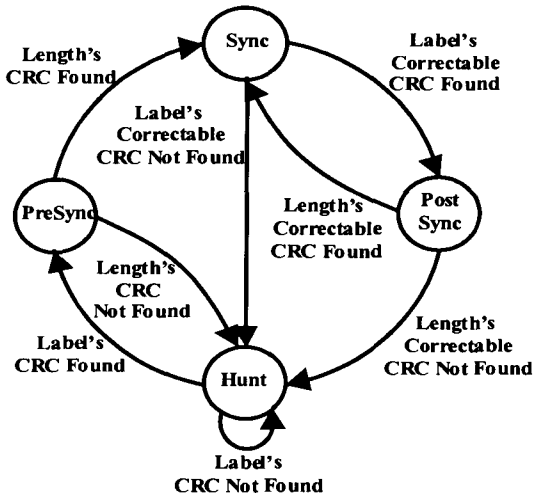


Fig. 3. Framing and Delineation State Transition Diagram

The procedure of framing and delineation is as follows: When the framer is first powered up or in the state of re-framing, the framer tries to find out the first valid Label field in a byte-by-byte shifting and checking manner. If no valid header is found, the framer continues the shifting and check processes, i.e., staying in the "Hunt" state. Once upon finding such a valid CRC-16 sequence, the framer moves to the "Presync" state, and records the status of framing. The control logic of framer, at this state, continues to perform the second CRC-16 checking for the subsequent specific 4 bytes (Length field + second CRC-16 field) to find out the second correct CRC-16 pattern of Length field. If the framer fails in finding the second valid CRC-16 pattern after subsequent 4 bytes, it then returns back to the "Hunt" state. If it finds the second valid CRC-16 sequence, the framer goes into "Sync" state. At that point, the framer is considered in frame. The framer determines from the Length field to find out how many bytes there are until the beginning of next Label field of next frame.

Once the framer gets into the "Sync" state, the control logic circuit of framer enables the CRC-16 error checking circuit into a single-bit error correction circuit. The framer now detects the CRC-16 sequence with no error bit and with one error bit pattern for the Label and Length field of next frame. If it does not find the next single-error correctable CRC sequence, it goes back to the "Hunt" state to perform the searching process again. If the framer detects the first correctable CRC sequence of Label field of next frame, it goes into "Post Sync" state (with CRC-16 as a single-bit error correction circuit). The framer then continues identifying the subsequent 4 bytes with CRC-16 as single-bit error correction circuit to see if they are correctable for the Length field. If not correctable, the framer goes back into "Hunt" state to perform shifting and checking process. If the framer finds out

the correctable CRC-16 sequence of the Length field, it goes into "Sync" state again and maintains the CRC-16 circuit in single-bit error correction status. This process continues until lost of frame detected and all the re-framing processes go over again. We analyze the performance of the proposed protocol as follows.

II. Probability of Loss of Frame

It is worth to note that at the receiver side, the framer is first delineated using <Label> and <CRC-16> to arrive at "Presync" state and then using <Length> and second <CRC-16> fields to get into "Sync" state. Probability of loss of frame (PLF) is directly a function of optical channel bit error rate (BER). Once the synchronization is achieved, single-bit error correction is activated. Thus, for a loss of frame, more than one error in the <Label> and <CRC-16> fields or less than one error in <Label> and <CRC-16> fields but more than one error in <Length> and second <CRC-16> fields is required. Since all these fields are 2 bytes and the single-bit error correction is enabled, the probability of the frame loss due to the channel corruption in framing and delineation is:

$$PLF = PLF_{Label} + PLF_{Length}$$

where, PLF_{Label} is the probability that loss of frame happened in checking valid CRC-16 sequence of Label field and PLF_{Length} is the probability that loss of frame happened in checking valid CRC-16 sequence of Length field. These two probabilities are shown as follows:

$$PLF_{Label} = 1 - [(1 - P_e)^{32} + 32P_e(1 - P_e)^{31}]$$

$$PLF_{Length} = PLF_{Label}(1 - PLF_{Label})$$

where, P_e is the optical channel BER. Fig. 4 shows the probability of loss of frame of the proposed IP over WDM protocol.

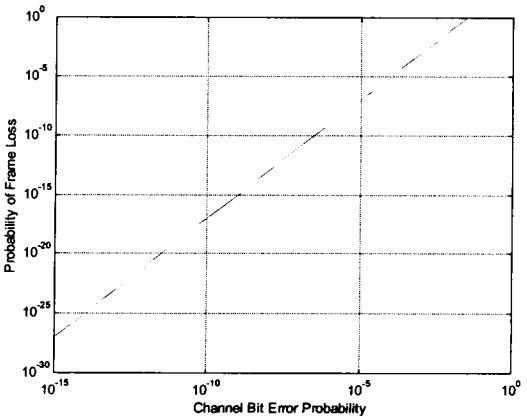


Fig 4. Probability of Frame Loss versus Channel BER.

After we obtain the PLF, we can roughly estimate the "Mean Time To Frame Loss" (MTTFL) which can be expressed as follows:

$$MTTFL = \frac{1}{PLF * FrameRate}$$

where, *FrameRate* is equal to *Channel Bit Rate* / *Frame Length*. However, since the transport is asynchronous and with variable packet length, we can't get the accurate "Mean Time To Frame Loss", unless we know the packet length distribution. We only can roughly estimate this parameter with a given packet length. Fig. 5 shows the graph of *MTTFL* (in Years) versus packet length (in Bytes) at optical channel bit rate 10Gbps and BER 10^{-12}. We can see *MTTFL* even at very small packet length distribution can still get almost a billion years!

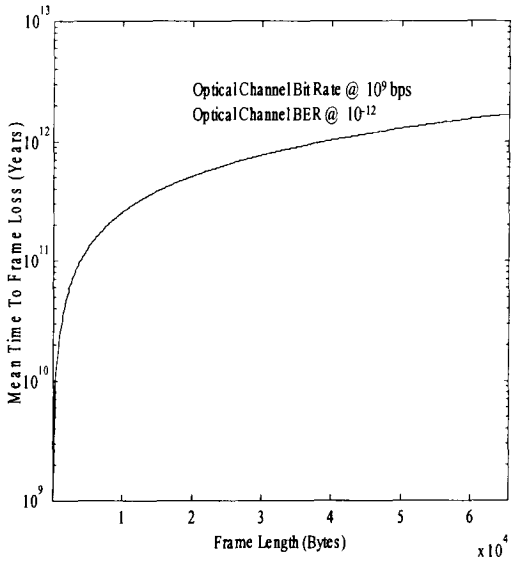


Fig. 5. Mean Time To False Frame versus Packet Length

III. Probability of False Frame

When parsing the frame byte-by-byte in search of a valid header, it is possible that the valid CRC-16 pattern is found inside the protocol data unit. This is the case that the CRC-16 header is simulated by the payload data. And the probability of false frame (*PFF*) is equal to the probability of two consecutive CRC-16 matches and is shown as follows:

$$PFF = \left(\frac{1}{2^{16}}\right)^2 = \left(2^{-32}\right) = 2.33 \times 10^{-10}$$

Even if "false framing" happens, the

framer does not stay in "Sync" state for the coming frames unless it finds more consecutive two CRC-16 matches at the right places. It could be an exact match or a match with only one bit in error. This means that the probability of falsely staying in "Sync" state is *Prob* (false sync) = $(33x2^{-16})^{2} = 25x10^{-8}$. This makes the probability of rejecting false frame becomes 1 - $25x10^{-8}$ and is equal to 0.99999975. From this, we can see that the probability of rejecting staying in false "Sync" state of the proposed protocol is nearly equal to 1.

IV. Mean Time To Frame

Re-frame refers to the transition from "Hunt" state to "Sync" state as shown in Fig. 3. Re-frame time, or mean time to frame (MTTF), is the duration of time required for a frame alignment circuit to re-establish frame alignment from being powered up or from a misalignment condition. A re-frame procedure consists of two separate processes: search and confirmation. A search process sequentially searches the CRC-16 sequence for Label header to identify the beginning of each packet. If there is a match, the position is provisionally accepted as a candidate and a confirmation process is called in to test validity of the candidate position. If the confirmation criterion is met with the candidate position, frame alignment is declared. If not, a search process is reinitiated, followed by another confirmation process and so on.

1. MTTF Search Process

To consider the worst case, a search process is initiated just one byte after the correct frame alignment position. Therefore, all the n bytes in each packet, where n is variable for each packet, have to be checked. For the first (n -1) byte position, a search process takes time to verify them wrong. If the payload simulates the CRC-16 sequence, a confirmation process is called in to detect its falsehood. After the falsehood

is detected, the search process is re-initiated and the next byte position is checked. At the $n\,th$ byte position the search process accepts it. Using the technique of probability of generating function (PGF) in the generalized state transition diagram, we can derive the state transition diagram of search process as shown in Figure 6.

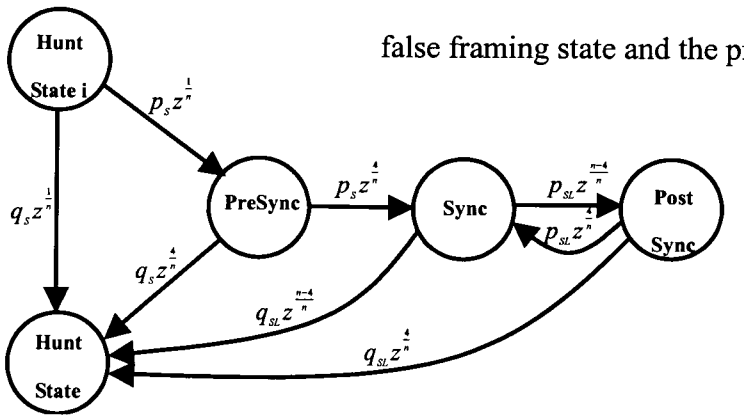


Fig. 6 Falsehood Detection Process in Search Process

The search process moves from the byte position i to $i+1$ in one byte duration. Since the proposed protocol is a packet based framing and delineation architecture, the more meaningful units in qualifying the MTTF duration is in packets. So, we denote one byte duration as $1/n$ packet, where n is the packet length in bytes. If the CRC-16 pattern of Label header is simulated, the confirmation process takes $4/n$ packet to detect its falsehood. The probability of CRC-16 pattern is simulated by the payload is p_s which is equal to $1/2^{16}$. The probability of not being simulated as the CRC-16 pattern is q_s which is equal to $1-p_s$. If the simulation continues happening, the framer will go into the false framing state and the probability of staying in "Sync" state and "Post Sync" state is p_{SL} which is equal to $33\times p_s$ and the probability of back to "Hunt" state is q_{SL} which is equal to $1-p_{SL}$. From Fig. 6, we then can show that the PGF of the state transition from state i to $i+1$ is given by:

$$q_S Z^{\frac{1}{n}} + P_S Z^{\frac{1}{n}}[q_S Z^{\frac{4}{n}} + p_S Z^{\frac{4}{n}} q_{SL} Z^{\frac{n-4}{n}} + p_S Z^{\frac{4}{n}} \frac{p_{SL} Z^{\frac{n-4}{n}}}{1-p_{SL} Z} q_{SL} Z^{\frac{4}{n}}]$$

$$\cong q_S Z^{\frac{1}{n}} + P_S q_S Z^{\frac{5}{n}}$$

The maximum mean time to frame requires the falsehood detection for the first $(n-1)$ byte position and a correct CRC-16 pattern for the last $n\,th$ byte position. Therefore if the correct CRC-16 pattern is not corrupted by optical channel bit errors, the search process takes the signal transfer function $\tau(Z)$ to the right position, where $\tau(Z)$ is given by:

$$\tau(Z) = [q_S Z^{\frac{1}{n}}(1 + P_S Z^{\frac{4}{n}})]^{n-1} Z^{\frac{1}{n}}$$

$$= q_S{}^{n-1} Z(1 + P_S Z^{\frac{4}{n}})^{n-1}$$

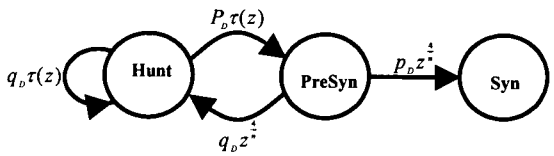


Fig. 7. Reframe Declaration Procedure

2. MTTF Confirmation Process

After the correct position of CRC-16 pattern is found, the confirmation process is called in. Fig. 7 shows the state transition diagram of the reframe declaration procedure. After the correct position is found, the probability that the framer goes into "Presync" state is P_D. And P_D is equal to the $(1-P_e)^{32}$, where P_e is the optical channel BER. The probability that the CRC-16 pattern is corrupted by the channel noise is q_D, which is equal to $1-P_D$.

The overall transfer function from the "Hunt" state to "Sync" state is given by

$$P_{RF}(Z) = \frac{p_D^2 \tau(Z) Z^{\frac{4}{n}}}{1 - q_D \tau(Z)[1 + p_D Z^{\frac{4}{n}}]}$$

which is the PGF of maximum average MTTF. Therefore, the maximum average MTTF, T_{RF} is given by

$$T_{RF} = P'_{RF}(1)$$

$$= \frac{1}{p_D^2}[\tau'(1) + \frac{4}{n} p_D]$$

Fig. 8 (a) to (d) show the maximum MTTF with packet length 65535 bytes, 1500 bytes, 354 bytes and 576 bytes, respectively. The packet length of 1500 bytes is the maximum transfer unit (MTU) of Ethernet and the length of 576 bytes is the length that all routers must support.

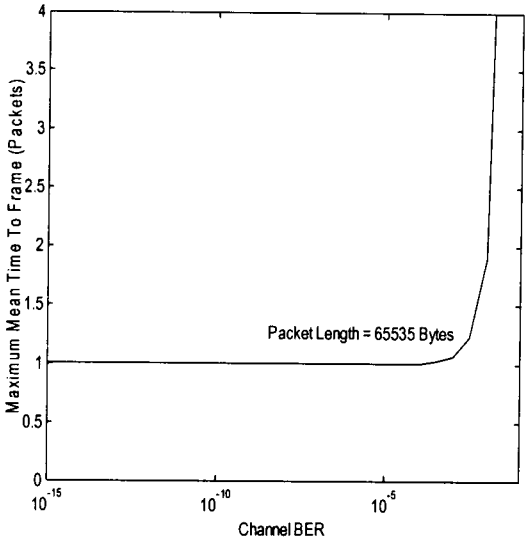


Fig. 8(a) Maximum Mean Time To Frame versus Channel BER with Packet Length 65535 Bytes

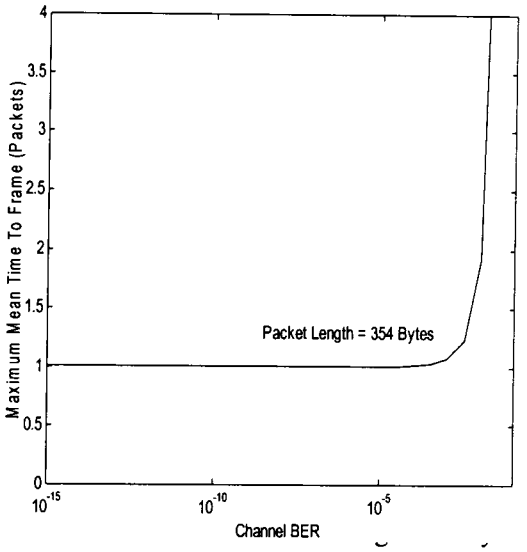


Fig. 8(c) Maximum Mean Time To Frame versus Channel BER with Packet Length 354 Bytes

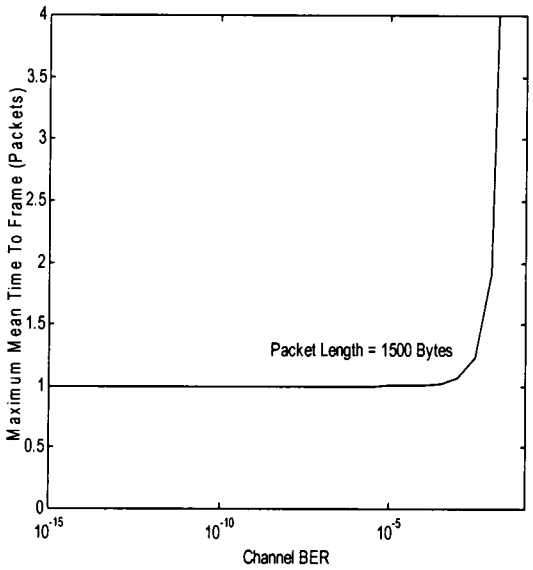


Fig. 8(b) Maximum Mean Time To Frame versus Channel BER with Packet Length 1500 Bytes

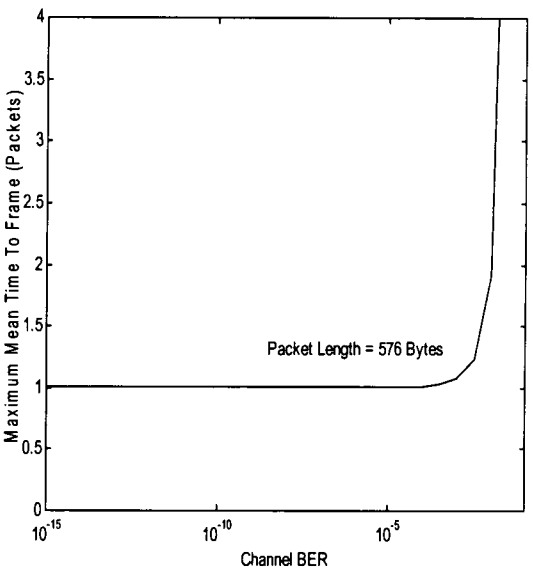


Fig. 8(d) Maximum Mean Time To Frame versus Channel BER with Packet Length 576 Bytes

From Fig. 8 (a) to (d), we can see that when optical channel BER is less than 10^{-5}, the maximum average MTTF is almost equal to one frame. This is quite straightforward, since the proposed protocol is actually in "Sync" state in less than one packet duration. The maximum average MTTF will be at most one packet to get into "Sync" state. This is much faster for the framing of variable packet length than in conventional one. Each different length packet arriving at "Sync" state is almost the same at channel BER less than 10^{-5}. We can see that although the packet length is a random variable and will affect the accuracy of actual MTTF results, we still can conclude that the MTTF will be almost less than one packet.

B. *Performance Analysis of Asynchronous WDM Optical Router under Long Range Dependence Traffic*

We first analyze the performance of the feed forward type asynchronous WDM optical routers under LRD traffic. In order to validate the impact of LRD traffic on the contention resolution techniques used in optical packet switching and the benefit of void filling algorithm, we have done programming in an ad-hoc discrete event simulator (DES) using raw C++ code to perform the analysis. In all the following simulations, we use an aggregated ON-OFF process modulated with a Pareto distribution to simulate LRD traffic. The Pareto distribution with cumulative distribution function (CDF), $P[X \le x] = 1 - (b/x)^{\alpha}$, for $\alpha > 0$, has the degree of self-similarity, $H = (3-\alpha)/2$, where $H \in$ (0.5, 1) and $\alpha \in$ (1, 2). All the traffic parameters for our simulations are Hurst parameter, $H = 0.9$, load of each source, $L_i = 0.8$, and the location parameter of the ON time Pareto distribution, $b = 400$ bytes. The switching parameters are 16x16

input/output ports, 16 wavelengths/port. The buffer depth is limited to 10 fiber delay lines for both of the main output buffers and common shared buffers. Wavelength conversion is activated in all input and re-circulation ports. The number of re-circulation allowed in the router is 30, which is a reasonable limit. The system void records for each output and re-circulation port are 40.

I. Feed Forward Type Asynchronous WDM Optical Router

We first validate the effect of excess load caused by asynchronous and variable length packets on optical buffers with fixed delay time, and the benefit of using void filling algorithm as shown in Fig. 9. We can see that switching performance is hardly improved by using wavelength conversion solely, which can improve the performance significantly for the synchronous and asynchronous optical packet switch under Bernoulli and bursty traffic. We also see the effect of excess load caused by voids in the fiber delay line buffers. The longer delay line length for each fiber delay line is, the more packet contention occurs and more packets are dropped.

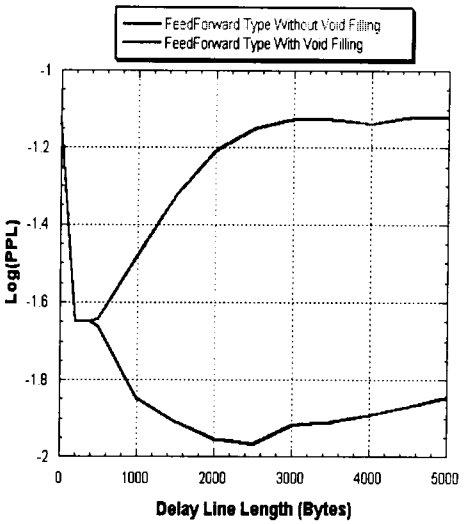


Fig. 9 Performance of feed forward type WDM optical router under LRD traffic

Fig. 10 is the performance analysis for the same router using multiple path routing (MPR). The three-category means that input packets are divided into three different types of traffic: type 1, 80% of the total traffic, not allowing to adopt MPR, type 2, 10% of the total traffic, allowing to have one chance to choose the second output port while the

original output buffers are full and type 3, 10% of the total traffic, allowing to have two chances to choose the alternate output port. On the other hand, the percentages of six-category traffic are $P1 = 0.2\%$ for type 1, $P2 = 30\%$ for type 2, $P3 = 23\%$ for type 3, $P4 = 20\%$ for type 4, $P5 = 25\%$ for type 5 and $P6 = 1.8\%$ for type 6. From Fig. 10, we believe that MPR is a good technique for asynchronous WDM optical router.

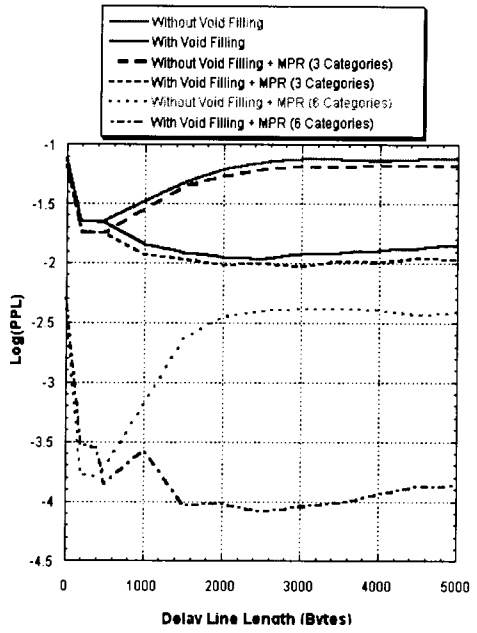


Fig. 10 Performance of feed forward type WDM optical router using MPR under LRD traffic

PSB Type Asynchronous WDM Optical Router

From our architecture of asynchronous WDM optical routers shown in Fig. 11, we force the delay line length of the re-circulation ports equal to the delay line length of the main output buffers in order to share the common fiber delay lines. In the following simulation we set the initial unit delay the same as the delay line length of fiber delay lines of the main output buffers. We simulate two extreme cases for the PSB type WDM optical routers. One is with single re-circulation port and the other is to use 16 re-circulation ports.

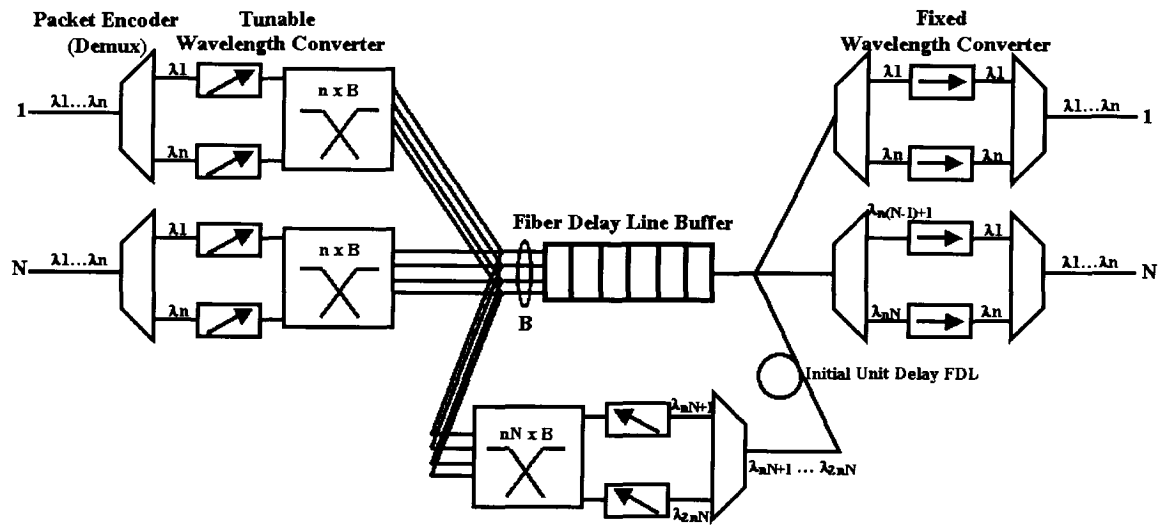


Fig. 11 Architecture of Asynchronous WDM router

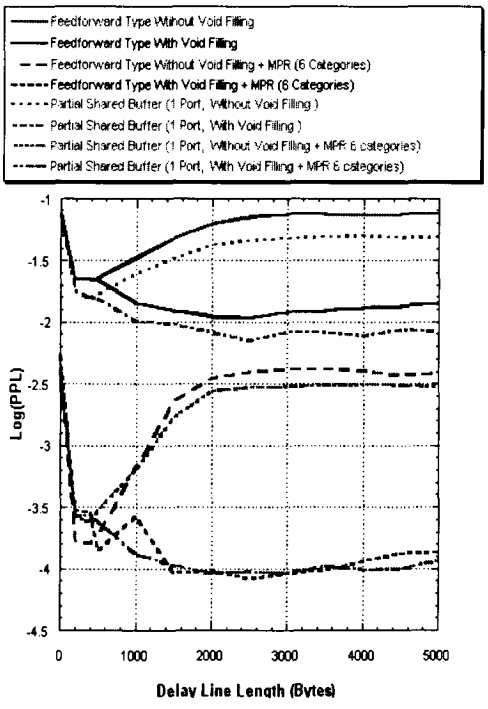


Fig. 12 Performance of feed forward type optical router versus PSB type optical router using 1 re-circulation port under LRD traffic

Fig. 12 shows the performance of feed forward type WDM optical routers with and without activating void filling algorithm versus PSB type WDM optical routers using single re-circulation port operating at the same modes. We can see that the performance of PSB type routers using single re-circulation port looks like the performance of feed forward type routers in various operating modes. Although, the performance of PSB type routers using single re-circulation port is superior to the feed forward type routers, the improvement is only a little. The function of the re-circulated common shared buffers acts as buffers with larger

buffer depth for the feed forward type routers. The improvement of switching performance is regulated by LRD traffic. In this case, the PSB type routers using single re-circulation port can only play the role of buffer depth extension. It doesn't show any benefit of coping with asynchronous and variable length packet switching. However, during the period of packet-by-packet verification for our simulator, we find that there are two driving forces in the PSB type routers affecting the probability of packet loss. One is the increase of delay line length in the buffers. The longer the delay line length is, the higher probability of packet contention is. The other is traffic smoothing capability due to the common shared re-circulation buffer. The longer the delay line length is, the more powerful the traffic smoothing capability is. These two driving forces cancel the effect of each other on the probability of packet loss due to LRD traffic.

In order to compare this phenomenon thoroughly, we analyze another extreme case of the PSB type routers using 16 re-circulation ports and without using void filling algorithm in Fig. 13. To compare with various operation modes in feed forward type routers, we depict the performance of 16 re-circulation ports PSB type routers without using void filling and with MPR activated in the same figure. We see that the PSB type WDM optical routers using 16 re-circulation ports without void filling can have comparable performance as the feed forward type routers without void filling but with MPR activated in the region of longer delay line length. We also find the same result in Fig. 13, for the performance comparison between the PSB type routers using 16 re-circulation ports with MPR activated but without void filling and the feed forward type routers with both void filling algorithm and MPR capability activated in the region of longer delay line length. Surprisingly, we find that the performance of the PSB type WDM

optical routers using 16 re-circulation ports has the similar properties with the feed forward type routers with void filling algorithm activated for the delay line length in both of with and without MPR activated modes.

This phenomenon is not straightforward as the void filling algorithm. We observe that during the period of packet-by-packet verification for our simulator, the driving force from the traffic smoothing is more powerful than the increase of packet contention probability resulted from the longer delay line length of the fiber delay line buffers. When there are more packets coming into the routers by using larger delay line length, it will cause more packets dropped in feed forward type routers. If we use shorter delay line length, the packet will come out from the re-circulation buffers more quickly. However, since the input traffic is LRD, it means that the traffic has a heavy tailed probability of having very long packet length. These long packets will occupy the output buffers for a long time. Then, the packets coming out from the re-circulation buffers have to re-circulate in the routers again and again. This will cause higher packet loss probability. These results show that we can use the PSB type router with suitable re-circulation ports to handle asynchronous and variable length packets without using the computing-consuming void filling algorithm at the expense of system complexity.

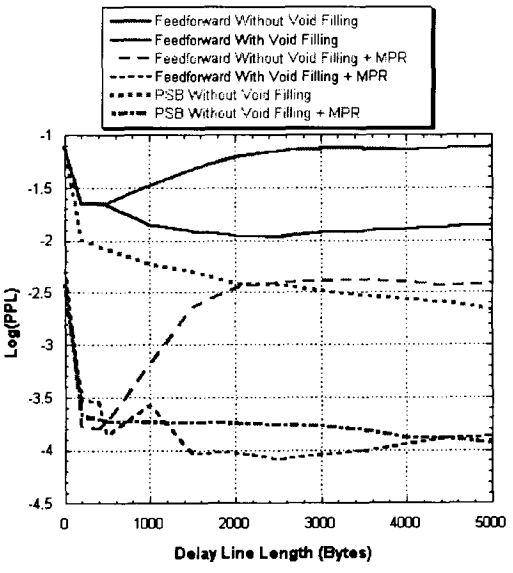


Fig. 13 Performance of PSB type WDM optical router with MPR activated under LRD traffic

【第二年】

A. Unicast Switch

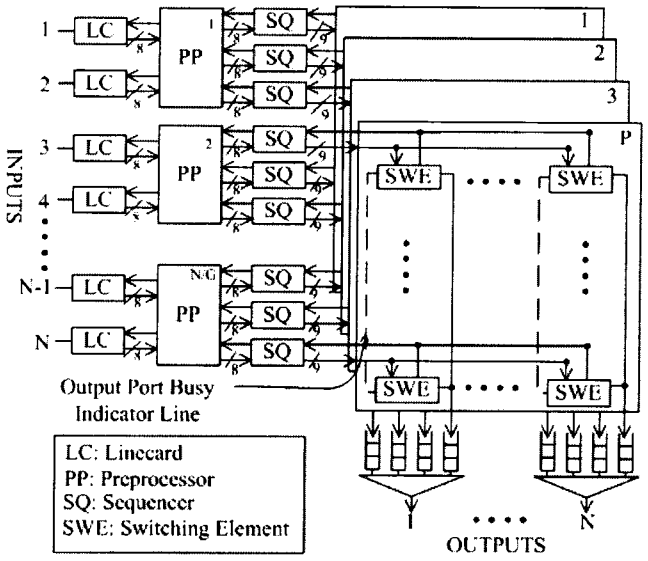


Fig. 1. The architecture of the NTU switch.

Fig. 1 depicts the architecture of the proposed optical switch, called NTU Switch. There are N input ports (output ports). The transmission speed of an input line and an output line is the same as that of the switching fabric. A header and a trailer will be attached to each incoming IP packet in linecards. The output port address, which is assigned according to the IP address, is placed in the header. The trailer acts as an ending delimiter. We let every G input ports of the switch group together to share a preprocessor. The number of output lines of the preprocessor is $G \cdot R$, where R is an expansion ratio. An output line of the preprocessor is connected to an input line of the sequencer followed by the crossbar switching planes. We denote the number of planes and the number of sequencers of each plane by P and NR/P respectively. Each crossbar switching plane consists of N^2R/P switching elements, which form an $(NR/P)\times N$ matrix. The structures of all switching elements are identical except with differently stored output port addresses. The outputs of switching planes are connected to N multiplexers. Each multiplexer has P inputs and single output.

I. Preprocessor

Each preprocessor accepts the packets from a group of G input ports and distributes them to appropriate sequencers associated with the preprocessor. The bit stream of an incoming packet arrives on an eight-bit

word bus. The sequencer will append a control bit to this word to form a nine-bit word before sending this packet into the crossbar switching plane. The control bit of a header and a trailer is set to 1. For a data byte, it is set to zero. The packet format with nine-bit words is illustrated in Fig. 2. Note that the format of the trailer is identical with that of the header but it consists of an invalid output port address.

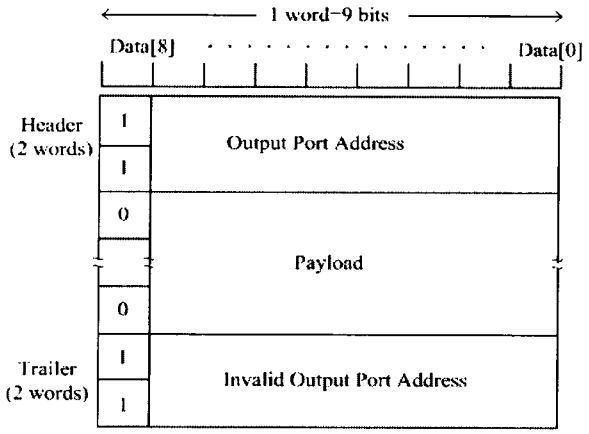


Fig. 2. The packet format of the NTU switch.

II. Sequencer

The sequencers associated with the same plane send the packets into the plane in a round robin fashion. All sequencers can start to send the packets only at pre-assigned time points, which are illustrated in Fig. 3, where τ is the propagation delay of an output port busy indicator signal in the plane from the first switching element to the last switching element in the same column. When the packet is successfully transmitted to the destination port via a certain switching element, the switching element will send a feedback signal to inform the sequencer continuing the transmission until finished. Otherwise the sequencer will retransmit the packet at the next allowable entering time point. The sequencer may buffer more than one packet. If a packet is not successfully transmitted after several attempts, the sequencer may send the next packet waiting in the buffer to the switch plane at the following allowable entering time points.

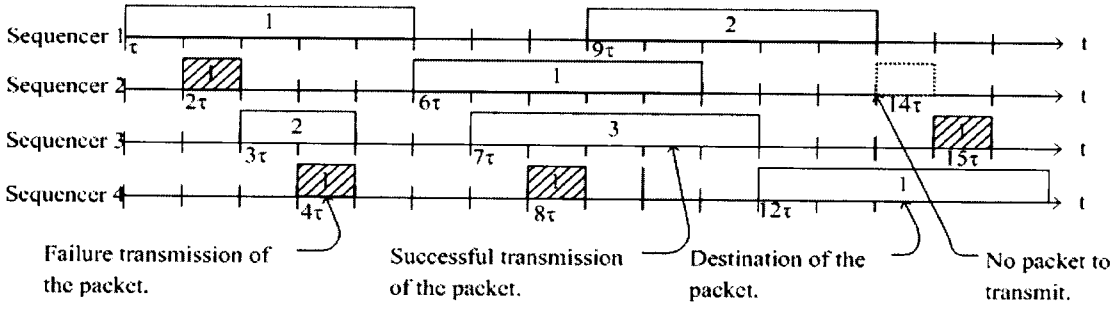


Fig. 3. The timing diagram of the sequencers in the same plane. We show the case with 4 sequencers.

III. Crossbar Switching Plane

The switching fabric is composed of P planes. For each plane, the switching elements in the same column corresponding to a single output port are connected to the output buffer via a vertical bus and an output port busy indicator line. The switching elements in the same row are connected to the horizontal bus and a feedback signal line associated with the sequencer. The schematic diagram of one switching element is shown in Fig. 4. It has two states: cross and bar. Initially, all the switching elements are in the bar state. The feedback signal and the output port busy indicator signal are idle. The switching element always listens to the horizontal bus. The control bit 1 will activate the correlator to examine the address in the nondata word. If the two-word packet address matches the prestored output port address and the output port busy indicator signal is idle, then the switching element changes its state to cross and activates its connector to route the packet to the output port. At the same time it sends a feedback signal to the sequencer and sets the output port busy indicator. When any nondata words, which do not match the output port address, are detected, the switching element releases its connector and resets the output port busy indicator line and the feedback signal line idle. Simultaneously the state is changed back to its initial bar state. Because the trailer

of a packet is designed to consist of an invalid output port address, it will always reset the switching element to the initial state and therefore finish transmission of the current packet.

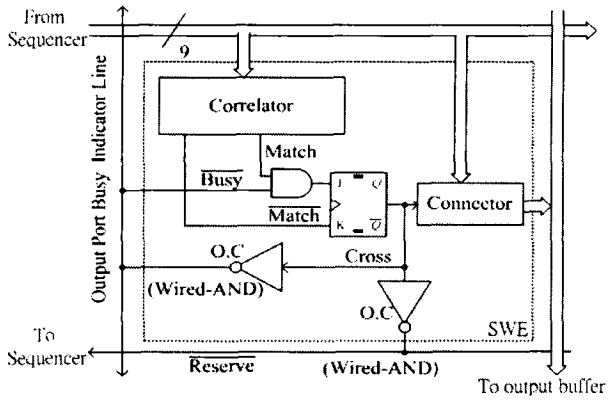


Fig. 4. The block diagram of the switching element.

IV. Analysis of the crossbar switching plane

We analyze a single crossbar switching plane first because it is the place where blocking of the packet streams occurs. Then we use the result of this analysis to evaluate the performance of the overall system.

1. Model Assumptions

In an $M \times N$ crossbar switching plane, each sequencer has a buffer with storage capacity of K packets, including the transmitting one. The system time is divided into slots, and M slots are grouped into one frame. Each slot is equal to τ and set to one. Each sequencer is assigned permanently to a slot in the frame. The backlogged sequencer can send one packet in its assigned slot to the plane to perform self-routing switching. The packet arrival process of each sequencer is i.i.d. (independent and identically distributed) Bernoulli process with rate σ and the length of each packet is i.i.d. geometric distributed with mean μ^{-1} slots. Each packet chooses one of the N output ports as its destination with equal probability. The late arrival model with delayed access is used in our analysis and simulation [8]. We use delayed replacement rule when a packet arrives in the last slot of a packet transmission time and the buffer is full, that is, the sequencer does not accept such a packet and packet loss occurs [9].

2. State Definition

The modes of a sequencer are: idle, backlogged, and transmission. The sequencer with empty buffer is in the idle mode, and it is in the transmission mode while transmitting a packet. The backlogged sequencer has packets in its buffer and is waiting for its assigned slot to start self-routing. In addition, the number of packets in the buffers is necessary for the sequencer to determine whether to accept the arriving packets or not. Moreover, after a packet departure, the mode of the sequencer is determined by its buffer status. Hence, we classify the sequencers as idle, backlogged or transmission with k packets in its buffer, where $k = 1, 2, \cdots, K$.

Packet arrival and departure in each sequencer occur during one slot with the probabilities σ and μ respectively. However, at the beginning of one slot, only the assigned sequencer can start self-routing. The self-routing sequencer will start transmitting if it doesn't collide with others, otherwise, it will enter backlogged mode. For convenience, we number all the sequencers from 1 to M and rotate the number with progress of slots, i.e., in the next slot, Sequencer i in the current slot will become Sequencer $i+1 \pmod{M}$. Without loss of generality, the current slot is assigned to Sequencer 1 which can start self-routing if it has packets waiting in its buffer. We observe the system at the beginning of slot t and define the state vector as $\mathbf{V}^{(t)} = \left(\mathbf{V}_1^{(t)}, \cdots, \mathbf{V}_j^{(t)}, \cdots \mathbf{V}_M^{(t)}\right)$, where $\mathbf{V}_j^{(t)} = (i_j^{(t)}; b_{j,1}^{(t)}, \ldots, b_{j,K}^{(t)}; t_{j,1}^{(t)}, \ldots, t_{j,K}^{(t)})$ and $i_j^{(t)}$ =Pr[Sequencer j is idle in the time slot t], $b_{j,k}^{(t)}$ =Pr[Sequencer j is backlogged with k packets in its buffer in the time slot t], $t_{j,k}^{(t)}$ = Pr[Sequencer j is transmitting packets with k packets in its buffer in the time slot t]. For Sequencer j, we have

$$i_j^{(t)} + \sum_{k=1}^{K} b_{j,k}^{(t)} + \sum_{k=1}^{K} t_{j,k}^{(t)} = 1 \qquad \forall\ j = 1, 2, \cdots, M.$$

Note that we have defined a multi-dimensional Markov chain $\{\mathbf{V}^{(t)}, t = 1, 2, \cdots, \infty.\}$.

The stochastic behaviors of Sequencers 1,..., M are identical and independent. Therefore, we can determine the state probabilities of Sequencer 2, ..., M in the next slot given the state probabilities of Sequencer 1,..., $M-1$ in the current slot:

$$\mathbf{V}_j^{(t+1)} = \mathbf{V}_{j-1}^{(t)} \cdot \mathbf{C}, \qquad \forall\, j = 2, 3, \cdots, M$$

where $\mathbf{C}$ is the routing probability matrix. Because self-routing is possible for Sequencer 1, and the probability of successful switching is determined by the modes of the transmitting sequencers, we have

$$\mathbf{V}_1^{(t+1)} = (1-s)\mathbf{V}_M^{(t)} \cdot \mathbf{D}_1 + s\mathbf{V}_M^{(t)} \cdot \mathbf{D}_0$$

where s =Pr[self-routing is successful], and $\mathbf{D}_1$ ($\mathbf{D}_0$) is the routing probability matrix given that self-routing fails (succeeds). The collision will occur if the self-routing packet is destined for an output port to which one of Sequencer 2,..., M is transmitting. Based on the assumption of uniform traffic and using the law of total probability, we have

$$s = \sum_{m=0}^{M-1} T_m \left(1 - \tfrac{1}{N}\right)^m,$$

where T_m =Pr[There are m transmitting sequencers among Sequencer 2,..., M]. When the buffer of a sequencer can store only one packet, the corresponding routing matrices are given as follows.

$$\mathbf{C} = \begin{pmatrix} 1-\sigma & \sigma & 0 \\ 0 & 1 & 0 \\ \mu & 0 & 1-\mu \end{pmatrix},$$

$$\mathbf{D}_1 = \begin{pmatrix} 1-\sigma & \sigma & 0 \\ 0 & 1 & 0 \\ \mu & 0 & 1-\mu \end{pmatrix},$$

$$\mathbf{D}_0 = \begin{pmatrix} 1-\sigma & 0 & \sigma \\ 0 & 0 & 1 \\ \mu & 0 & 1-\mu \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11}[p,q] & 0 \\ \mathbf{C}_{21}[p,q] & \mathbf{C}_{22}[p,q] \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11}[p,q] & \mathbf{D}_{12}[p,q] \\ \mathbf{D}_{21}[p,q] & \mathbf{D}_{22}[p,q] \end{pmatrix}$$

where

$$\mathbf{C}_{11}[p,q] = \begin{cases} 1-\sigma, \text{if } p = q, 1 \le p \le K, \\ \sigma, \quad \text{if } p = q-1, 1 \le p \le K, \\ 1, \quad \text{if } p = q = K+1, \\ 0, \quad \text{otherwise.} \end{cases} \quad \mathbf{C}_{21}[p,q] = \begin{cases} \mu(1-\sigma), \text{if } p = q, 1 \le p \le K-1, \\ \mu\sigma, \quad \text{if } p = q-1, 1 \le p \le K-1, \\ \mu, \quad \text{if } p = q = K, \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\mathbf{C}_{22}[p,q] = \begin{cases} (1-\sigma)(1-\mu), \text{if } p = q, 1 \le p \le K-1, \\ \sigma(1-\mu), \quad \text{if } p = q-1, 1 \le p \le K-1, \\ 1-\mu, \quad \text{if } p = q = K, \\ 0, \quad \text{otherwise,} \end{cases}$$

$$\mathbf{D}_{11}[p,q] = \begin{cases} 1-\sigma, \quad \text{if } p = q = 1, \\ \mathrm{R}(1\text{-}\sigma), \text{if } p = q, 2 \le p \le K, \\ R, \quad \text{if } p = q = K+1, \\ \mathrm{R}\sigma, \quad \text{if } p = q-1, 1 \le p \le K, \\ 0, \quad \text{otherwise.} \end{cases} \quad \mathbf{D}_{12}[p,q] = \begin{cases} \sigma(1-R), \quad \text{if } p = q, 1 \le p \le K, \\ (1\text{-}\sigma)(1-\mathrm{R}), \text{if } p = q+1, 2 \le p \le K, \\ 1-R, \quad \text{if } p = K+1, q = K, \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\mathbf{D}_{21}[p,q] = \begin{cases} \mu(1-\sigma), \text{if } p = q = 1, \\ \mathrm{R}\mu(1\text{-}\sigma), \text{if } p = q, 2 \le p \le K-1, \\ R\mu, \quad \text{if } p = q = K, \\ \mathrm{R}\mu\sigma, \quad \text{if } p = q-1, 1 \le p \le K-1, \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\mathbf{D}_{22}[p,q] = \begin{cases} (1-\sigma)(1-\mu) + \mu\sigma(1-R), & \text{if } p = q, 1 \le p \le K-1, \\ 1\text{-}\mu, & \text{if } p = q = K, \\ \mu(1-\sigma)(1-R), & \text{if } p = q+1, \ 2 \le p \le K-1, \\ \mu(1-\mathrm{R}), & \text{if } p = K, \ q = K-1, \\ \sigma(1-\mu), & \text{if } p = q-1, \ 1 \le p \le K-1 \\ 0, & \text{otherwise,} \end{cases}$$

where σ and μ are defined in the assumption of traffic model. According to the definition, $\mathbf{D}_1$ ($\mathbf{D}_0$) is equal to $\mathbf{D}$ with $R = 1$ (0).

V. *Limiting Probabilities*

We assume the system is in the steady state and denote the steady state vector by $\mathbf{V}$. Using the above equations, we have

$$\mathbf{V}_j = \mathbf{V}_1 \cdot \mathbf{C}^{j-1}, \forall j = 2, \cdots, M,$$

and

$$\mathbf{V}_1 = (1-s)\mathbf{V}_M \cdot \mathbf{D}_1 + s\mathbf{V}_M \cdot \mathbf{D}_0 \ .$$

The steady state probabilities are in terms of $\mathbf{V}_1$, which satisfies a nonlinear equation

$$\mathbf{V}_1 = \mathbf{V}_1 \cdot \mathbf{C}^{M-1} \cdot \left[(1-s)\mathbf{D}_1 + s\mathbf{D}_0\right]$$

We define the generating function[10] to count the events that m sequencers are transmitting packets: $S(x) = \prod_{j=2}^{M}\left[1+(x-1)\sum_{k=1}^{K} t_{j,k}\right]$. In $S(x)$, we have counted the probabilities of transmission for Sequencer j, i.e., $t_{j,k}$ by associating them with x. Hence, T_m is equal to the coefficient of x^m in $S(x)$ and hence $S^{(m)}(0)/m!$. Because $\mathbf{C}$ is a stochastic matrix, we have

$$i_1 + \sum_{k=1}^{K}\left(b_{1,k} + t_{1,k}\right) = 1$$

Finally, we can solve for $\mathbf{V}_1$ by using Newton's Method for systems of nonlinear equations.

VI. *Analysis of performance metrics*

1. *Definition*

We denote the throughput of the switch by γ, which is defined as the average utilization factor per output port. The generating function which counts the number of sequencers in the transmission mode is $\Gamma(x) = \prod_{j=1}^{M}\left\{i_j + \sum_{k=1}^{K}\left(b_{j,k} + x \cdot t_{j,k}\right)\right\}$. The coefficient of x^m in $\Gamma(x)$ is the probability that there are m sequencers in transmission mode. Hence the expected number of transmitting sequencers is $\Gamma'(1)$. After using the normalization condition, we obtain $\gamma = \frac{\Gamma'(1)}{N} = \sum_{j=1}^{M}\sum_{k=1}^{K}\frac{t_{j,k}}{N}$. Because of conservation of the traffic flow, we have $\frac{\sigma}{\mu}(1-\beta) = \gamma\frac{N}{M}$ in the steady state, where we denote β as the blocking probability. Hence,

$$\beta = 1 - (\gamma\mu\sigma / \sigma M)$$

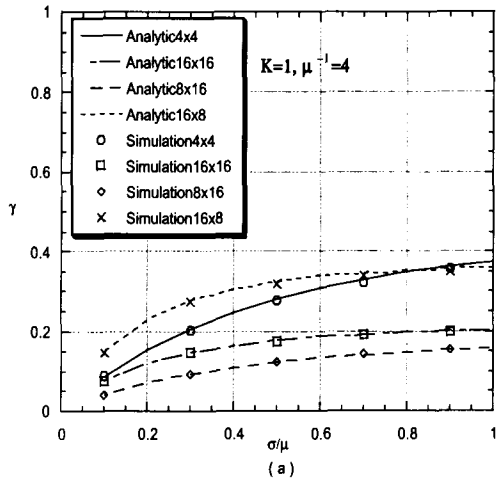
We define the delay experienced by each packet in the sequencer as the total time it spends from the instant it arrives at the sequencer buffer to the instant when it finishes transmission. The associated generating function for the system size is $Q(x) = \prod_{j=1}^{M} \left\{ i_j + \sum_{k=1}^{K} (b_{j,k} + t_{j,k}) x^k \right\}$. The coefficient of x^m in $Q(x)$ is the probability that there are m packets in the system in the steady state. Therefore the expected systems size is $Q'(1)$. Using the well-known Little's Law [10], we obtain the packet delay as follows

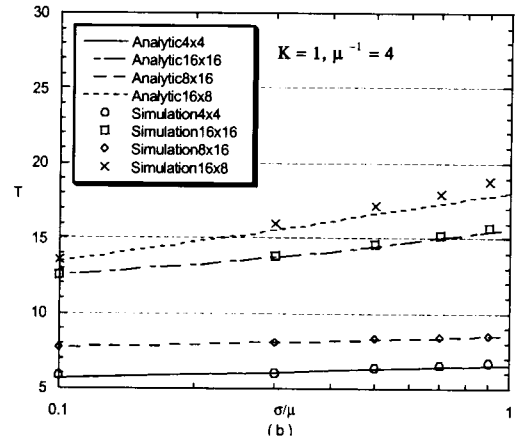
$$T = Q'(1) \Big/ N\gamma\mu = \sum_{j=1}^{M} \sum_{k=1}^{K} k\left(b_{j,k} + t_{j,k}\right) \Big/ N\gamma\mu$$

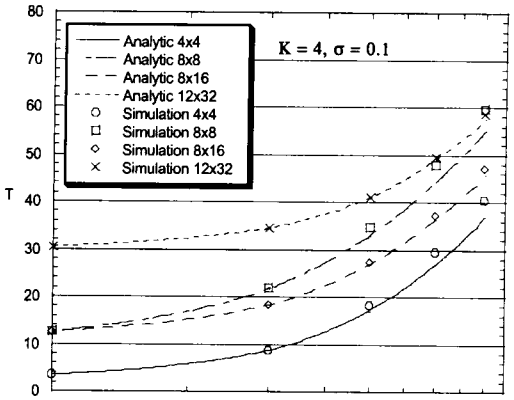
.

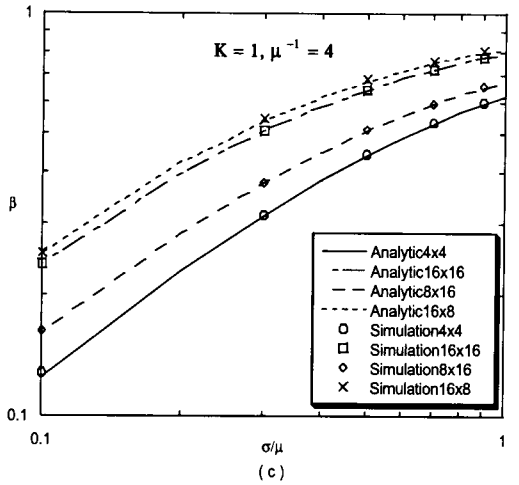
2. *Validation of the Model*

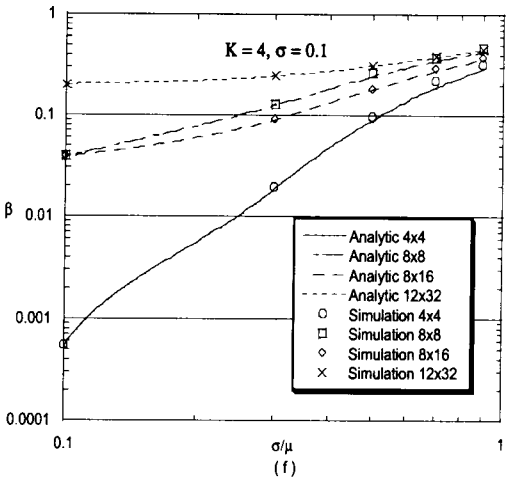
We use a stochastic self-driven discrete event simulation to verify the correctness of our analytic model. The assumption of simulation is identical with the analytic model. We measure the time average utilization per output port as the throughput. The delay is measured by averaging over all the packets. Finally, the blocking probability is defined as the ratio of the total number of discarded packets due to overflow of the buffers to the total number of arriving packets. Fig. 5 plots the comparison of simulation results with the analytic model. Both the unbuffered case ($K = 1$) and general case ($K = 4$) are verified. It can be seen that the results obtained through the analytic model agree with simulation very well.











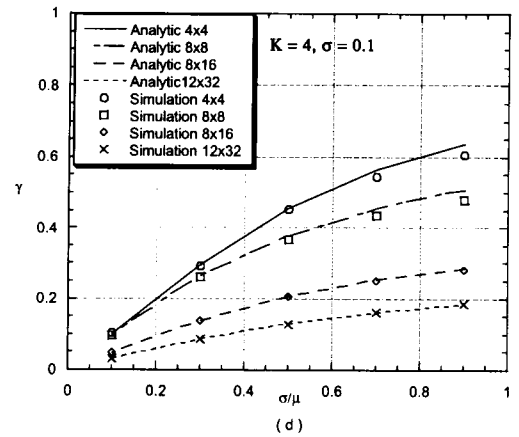


Fig. 5. Comparison of the simulation and analytic result for (a)throughput, K=1, μ^{-1} =4, (b)delay, K=1, μ^{-1} =4, (c)blocking probability, K=1, μ^{-1} =4, (d)throughput, K=4, σ =0.1, (e)delay, K=1, σ =0.1, (f)blocking probability, K=1, σ =0.1.

VII. Performance of the multiplane switch

We assume that the packet arrival processes at the linecards are i.i.d. Bernoulli process with rate λ. The incoming packet will be sent to the connected preprocessor, where it chooses one of the connected sequencers with the same probability. Hence, the arrival rate of the Bernoulli process at each sequencer σ is equal to λ/R, where R is the expansion ratio.

1. Throughput for Infinite Buffer

Let us assume for the moment that the storage capacities of input and output buffers are infinite. Thus, we are considering a no-loss system. We denote the throughput per output port of the crossbar switching plane by γ. The overall switch throughput per output port is the aggregate throughput of the outputs of the planes $\gamma_o = \sum_{l=1}^{P} \gamma_l = P\gamma$. We are interested in the maximum throughput $\gamma_{\max}$ which occurs at saturation. When the system is at saturation, there are always packets waiting in the buffer of the sequencer to be sent. Thus, there are no idle modes and the state vector $\mathbf{V}$ equals $(b_{1,1},\ldots,b_{M,1};t_{1,1},\ldots,t_{M,1})$. We have the routing probability matrix $\mathbf{C}^{(s)} = \begin{pmatrix} 1 & 0 \\ \mu & 1-\mu \end{pmatrix}$ and $\mathbf{D}^{(s)} = \begin{pmatrix} R & 1-R \\ R\mu & 1-\mu+\mu(1-R) \end{pmatrix}$. The associated generating function with throughput is $S(x) = \prod_{i=2}^{M}\left[1+(x-1)t_{i,1}\right]$. Then we can use the state transition equations and the normalization condition $\sum_{i=1}^{M}(b_{i,1}+t_{i,1}) = 1$ to solve for

the steady state vector $\mathbf{V}$. Finally, the maximum throughput is $\gamma_{\max} = \frac{P}{N}\sum_{j=1}^{M} t_j$.

Fig. 6 shows the maximum throughput with different expansion ratio R and number of planes P. We see that the HOL blocking is alleviated by increasing the number of planes and expansion ratio. Moreover, the mean packet length will also affect the maximum throughput. For a 32×32 switch, we see that 100% throughput is attained with four planes and expansion ratio 1.5 if the mean packet length is larger than 20τ. This is because when the packet length is larger compared to the frame duration, the fraction of time the sequencer spends in waiting its turn to perform self-routing will be smaller. However, the large packet will also block the transmission of the other packets destined for the same output port. Hence, the gain of throughput will decrease as the packet length increases.

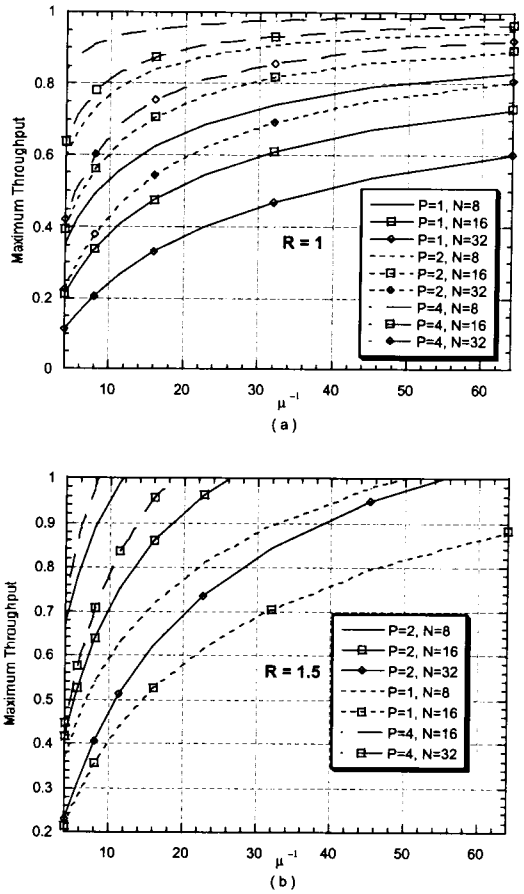


Fig. 6. The maximum throughput for infinite buffer for (a)R=1, (b)R=1.5.

2. Blocking Probability

In this and next subsection, we consider a loss system. Assume that the buffer capacity in each sequencer is K packets and the output buffer of each output port can store K_o packets. An output queue for each output port is shared by all P planes and no backpressure mechanism is used. We model the output buffer by a discrete time $Geom/Geom/1/K_o$ queueing

system. The offered load of the output buffer is $P \cdot \gamma$, and the departure rate is μ. We can use a Markov chain analysis to derive the state probability of the $Geom/Geom/1/K_o$ system [11]

$$P_n = \frac{P_0}{1-\mu}\left(\frac{\lambda_o(1-\mu)}{\mu(1-\lambda_o)}\right)^n, \quad n = 1, \cdots, K_o,$$

and

$$P_0 = (\mu - \lambda_o) \Big/ \left[\mu - \lambda_o\left(\frac{\lambda_o(1-\mu)}{\mu(1-\lambda_o)}\right)^{K_o}\right],$$

where $\lambda_o = P \cdot \gamma \cdot \mu$.

Because of the BASTA property of the arrival process [12], the blocking probability at the output buffer (P_{B2}) is equal to the probability of state K_o, that is, $P_{B2} = P_{K_o}$. Using conservation of packet flows for an $N \times N$ switch, we have $\lambda/\mu(1-P_B) = P\gamma(1-P_{B2})$, where P_B is the overall blocking probability. Therefore, we have $P_B = 1 - \frac{P \cdot \gamma \cdot \mu}{\lambda}(1-P_{B2})$.

Fig. 7 shows the blocking probability of a 16×16 switch, where we denote by L the mean packet length. We find that the blocking probability is almost the same when the packet length is 10 or 100 for R=1.5 and P=4. This is because when packet length is larger than 8 the maximum throughput of a 16×16 switch has attained 100% if $R = 1.5$ and $P = 4$, and hence the performance of the switch doesn't change greatly with the packet length. The blocking probability is more sensitive to the change of output buffer capacity than that of input buffer capacity. Increasing the capacity of the output buffer will obtain more gain in blocking probability than increasing the capacity of the input buffer. However, the more the input buffer capacity is, the more the gain is obtained when we increase the output buffer capacity. There exists an upper bound beyond which the gain doesn't increase especially in the light load. Finally, we see that the maximum throughput with infinite buffer capacity can be attained if we increase both the capacities K and K_o. When $R = 1$ and $P = 2$, the

switch is input buffer limited and therefore the gain of the blocking probability doesn't increase as the output buffer capacity increases when the input buffer capacity is small. The blocking probability will be smaller when the mean packet length is 100. Comparing with the case R =1.5 and P =4, we see that the more alleviation of blocking effect using more planes and larger expansion ratio.

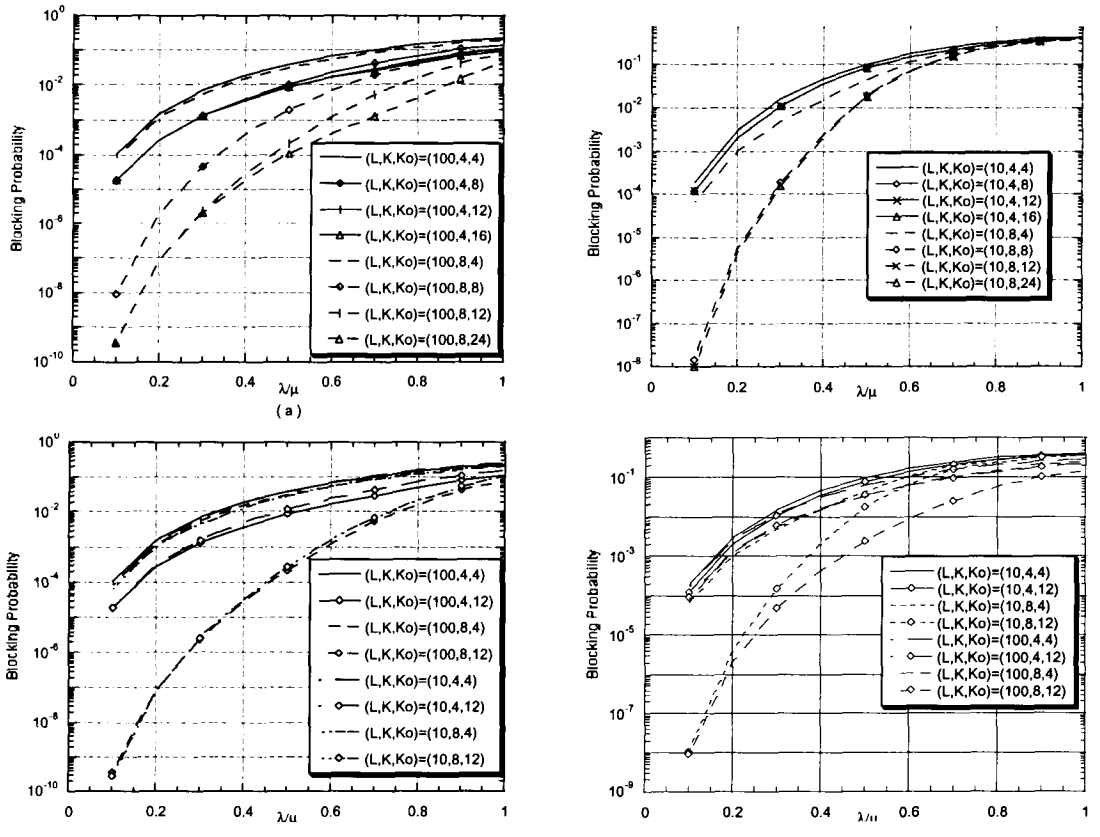


Fig. 7. The blocking probability of a 16×16 NTU switch for (a)P=4, R=1.5, L=100, (b)P=4, R=1.5, L=100 or 10, (c)P=2, R=1, L=10, (d)P=2, R=1, L=100 or 10.

3. *Waiting Time*

The waiting time is defined as the time interval from the time an arbitrary packet enters the buffer to the time the packet leaves the buffer. Storing-and-forwarding switching is used in our model, hence the minimum waiting time is twice the packet length. We denote the mean waiting time of the switch by W . Clearly, $W = T + T_o$, where T is the mean waiting time spent in the sequencer buffer and T_o is the mean waiting time spent in the output buffer. T_o is obtained by using Little's Law:

$$T_o = \sum_{n=1}^{K_o} nP_n \Big/ \lambda_o (1 - P_{B2}) .$$

Fig. 8 shows the mean waiting time of packets in units of mean packet length in a 16×16 switch. We see that packets with mean length of 100 will spend less time in units of mean packet length waiting then to the packets with mean length of 10. The difference in the behavior of the waiting time is more

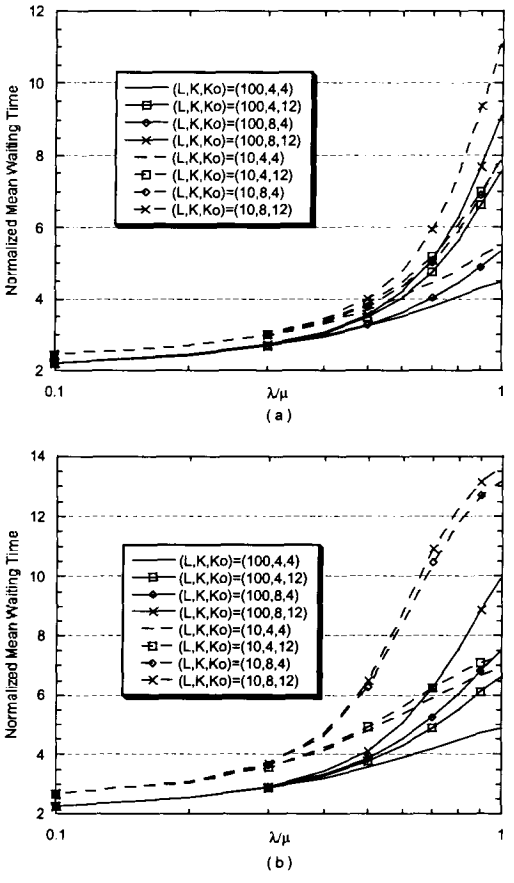


Fig. 8. The normalized mean waiting time of a 16×16 switch for (a)P=4, R=1.5, (b)P=2, R=1.

obvious when the switch is input buffer limited, i.e., $R = 1$ and $P = 2$. Therefore, it is more efficient to operate in longer packet length when the switch is input buffer limited. Note that when the switch is output buffer limited, that is, $R = 1.5$ and $P = 4$, increasing the input buffer capacity will obtain much gain in the blocking probability if the sufficient output buffer capacity is used. This is attained at the expense of increase in the mean waiting time.

VIII. CONCLUSION

We analyze the queueing system of the crossbar switching plane, where blocking of packet streams occurs. The maximum throughput, blocking probability, and waiting time of the overall system are obtained based on this analysis. As seen in the analysis, although the input port expansion and multiplane structure increase the switch complexity a little, it will significantly improve the throughput. Our result shows that a reasonably sized NTU switch with four planes and 1.5 expansion ratio will have 100% maximum throughput, which matches OQ switches. Moreover, the result shows that the blocking probability and waiting time of the switch is output buffer limited when enough planes and expansion ratio are used.

B. Multicast Switch

There are many multicast switches proposed and analyzed in the literature [13]. However, a little is done about the multicast scheduling of variable length packets. In this report we assume that an incoming packet has multiple destinations and is of variable length without being fragmentized into cells. Hence a new architecture must be designed to solve output contention among all the multicast destinations, to replicate the packets and finally to route these copies to their destinations respectively. The complication arises because the availability of outputs is of random nature.

Multicast queueing disciplines were discussed in [14] for fixed length packets. It was shown that the multicast service with Fanout splitting has better throughput and delay performance. However, the performance of a switch capable of handling variable length packets is not discussed. In this report, we assume a random HOL service with Fanout splitting. Multiple copies of an

incoming packet can be served at the same time, depending on the availability of the destinations. The HOL packet in an input is completely served if all its copies are transmitted. Then the next fresh packet in the queue can move into the HOL position. It will be shown that the maximum throughput depends on not only the mean of the Fanout distribution as expected, but also on a quantity termed Fanout Function, whose characteristics will be surveyed in detail. The mean waiting time will also be derived.

The HOL queueing discipline for multicast packets with variable length is more complicated than that for fixed length packets because of the asynchronous nature of output availability. Hence, it is difficult to schedule the routing of all the multicast copies by a central arbitrator. Instead, distributed packet replication, decentralized output contention resolution and self-routing are very desirable. In this report we propose a multicasting switch based on the previous architecture to provide multicast switching function with random HOL service and Fanout splitting. This switch, which is a combined input output queue switch, can achieve 100% maximum throughput and hence emulates an Output Queue switch by using multiple planes, input grouping and expansion techniques.

I. Multicast Switch Operation

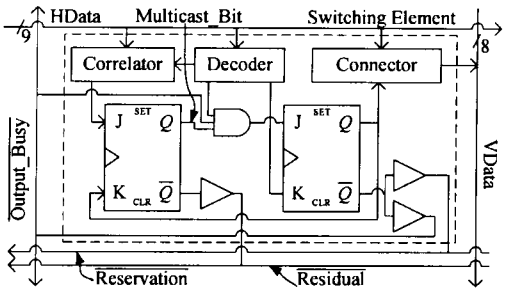


Fig.9. The switching element of multicast switch.

The architecture is Fig.9. For multicast switching, the linecard generates a multicast address packet and a data packet for each incoming packet. The multicast address packet consists of a truncated bit map of the incoming packet's multicast destination set, and

the multicast data packet carries the data unit of the incoming packet. Each output has a preassigned bit in the bit map. For example, in a bit map $\{b_i : i=1,2,\cdots,N\}$, b_i is set to 1(0) if output i is(not) in the multicast destination set, for i =1,2,…N. The transmission of the whole bit map is not necessary when output i are not in the multicast destination set for $i = I+1,\cdots,N$, where I is the last output port number in the multicast destination set. Because both kinds of packets are of variable length, trailers are appended to indicate the ending.

The linecard sends the multicast address and data unit of its HOL packet to one of the shared sequencers via the coordination of the preprocessor. The sequencer can broadcast a packet to all the switching elements connected to it. The schematic diagram of the switching element is shown in Fig.9. A switching element consists of a multicast bit, which is represented by the output of a J-K Flip-Flop as shown in Fig.9. The switching elements in the same column of a plane are connected to the same output and interconnected by an output busy indicator line. Once the multicast address packet is sent to a sequencer, the sequencer will broadcast this packet immediately. The connected switching element will check whether it is in the multicast destination set and if so, its multicast bit will be set to high, or else it is set to low. Note that the multicast address packet of an incoming packet is sent only one time to solve the multicast addressing problem.

The sequencers in the same plane must send a data packet in a round-robin fashion to prevent output conflict. The sequencer waits for its preassigned time point to broadcast the multicast data packet. When a data packet arrives, the switching element, whose multicast bit is high, will check the output busy indicator line output busy. If this line is high, the switching element will route the incoming data packet to the

connected output, reset its multicast bit to low, and set the connected output busy indicator line output busy to low. Otherwise, the switching element discards the data packet and sends a NACK signal to the sequencer through the signal Reservation. The switching elements must also send the updated multicast bit to the sequencer through the signal Residual.

Then, the sequencer checks signals Reservation and Residual. If Reservation is high, it means that there are no switching elements able to route the data packet and hence the sequencer will terminate this transmission. This data packet will be transmitted again at the next allowable time point. If Reservation is low, the sequencer will check Residual. If Residual is high, the data packet is already sent to all the multicast destinations and the sequencer completes the processing of the packet after this transmission. Otherwise, the sequencer will wait for the next allowable time point to transmit the data packet again. This process is repeated until transmission completion, i.e. Residual is high.

II. Performance Analysis

Assume that the packet arrival process at each input is independent and identical Poisson process with rate λ_{in}. The length of each packet is independently and identically exponentially distributed with mean $1/\mu$. For multicast switching, a packet may request delivery to multiple destinations. This multiplicity denoted by F is termed Fanout, and its probability mass function is $p_f \triangleq \Pr[F= f\ ], \forall\ f = 1,..., N$. The multicast destination set of each packet is a random sample without replacement of all the N outputs. That is, the first copy chooses one of the N outputs randomly, the second copy chooses one of the remaining N-1 outputs randomly, and so on. When N is much larger than the mean Fanout, we can approximate the destinations of a multicast packet to

be independently and uniformly distributed for all outputs, since the resulting event of having two identical destinations in a packet becomes unlikely for large N. Therefore, the packet copy flows are spread evenly into all of the N outputs.

The buffer capacity of each linecard is assumed to be infinite. Ignoring the overhead of preprocessor, we can model the G linecards and the associated G·R sequencers by an M/G/c queueing system, where the number of servers c is G·R and packet arrival rate is G·λ_{in}. The service time is defined as the time each packet spent in the sequencer.

1. *Service Time in Sequencers*

The service time distribution of each multicast copy, denoted by X_i,i=1,...,F, is exponential with rate q:

$$q = \frac{\sum_{j=1}^{N} e(V_j(t)) \cdot \mu}{\sum_{j=1}^{N} V_j(t)} \approx \frac{E[e(V_j)] \cdot \mu}{V} \quad (\text{as } N \to \infty)$$
$$= \frac{\lambda}{P \cdot V} = \mu \frac{\lambda}{P}.$$

And an incoming packet spends time X in the sequencer. The mean of X is

$$E[X] = M_X^{(1)}(0) = \frac{1}{q}\sum_{f=1}^{N} H_f \cdot p_f$$
$$= \frac{1}{q}E[H_F] = \frac{\eta \cdot E[F]}{\mu}\frac{P}{P - \rho_{in}},$$

where

$$\eta \triangleq \frac{E[H_F]}{E[F]}, \quad H_F \triangleq \sum_{i=1}^{F} i^{-1}$$

. $\rho_{in} \triangleq (\lambda/\mu)$ is the input load, and η is the Fanout Function. The coefficient of variation of X depends only on the random variable F:

where

$$C_X^2 \triangleq \frac{\sigma_X^2}{E[X]^2} = \frac{E\left[H_F^2 - H_F^{(2)}\right]}{E[H_F]^2} + 1,$$

$H_F^2 \triangleq (H_F)^2$ and $H_F^{(2)} \triangleq \sum_{i=1}^{F} i^{-2}$

2. *Maximum Throughput*

The maximum throughput per output, denoted by S, is attained when

$$\frac{G \cdot \lambda_{in}}{c} E[X] = 1.$$

We can solve λ_{in} under the saturation condition as

$$\lambda_{in} = \frac{R \cdot P \cdot \mu}{P \cdot E[H_F] + R \cdot E[F]}.$$

Then, we have the maximum throughput as

$$S \triangleq Min\left(\frac{\lambda_{in} \cdot E[F]}{\mu}, 1\right) = Min\left(\frac{R \cdot P}{R + P \cdot \eta}, 1\right).$$

3. Output Buffering Condition

For given multicast traffic with Fanout Function η and the 100% maximum throughput requirement, i.e. S=1, we can express the optimal values

$$P_o = \frac{R}{R - \eta} \text{ or } R_o = \frac{P \cdot \eta}{P - 1}.$$

of R_o and P_o as

For multicast switching with finite mean Fanout, $0<\eta<1$, and without expansion, R=1, it is necessary to use multiple planes to obtain 100% throughput. Moreover, the smaller η is, more planes are required.

The required R_o or P_o is smaller in the multicast switching than in the unicast switching. Therefore, a configuration of our system to emulate output buffering in unicast switching always also attains 100% maximum throughput in the multicast switching case.

4. Mean Waiting Time in Linecards

Because there is no exact solution of the M/G/c queueing system, we use the following approximations to evaluate the mean waiting time

$$E[W_q] = \left\{(1-\rho)\gamma\frac{c}{E[X]} + \rho\frac{1+C_X^2}{2}\right\}E[W_q^{\exp}],$$

where

$$E[W_q^{\exp}] = \frac{(c\rho)^c \rho}{c!\lambda_{in}(1-\rho)^2}\left\{\sum_{k=0}^{c-1}\frac{(c\rho)^k}{k!} + \frac{(c\rho)^c}{c!(1-\rho)}\right\}^{-1},$$

$$\gamma = (1 - C_X^2)\frac{E[X]}{c+1} + C_X^2\frac{E[X]}{c},$$

$$c = G \cdot R,\ \rho = \frac{G\cdot\lambda_{in}}{c}E[X].$$

The normalized mean waiting time, $E[W_q] \cdot \mu$, is a function of the input load ρ_{in} and the Fanout distribution. The Fanout will affect the mean waiting time through C_X^2.

C. Numerical Result

Assuming a 64×64 switch with P planes and R expansion ratio, the Fanout distribution can be deterministic or truncated Geometric distributed, both with the same mean Fanout E[F]. The deterministic Fanout is a fixed value of E[F].The truncated Geometric Fanout has possible values of 1,2,3,...,64. We use these two distributions to evaluate the maximum throughput and optimal expansion ratio as shown in Tables 1 and 2 respectively. The simulation values in Table 1 are obtained by using deterministic Fanout and the bound of 95% confidence interval given as (SimL,SimU). As a result, the deterministic distribution is a good candidate for Fanout distribution in performance evaluation because its Fanout Function is easy to evaluate and doesn't differ very much from other distributions.

A multicast switch with small expansion ratio and multiple planes will eliminate the HOL blocking, as shown in Table 2. As the mean Fanout E[F] increases, it can lower the switch complexity to less than N^2 by using enough planes. For example, if we use three planes when the mean Fanout E[F] is four, the switch complexity is $0.6931N^2$ for Geometric Fanout.

Although the number of linecards per preprocessor, G, doesn't affect the maximum throughput, it can lower the mean waiting time in the linecards as shown in Fig.10. When four linecards are grouped to share a preprocessor,the normalized mean waiting time is smaller than that when two linecards share a preprocessor.

EF	R	$P=1$		$P=2$	
		1	1.2	1	1.2
1	*All*	0.5000	0.5455	0.6667	0.7500
	SimL	0.4762	0.5150	0.6471	0.7156
	Sim	0.4920	0.5323	0.6699	0.7406
	SimU	0.5079	0.5497	0.6928	0.7657
	Det.	0.5714	0.6154	0.8000	0.8889
	Geo.	0.5906	0.6339	0.8381	0.9280
2	*SimL*	0.5557	0.5945	0.7984	0.8910
	Sim	0.5657	0.6056	0.8134	0.8975
	SimU	0.5758	0.6168	0.8284	0.9141
	Det.	0.6207	0.6626	0.9000	0.9910
	Geo.	0.6455	0.6860	0.9530	1.0000
3	*SimL*	0.6058	0.6390	0.9032	0.9759
	Sim	0.6124	0.6468	0.9150	0.9892
	SimU	0.6190	0.6546	0.9269	1.0000
	Det.	0.6575	0.6973	0.9796	1.0000
	Geo.	0.6839	0.7220	1.0000	1.0000
4	*SimL*	0.6442	0.6738	0.9759	1.0000
	Sim	0.6497	0.6807	0.9860	1.0000
	SimU	0.6551	0.6877	0.9960	1.0000

Table 1. The maximum throughput.

EF	P	2	3	4
1	*All*	2	1.5	1.3333
2	*Det.*	1.5000	1.1250	1.0000
	Geo.	1.3863	1.0397	0.9242
3	*Det.*	1.2222	0.9167	0.8148
	Geo.	1.0986	0.8240	0.7324
4	*Det.*	1.0417	0.7813	0.6944
	Geo.	0.9242	0.6931	0.6161

Table 2. The optimal expansion ratio.

D. Conclusion

We generalized the performance analysis of variable length packet unicast switching to random HOL multicast service with Fanout splitting. The Fanout F affects the maximum throughput through the Fanout Function, η. We formulate the maximum throughput analytically as a design guide.

The performance simulation results match the analytical calculation. The optimal number of planes and expansion ratio are derived to emulate the output buffer switch. The input grouping with larger G can lower the mean waiting time. It is known that the condition of the switch for unicasting is more stringent than multicasting and we can lower the number of switching elements to provide multicasting. In particular, the complexity of the switch can be lowered to less than N^2 by using multiple planes in multicast switching.

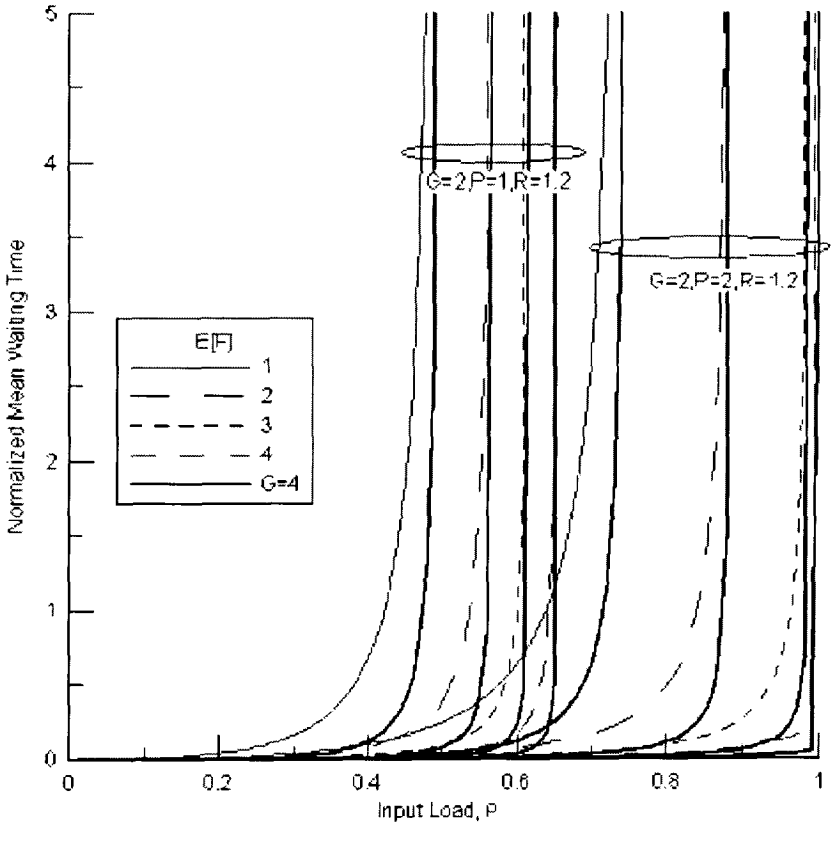


Fig. 10. The normalized mean waiting time.

【第三年】

A. Perfect Difference Code

Definition: Let W be the v-set of the integers 0, 1, 2, …, v-1 modulo v. A set $D=\{d_1, d_2, ..., d_k\}$ is a k-subset of W. For every $a \neq 0 \pmod{v}$, there are exactly λ ordered pairs (d_i, d_j), $i \neq j$, such that

$$d_i - d_j \equiv a \pmod{v}. \quad (1)$$

A set fulfilling these requirements is called a *perfect difference set*. Since there are $k(k\text{-}1)$ pairs and the total number for $a \neq 0 \pmod{v}$ is (v-1), the relation among v, k, λ is

$$\lambda = \frac{k(k-1)}{v-1}. \quad (2)$$

We can construct perfect difference codes $C_\tau = \{c_{\tau,0}, c_{\tau,1}, ... c_{\tau,(v-1)}\}$, $\tau \in W$, based on the cyclic difference set with the rule

$$c_{\tau,i} = \begin{cases} 1, & \text{for } i \in D_\tau \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The code weight, code length, and total number of codes are k, v, v, respectively, where $v = k^2 - k + 1$. The cyclic-shifted property of perfect difference codes makes the design of a tunable transmitter or a tunable receiver simpler.

The application of the perfect difference sets in the SOCDMA systems is investigated first time. We also propose a new SOCDMA system using these sets as the signature codes. Based on the cyclic-shifted and unity cross correlation properties of these codes, we propose the transmitter and receiver structures as shown in Fig.1 and Fig. 2.

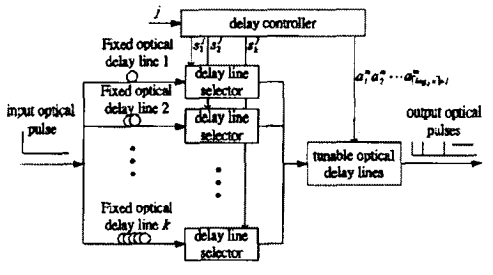


Fig. 1. Tunable transmitter structure.

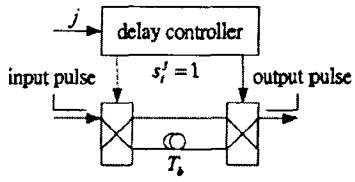


Fig. 2. Delay line selector in the *i*th branch.

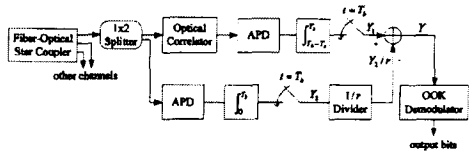


Fig. 3. The fixed receiver structure with optimal divider.

An optimal divider shown in Fig. 3 is used in the receiver to enhance the system performance. The numerical result with consideration of shot noise, thermal noise, APD bulk, and surface leakage currents is shown in Fig. 4. PDC with the proposed receiver structure can effectively eliminate the MUI and the asymptotic error floor, which appears in the SOCDMA systems with modified prime sequence codes. Moreover, the proposed system can provide reliable communication even under heavy load without reducing the number of usable codes. It is believed that the proposed system performs better than other SOCDMA systems.

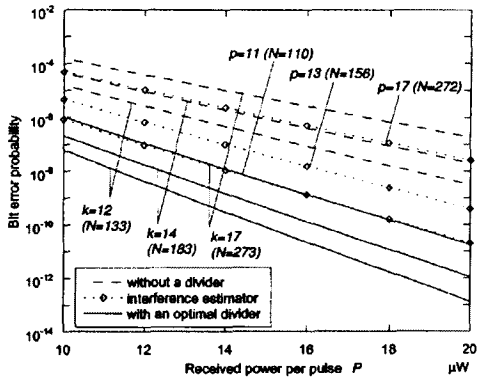


Fig. 4. Bit error probability comparisons under different code weights and full load.

B. Optical orthogonal codes with nonideal cross correlation

Recently, the construction and performance analysis of OOCs for optical code division multiple access (OCDMA) systems have been investigated widely. A $(v, w, \lambda_a, \lambda_c)$ - OOC is a family of (0, 1) sequences with code length v, code weight w, the

maximum value of off-peak autocorrelation λ_a, and the maximum value of cross correlation λ_c. Most studies paid attention to $(v,w,1)$-OOCs with $\lambda_a = \lambda_c = 1$ for the sake of synchronization and minimizing interference. However, under the constraint of the ideal auto- and cross-correlation properties, the code size is upper bounded by $\lfloor (1/w)\lfloor (v-1)/(w-1) \rfloor \rfloor$, which is linear to the code length, where $\lfloor x \rfloor$ denotes the greatest integer not greater than x. Therefore, the code size is very sparse with respect to the code length. In order to obtain a larger code size, we should relax the constraint. We drop some marks appropriately from k marks of each code, such that the modified code weight becomes w and the code length is still equal to v. Moreover, the cross-correlation property should satisfy the constraint of $\lambda_c = 2$, even when the two codes are aligned with each other. For example, the set {0, 1, 3, 7, 15, 31, 36, 54, 63} is a (73, 9, 1)-perfect difference set, where each element means the mark position. We can drop the first and the last five elements from the set, respectively, to form two subsets. Therefore, the two subsets form a family of (73, 4, 1, 2)-OOCs. We can further use the follow algorithm to obtain the code with maximal code size.

1) Construct a perfect difference set with elements according to (A), where $k = q^r + 1$ and $v = q^{2r} + q^r + 1$, q is a prime number and r is a positive integer.
2) Each element of an extended field $F = \mathrm{GF}(q^r) \cup \{\infty\}$, which also contains k elements, is related to one element of the perfect difference set.
3) Choose all the possible subsets with w elements from F and reserve the subsets fulfilling the property of a circle, where $w = q + 1$.
4) Each circle is related to w elements of the perfect difference set. The w elements form a code word with code length v and code weight w,

where each element represents the mark position. The codes form a family of $(v, w, 1, 2)$-OOCs.

The total number of codes is given by

$$T = q^{r-1} \cdot \frac{q^{2r} - 1}{q - 1} \quad . \tag{4}$$

Fig. 5 shows the bit error probability versus the received power per pulse of the systems using the four classes of OOCs under the number of simultaneous users N=200 and code length v=6643. When the received power P_W is small, these systems are power limited. When P_W is large, these systems are MUI limited. It also shows that the proposed codes with reversed codes perform better than ideal OOCs and Yang's codes. Moreover, the proposed codes have the largest code size. Fig. 6 shows the bit error probability versus the number of simultaneous users under the fixed received power $P_W = 0.5\mu$ W. At BER$\leq 10^{-9}$, the systems with the ideal OOCs have only about eight (with w=3) or 25 (with w=4) simultaneous users. The systems with Yang's codes have about 52 (with w=5), 83 (with w=6), or 115 (with w=7) simultaneous users. On the other hand, the systems with the two proposed codes accommodate about 190 users and have more users than the conventional coding schemes.

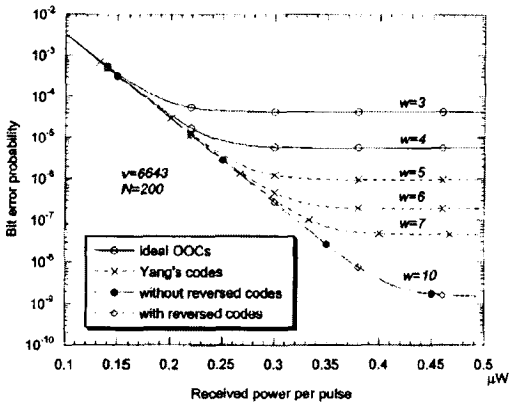


Fig. 5. The bit error probabilities versus the received power under v=6643 and N=200.

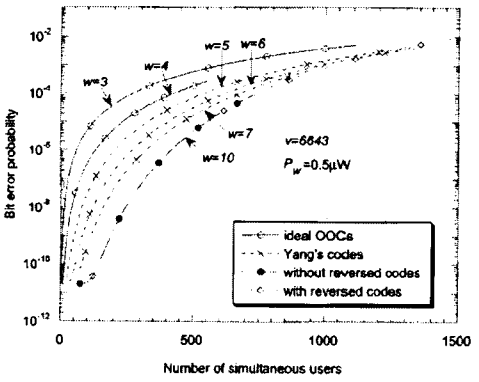


Fig. 6. The bit error probabilities versus the number of simultaneous users under v=6643 and $P_W = 0.5\mu$W.

We propose two new classes of OOCs. For the first time, we show that it is possible to construct a code family with code size larger than the code length. Compared with ideal OOCs and Yang's codes $(v, w, 1, 2)$ under the same code length and code weight, the performances of the systems with the four classes of codes are almost the same as one another. However, the proposed codes accommodate the users 10 times (which depends on the values of k and w) more than ideal OOCs for larger code length. To increase the number of simultaneous users for a given BER, we may increase the code weight and reduce the code size. We show that it is possible to accommodate many more simultaneous users with the proposed codes than ideal OOCs and Yang's codes.

C. Jitter Performance Analysis of SOCDMA-Based EPON Using Perfect Difference Code

For the spread spectrum technique of wireless communication system, PN sequence is usually used as a signature pattern. However, the SOCDMA code structure is designed under some correlation constrain and we can't further assume it at random. Therefore, we discuss the correlation property of PDC, which has better performance than other codes in the literature. We have the following observations. The codes with weight k and length $v = k^2 - k + 1$ have k-1 codewords with first chip of value "1" and k-1 codewords with last chip of value "1" but not both. These $2k$-2 codes are called margin_1 codes. There

exists only one "special code" with both first chip and last chip of value "1". The rest codes are named margin_0 codes as shown in Table. I. Note that the margin_1 codewords do not include the special code, which has chip "1" at both first and last chip position. We also observe that each codeword has two corresponding *burden codes*. For example, the desired receiver codeword "00000101000011001000" has corresponding burden codes "00000010100001100100" and "00001010000110010000" which are, respectively, the right shifted and left shifted versions of the desired code. We call it *burden code* because the cross-correlation between desired code and burden code is larger than the other codes. In the following analysis we will show that the burden code jitter has major impact on BER performance.

We can classify the receivers into three categories: the margin_1 receiver, the margin_0 receiver, and the special code receiver. A margin_1 receiver has the margin_1 code as the signature codeword. When time error occurs, the correlator output of the last_chip_1 receiver is influenced by the next transmitted data bit of the interfering users. On the other hand, the first_chip_1 receiver is influenced by the previous transmitted data bit. The two kinds of receivers have the same correlation properties. Without loss of generality, we consider only the last_chip_1 receiver in the following analysis and the analysis of the other kinds is similar.

We suppose that the signal waveform of APD output is rectangular. Moreover, the output waveform after entering an ideal low pass filter with bandwidth much larger than the signal bandwidth (e.g. filter bandwidth is five times of signal bandwidth) is very closed to rectangular. The thermal noise and shot noise in this receiver are much smaller than MUI and will not significantly influence the system performance. Therefore, we can assume

a rectangular signal waveform in our analysis without loosing the accuracy. We evaluate the jitter properties of three different receivers as the following.

A. The Receiver with Correlator

For the receiver using a correlator, if the desired signal has time error ε_d and the correlation error in each chip is Δ_d, the autocorrelation has the following property:

$$A=\begin{cases}(1-\Delta_d)\cdot k+y\cdot\Delta_d, \text{if specialcode}\\ (1-\Delta_d)\cdot k+\Delta_d, \text{otherwise}\end{cases} \quad (1)$$

where y is binomial distributed with equal probability that the desired signal has bit 1 or 0 in neighbor transmitted data bit.

The calculation of cross-correlation of PDC is more complicated than that of autocorrelation because the cross-correlations of codewords with first_chip_1 and last_chip_1 are different. Assume N_1 is the number of interfering users transmitting bit 1 in present bit duration and N_2 is the number of interfering users transmitting bit 1 in next bit duration, then N_1 and N_2 are statistically independent. The cross-correlation between the i-th interfering user and the desired last_chip_1 receiver in present data bit duration C_1 is given by:

$$C_1=\begin{cases}1 & ,\text{if } i \text{ hasfist\_chip0}\\ 1-\Delta_i x_i & ,\text{if } i \text{ hasfirst\_chip1}\\ 1+(k-2)\cdot\Delta_i x_i & ,\text{if } i \text{ isa burdencode}\end{cases}. \quad (2)$$

The cross-correlation of the last_chip_1 receiver with the i-th user in next transmitting bit is

$$C_{1n}=\begin{cases}0 & ,\text{if next bit transmits 0}\\ \Delta_i z_i & ,\text{if next bit transmits 1}\end{cases}, \quad (3)$$

where

$$x_i=\begin{cases}1 & ,\text{if late time error}\\ 0 & ,\text{if early time error}\end{cases} \quad (4)$$

$$z_i=\begin{cases}1 & ,\text{if next bit 1}\\ 0 & ,\text{if next bit 0}\end{cases}. \quad (5)$$

Assuming that bit "1", bit "0", late time and early time errors have the probability equal to $\frac{1}{2}$. Let n and m be the number of interference users with first_chip_1 transmitting bit 1, then the total cross-correlation is given by:

$$C=\begin{cases}\sum_{i=1}^{n}(1-\Delta_i x_i)+\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ \qquad ,\text{if no burden code appears} \\ \sum_{i=1}^{n}(1-\Delta_i x_i)+\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ +(k-2)\cdot\Delta_b x_b \\ \qquad ,\text{if one burden code apprears} \\ \sum_{i=1}^{n}(1-\Delta_i x_i)+\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ +(k-2)\cdot\Delta_{b'} x_{b'}+(k-2)\cdot\Delta_{b''} x_{b''} \\ \qquad ,\text{if two burden codes appear}\end{cases} \tag{6}$$

The special code receiver has the special code as its signature codeword. It has similar correlation property as (2) of the margin_1 receiver. The total cross-correlation of the special code receiver can also be derived as:

$$C=\begin{cases}\sum_{i=1}^{n}(1-\Delta_i x_i)+2\cdot\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ \qquad ,\text{if no burden code appears} \\ \sum_{i=1}^{n}(1-\Delta_i x_i)+2\cdot\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ +(k-2)\cdot\Delta_b x_b \\ \qquad ,\text{if one burden code appears} \\ \sum_{i=1}^{n}(1-\Delta_i x_i)+2\cdot\sum_{i=1}^{m}\Delta_i z_i+(\mathrm{N}_1-n) \\ +(k-2)\cdot\Delta_{b'} x_{b'}+(k-2)\cdot\Delta_{b''} x_{b''} \\ \qquad ,\text{if two burden codes appear}\end{cases} \tag{7}$$

A margin_0 receiver has the margin_0 code as the signature codeword. The cross-correlation property is given by:

$$C_0=\begin{cases}1+(k-1)\cdot\Delta_i x_i & ,\text{if } i \text{ is a burden code} \\ 1 & ,\text{otherwise}\end{cases} \tag{8}$$

The total cross-correlation of margin_0 receiver can be derived as:

$$C=\begin{cases}\mathrm{N}_1 \\ \qquad ,\text{if no burden code appears} \\ \mathrm{N}_1+(k-1)\cdot\Delta_b x_b \\ \qquad ,\text{if one burden code appears} \\ \mathrm{N}_1+(k-1)\cdot\Delta_{b'} x_{b'}+(k-1)\cdot\Delta_{b''} x_{b''} \\ \qquad ,\text{if two burden codes appear}\end{cases} \tag{9}$$

B. The Receiver using Simple Sampler

Here we analyze a receiver using simple sampler to reduce the jitter impact when the maximum jitter is less than one half of the chip duration. The receiver structure is shown in Fig. 4. The sampling point is at the center of the chip and the sampling rate is equal to the chip rate. In this receiver structure, the output is correct as long as the sampling instant falls within the corresponding chip duration. When the tracking loop is used in the receiver, the autocorrelation A will always be correct and $A=k$.

For the margin_1 receiver, we may follow the previous procedure and obtain the last_chip_1 receiver correlation value. The results of the first_chip_1 receiver and the special code receiver can be derived in the same way. The cross-correlation of the margin_1 receiver is given by:

$$C_1 = \begin{cases} 1 \\ \quad ,\text{if } i \text{ has first\_chip 0} \\ 1-\alpha_i\beta_i \\ \quad ,\text{if } i \text{ has first\_chip 1} \\ 1+(k-1-\beta_b)\alpha_b \\ \quad ,\text{if } i \text{ is a burden code with first\_chip 1} \\ 1+(k-1)\alpha_b \\ \quad ,\text{if } i \text{ is a burden code with first\_chip 0} \end{cases} \tag{10}$$

where α_b is the probability that the burden code has time error exceeding 0.5 chip and is binomial distributed with probability given by:

$$p(\alpha_b = 1) = \int_{-\varepsilon_{\max}}^{+\varepsilon_{\max}} \int_{0.5+\varepsilon_b}^{\infty} f(\varepsilon_t, \varepsilon_b) d\varepsilon_t d\varepsilon_b \quad , \tag{11}$$

where ε_t is the jitter of tracking loop and ε_b is the jitter of the burden code. Assuming ε_t and ε_b are statistically independent, then their joint probability density function is

$$f(\varepsilon_t, \varepsilon_b) = f(\varepsilon_t) \cdot f(\varepsilon_b) = \frac{1}{\sqrt{2\pi\sigma_{\varepsilon_t}^2}} e^{-\frac{\varepsilon_\varepsilon^2}{2\sigma_{\varepsilon_t}^2}} \cdot \frac{1}{2\varepsilon_{\max}} \quad . \tag{12}$$

and β is binomial distributed with equal probability

$$\beta = \begin{cases} 1 & ,\text{if next bit is 0} \\ 0 & ,\text{if next bit is 1} \end{cases} \quad . \tag{13}$$

The total cross-correlation is given as:

$$C = \begin{cases} \mathrm{N}_1 - n + \sum_{i=1}^{n}(1-\alpha_i\beta_i) \\ \quad ,\text{if no burden code appears} \\ \mathrm{N}_1 - n + \left(\sum_{i=1}^{n}(1-\alpha_i\beta_i)\right) + (k-1-\beta_b)\alpha_b \\ \quad ,\text{if only one burden code with first\_chip 1 appears} \\ \mathrm{N}_1 - n + \left(\sum_{i=1}^{n}(1-\alpha_i\beta_i)\right) + (k-1)\alpha_b \\ \quad ,\text{if only one burden code with first\_chip 0 appears} \\ \mathrm{N}_1 - n + \left(\sum_{i=1}^{n}(1-\alpha_i\beta_i)\right) + (k-1-\beta_{b'})\alpha_{b'} + (k-1)\alpha_{b''} \\ \quad ,\text{if two burden codes appear} \end{cases} \tag{14}$$

The cross-correlation of the margin_0 receiver is given by:

$$C_0 = \begin{cases} 1+(k-1)\alpha_b \\ \quad ,\text{if jitter} > 0.5\,T_c \text{ and } i \text{ is burden code} \\ 1 \\ \quad ,\text{otherwise} \end{cases} \tag{15}$$

The total cross-correlation of the margin_0 receiver is

$$C=\begin{cases} N_1 \\ \quad \text{, if no burdne code appears} \\ N_1+(k-1)\cdot\alpha_b \\ \quad \text{, if one burden code appears} \\ N_1+(k-1)\cdot\alpha_{b'}+(k-1)\cdot\alpha_{b''} \\ \quad \text{, if two burden codes appear} \end{cases} \tag{16}$$

C. The Receiver using Joint Jitter Estimator and Canceller

We have defined two burden codes, which are right shifted and left shifted version of the desired code. Burden codes will cause additional multiple-user interference (MUI) as (2), (8) in the receiver with correlator and as (10), (15) in the receiver using simple sampler. Note that a left shifted burden code has additional MUI effect when it has early time error but no effect when late time error occurs. The case is similar for the right shift burden code. In order to improve the system performance, such additional MUI due to burden code jitter must be cancelled. We propose a joint jitter estimator and canceller (JJEC) as shown in Fig. 5. The upper part is the data detection circuit with MUI canceller and the lower part is noncoherent delay-locked tracking loops. Two DLLs are employed in the lower part of the receiver, which act as two burden code jitter estimators. Let DLL1 track the time information of one of the burden codes $T_{b'}$, which is the right shift version of the receiver signature pattern. DLL2 tracks the time information of the other one $T_{b''}$. In the correlator receiver, JJEC estimates the time error and calculates the amount of jitter. Assuming that the burden code jitter is $\bar{\varepsilon}_\ell = T_b - T_\ell$, ($\ell = b', b''$) and correlation error is $\bar{\Delta}_\ell = \left|\bar{\varepsilon}_\ell\right|$. A decision circuit with zero threshold value is used to detect whether the burden code has the early or late time error, which may cause additional MUI. When additional MUI is induced, then the output of JJEC canceller is $\phi = \bar{\Delta}_\ell \cdot \xi$, where

$$\xi = \begin{cases} GT_c \dfrac{(k-1)}{k}\lambda_1 & \text{, if margin\_0 receiver} \\ GT_c \dfrac{(k-2)}{k}\lambda_1 & \text{, if margin\_1 receiver} \end{cases} \tag{17}$$

where G is the APD gain. In the simple sampling receiver, the average cross-correlation of a burden code to the desired receiver is k for the margin _0 receiver or $(k-\frac{1}{2})$ for the margin_1 receiver when the jitter exceeds $\pm 0.5T_c$. In order to detect and cancel the large cross-correlation of the burden code, the decision threshold of $\bar{\varepsilon}_{b'}$ is set to $-0.5T_c$ and $\bar{\varepsilon}_{b''}$ is set to $0.5T_c$. If $\bar{\varepsilon}_{b'} < 0.5\ T_c$ or $\bar{\varepsilon}_{b''} > 0.5\ T_c$, additional MUI will be induced. The amount of cancellation is set $\phi = GT_c \dfrac{(k-1)}{k}\lambda_1$ for the margin _0 receiver and $\phi = GT_c \dfrac{(k-\frac{1}{2}-1)}{k}\lambda_1$ for the margin _1 receiver.

The BER of the receiver using correlator under various signal power and maximum peak jitter Δ_{max} when the system is full-loaded and code weight k=12 is shown in Fig.8. We observe that BER degrades with increasing Δ_{max} and elevates as signal power increases. Figs. 9-10 compare the BER performance of the receiver using correlator and the receiver using simple sampler for various received signal power assuming ε_{max} =0.5 T_c, 0.4 T_c and $\sigma^2_{\varepsilon_t}$ =0.0025 T_c, respectively. Numerical results show that the performance of the correlator receiver is power limited when no JJEC is used. This is because the MUI of burden codes dominates the performance degradation. If we increase the signal power, it also enhances MUI. The receiver with JJEC can significantly improve the BER and the performance depends on the accuracy of NC/DLL. The smaller the tracking error of NC/DLL is, the better performance JJEC can be achieved. These figures also show that the receiver using simple sampler has better BER performance

than the correlator receiver. The use of JJEC in the simple sampling receiver has good BER performance closed to the ideal case (no jitter) under the condition of small time error ε and NC/DLL performing well. This is because the sampling point falls within the chip boundary and the effect of MUI caused by burden code can be accurately detected and cancelled. Figs. 11-12 show the BER performance as functions of number of simultaneous users. The BER degrades when the number of simultaneous users increases. In addition, the receiver without JJEC has poor BER performance even in low user population. JJEC can indeed make the system robust against jitter.

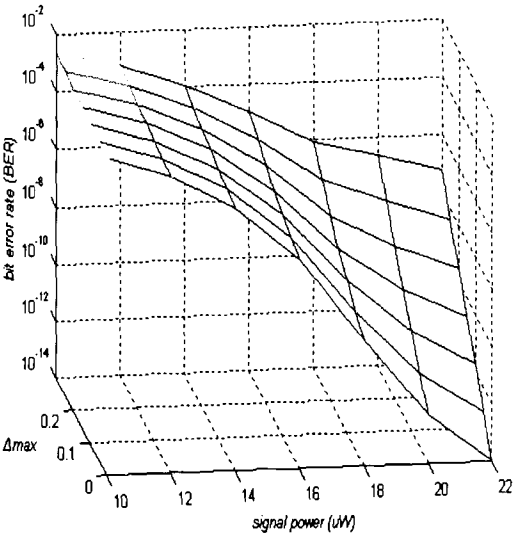


Fig. 8. Bit error rate (BER) of correlator receiver with various signal power and maximum peak-to-peak jitter $\varepsilon_{\max}$ when k=12.

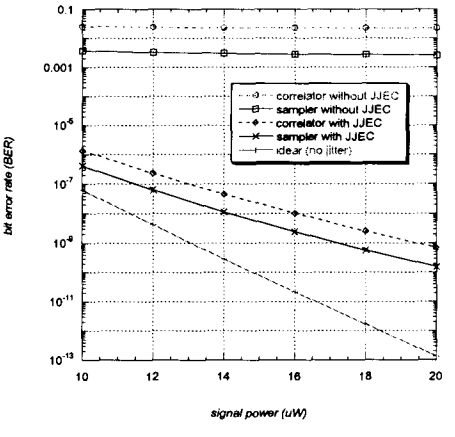


Fig. 9. Bit error rate (BER) with various signal power of correlator receiver and peak sampling receiver when k=12, $\varepsilon_{\max}$ =0.5 T_c and tracking error =0.0025 T_c.

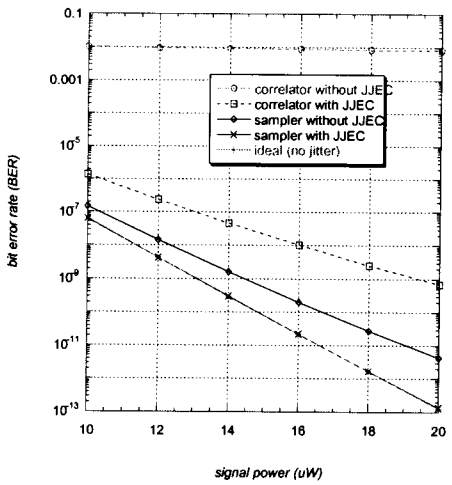


Fig. 10. Bit error rate (BER) with various signal power of correlator receiver and peak sampling receiver when k=12, $\varepsilon_{\max}$=0.4T_c and tracking error =0.0025T_c.

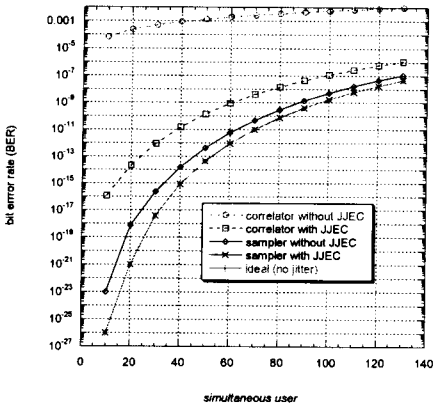


Fig. 12. Bit error rate (BER) with various simultaneous user of correlator receiver and peak sampling receiver when k=12, $\varepsilon_{\max}$=0.4T_c, signal power =10μW and tracking error =0.0025T_c.

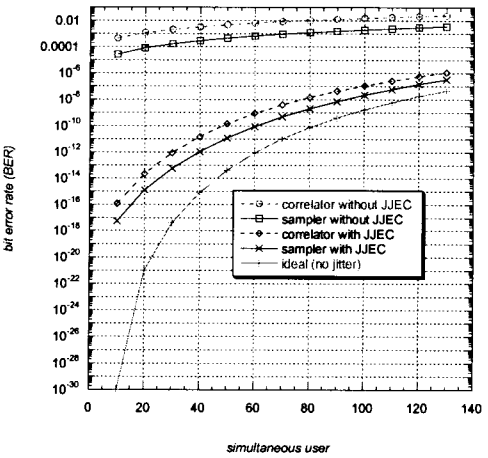


Fig. 11. Bit error rate (BER) with various simultaneous user of correlator receiver and peak sampling receiver when k=12, $\varepsilon_{\max}$=0.5T_c, signal power =10μW and tracking error =0.0025T_c.

class		code sequence
Margin_0 codeword		01000000001010000110
		00000001010000110010
		00000010100001100100
		00000101000011001000
		00001010000110010000
		00010100001100100000
		00101000011001000000
		00101000110010000000
		01010000110010000000
		00011001000000001010
		00110010000000010100
		01100100000000101000
Margin_1 codeword	First chip_1 code	11001000000001010000
		10000000010100001100
		10100001100100000000
		10000110010000000010
	Last chip_1 code	00100000000101000011
		00000000101000011001
		01000011001000000001
		00001100100000000101
Special code		
10010000000010100001		

四、 計劃成果自評與結論

【第一年】

我們提出一個新的 IP over WDM 架構，利用標籤交換(label switching)與 WDM 波長資訊的觀念，設計一個可以有效簡化網路規約的封包格式，同時也將 MPLS 流量控制概念與 IP 傳輸與交換平面做有效整合。另一方面我們也分析了兩種可以處理非同步變動長度封包的高速光網路路由器：部份分享暫存器(PSB)與多重路徑路由(MPR)，模擬結果顯示 PSB WDM 光路由器比較適合使用在高速 IP 光纖網路。

【第二年】

本研究計畫設計一高速可非同步交換長度可變之封包的交換器。由於模組化的設計以及分散式的自我路由決定封包的交換排程，此新架構可以以簡單易行的硬體實現之。另外，我們經由此系統的理論分析與模擬，證明其可達到最佳化的效能，即百分之百的最高流通率以及低延遲時間。本交換器可非同步地處理長度可變封包之多點發送交換，並且我們亦深入地分析可變長度封包的多點發送之效能，實爲此領域之創新。本研究計畫發展之技術已申請美國和台灣專利，同時其成果亦發表於國際會議論文。

【第三年】

本研究計劃設計出兩組光分碼多工之編碼:完美差異碼與非理想交互關係碼。經過數值分析結果證明出這兩組編碼方式比傳統的同步碼與非同步碼有更佳的性能，此外、我們也提出相互時間偏差估測與消除電路可以有效改善完美相差碼對於時間誤差的敏感性；將此兩組編碼應用被動光網路可以真正達到寬頻接取的服務。

五、 參考文獻

【第一年】

[1]EURESCOM Project P918-GI, "Integration of IP over Optical Networks: Networking and Management," Deiverable 1, Oct. 1999.

[2]R. Callon, P. Doolan, N. Feldman,

A. Fredette, G. Swallow and A. Viswanathan, "A Framework for Multiprotocol Label Switching," Internet Draft, Work in Progress, Sep. 1999.

[3] E. C. Rosen, A. Viswanathan and R. Callon, "Multiprotocol Label Switching Architecture," Internet Draft, Work in Progress, Aug. 1999.

[4]B. T. Doshi, S. M. Dravida, E. J. Hernandez-Valencia, W. A. Matragi, M. A. Qureshi, J. Anderson and J. S. Manchester, "A Simple Data Link Protocol for high-Speed Packet Networks", *Bell Labs Technical J.*, pp85-104, Jan.-Mar. 1999.

[5]Bijan Raahemi, "Comparison of Frame Delineation Performance Between 10GE WAN PHY and 8B/10B," Nortel Networks' submission to IEEE802.3 HSSG, ver. 1.0, Nov. 1999.

[6]R. D. Doverspike, S. Phillips, and J. R. Westbrook, "Future Transport Network Architecture," *IEEE Commun. Magazine*, vol. 37, issue 8, pp.96-101, Aug. 1999.

[7]D. O. Awduche, Y. Rekhter, J. Drake, R. Coltun, "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects," *Internet Draft*, Work in Progress, Nov. 1999.

[8]Tom Worst, "General Switch Management Protocol," *Internet Draft*, Work in Progress, Oct. 1999

[9]John Crowcroft, "IP over Photons: How not to Waste the Waist of the Hourglass," *IEEE Conference IWQoS '99, Seventh International Workshop on Quality of Service,* pp.9-11, April 1999.

[10] ITU-T Recommendation I.432, "B-ISDN user-network interface – Physical layer specification: General characteristics", Aug. 1996.

[11] DooWhan Choi, "Frame Alignment in a Digital Carrier System – A Tutorial," *IEEE Commun. Magazine*, vol. 28, issue 2, pp.47-54, Feb. 1990.

[12] R. W. Sitter, "Systems Analysis of Discrete Markov Processes," *IRE Trans. on Circuit Theory*, Vol.CT-3, pp.257-266, Dec 1956.

[13]Franco Callegati, Giorgio Corazza and Carla Raffaelli, "Exploitation of DWDM for optical packet switching with quality of service guarantees," *IEEE J. Select. Areas Commun.*, vol. 20, pp 190-201, Jan. 2002.

[14] Lubo Tancevski, Siva Yegnanarayanan, Gerardo Castanon, Lakshman Tamil, Francesco Masetti and Tom McDermott, "Optical Routing of Asynchronous, Variable length packets," *IEEE J. Select. Areas Commun.*, vol. 18, pp 2084-2093, Oct. 2000.

[15] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*,

vol. 2, no.1, pp. 1-15, Feb. 1994.

[16] M. E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," *IEEE/ACM Trans. Networking*, vol. 5, no.6, pp. 835-846, Dec. 1997.

[17] Meow C. Chia, David K. Hunter, Ivan Andonovic, P. Ball and I. Wright, "Optical packet switches: A comparison of designs," in *Proc. ICON'2000*, Singapore, Sept. 2000, pp. 365-369.

[18] David K. Hunter, Meow C. Chia and Ivan Andonovic, "Buffering in optical packet switches," *IEEE/OSA J. Lightwave Technol.*, vol 16, no. 12, pp. 2081-2094, Dec. 1998.

[19] David K. Hunter, Mohamed H. M., Meow C. Chia, Ivan Andonovic, Ken M. Guild, Anna Tzanakaki and Mike J. O'Mahony, John D. Bainbridge, Marc F. C. Stephens, Richard V. Penty and Ian H. White, "WAPNET: A Wavelength Switched Packet Network," *IEEE Commun. Mag.*, vol. 19, issue 9, pp. 120-129, March 1999

[20] Wen De Zhong and Rodney, S. Tucker, "A Wavelength-routed photonic packet buffer combining travelling delay lines with delay loops," *IEEE/OSA J. Lightwave Technol.*, vol. 19, no. 8, pp. 1085-1092, Aug. 2001.

[21] P. Gambini *et al.*, "Transparent optical packet switching: Network architecture and demonstration in the KEOPS project," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 45-59, Sept. 1998.

[22] Soeren L. Danielsen, Peter B. Hansen and Kristian E. Stubkjaer, "Wavelength conversion in optical packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 11, no. 12, pp. 2095-2108, Dec. 1998.

[23] Gerardo Castanon, Lubo Tancevski and Lakshman Tamil, "Optical packet switching with multiple path routing," *Computer Networks.*, vol. 32, issue 5, pp. 635-662, May 2000.

[24] Juan Diao and Pak. L. Chu, "Analysis of partially shared buffering for WDM optical packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 17, no. 12, pp. 2461-22469, Dec. 1999.

[25] Lubo Tancevski, An Ge and Gerardo Castnon, "Optical packet switchi with partially shared buffers: Design Principles," in *Proc. OFC '2001*, Anaheim, March 2001, paper TuK3.

[26] Franco Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, pp. 292-294, Sept. 2000.

【第二年】

[1] Kun-Tso Chen and Jingshown Wu, "An asynchronous IP switch for variable length packets," R.O.C. patent pending, U.S.A. patent pending.

[2] N. Chandhok, A. Durresi, R. Jagannathan, S. Seetharaman, K. Vinodkrishbnan, "IP over Optical Networks: A summary of issues," Internet Draft draft-osu-ipo-mpls-issues-00.txt, Work in Progress, July 2000.

[3] Lisong Xu, Harry Perros and George Rouskas, "A survey of optical packet switchin and optical burst switching," Technical Report, Department of Computer Science, North Carolina State University, 2000.

[4] L. Ceuppens, et al., "Performance monitoring in photonic networks in support of MP λ S," Internet Draft draft-ceuppens-mpls-optical-00.txt, March 2000.

[5] M.G. Hluchyj, and M. J. Karol, "Queueing in High-Performance Packet Switching," *IEEE J. Selected Areas Commun.*, Vol. 6, No. 9, pp. 1587-1597, Dec. 1988.

[6] T. E. Anderson, S. S. Owicki, J. B. Saxe, and C. P. Thacker, "High-Speed Switch Scheduling for Local-Area Networks," *ACM Trans. Computer Systems,* Vol. 11, No. 4, pp. 319-352, Nov. 1993.

[7] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% Throughput in an Input-Queued Switch," *IEEE Trans. Commun.*, Vol. 47, No. 8, pp. 1260-1267, Aug. 1999.

[8] J. J. Hunter, *Mathematical Techniques of Applied Probability. Volume 2. Discrete Time Models: Techniques and Applications.* New York: Academic Press, 1983.

[9] M. S. Schmookler, "Limited capacity discrete time queues with single or bulk arrival," IBM technical report TR 00.2048, New York, June 18, 1970.

[10] J. D. C. Little, "A simple proof of $L = \lambda W$," Operations Research, Vol. 9, pp. 383-387, 1961.

[11] Hideaki Takagi, *Queueing Analysis, A foundation of Performance Evaluation, Volume 3: Discrete-Time Systems*. Elsevier Science Publishers B.V., Amsterdam, 1993.

[12] S. Halfin, "Batch delays versus customer delays," the Bell System Technical Journal, Vol.62, No. 7. pp.2011-2015, Sep. 1983.

[13] Guo and Ruay-Shiung Chang, "Multicast ATM switches: survey and performance evaluation," Comput. Commun. Rev, vol. 28, pp.98-131, Apr. 1998.

[14] Joesph Y. Hui and Thomas Renner, "Queueing analysis for multicast packet switching," IEEE Trans. Commun., vol. 42, pp.723-731, Feb. 1994.

【第三年】

[1] Chi-Shun Weng and Jingshown Wu, "Perfect difference codes for synchronous fiber-optic CDMA communication systems," *J. Lightwave Technol.*, vol. 19, no. 2, pp. 186-194, Feb. 2001.

[2] Chi-Shun Weng and Jingshown Wu, "Optical orthogonal codes with nonideal cross correlation," *J. Lightwave Technol.*, vol. 19, no. 2, pp. 1856-1863, Dec. 2001.

[3] Jingshown Wu and Fong-Ray Gu, "Jitter performance analysis of SOCDMA-based EPON using perfect difference code," submitted to *J. Lightwave Technol.*

[4] Chi-Shun Weng and Jingshown Wu, "Optical orthogonal codes with large cross-correlation and their performance bound for asynchronous optical CDMA systems," accepted by *J. Lightwave Technol.*

[5] Che-Li Lin and Jingshown Wu," Channel interference reduction using random Manchester codes for both synchronous and asynchronous fiber-optic CDMA systems," *J. Lightwave Technol.* Vol. 18, no. 1, pp.26-33. Jan. 2000.

[6] Jingshown Wu and Che-Li Lin," Fiber-optic code division add-drop multiplexers," *J. Lightwave Technol.* Vol. 18, no. 6, pp.819-824. June 2000.

[7] Che-Li Lin and Jingshown Wu," A synchronous fiber-optic CDMA system using adaptive optical hardlimiter," *J. Lightwave Technol.*, Vol. 16, no. 8, pp. 1393-1403, Aug. 1998.

[8] Jyh-Horng Wu and Jingshown Wu, "Synchronous fiber-optic CDMA using hard-limiter and BCH codes," *J. Lightwave Technol.*, vol. 13, no. 6, pp. 1169 –1176, June 1995.

[9] W. C. Kwong, P. A. Perrier, and P. R. Prucnal, "Performance comparisons of asynchronous and synchronous code-division multiple-access techniques for fiber-optic local area network," *IEEE Trans. Commun.* vol. 39, no. 11, pp. 1625-1634, Nov. 1991.

[10] J. G. Zhang, "Improvement in bandwidth efficiency of asynchronous prime-code CDMA networks by error correction coding," *IEEE Electronics Letters*, vol. 30, no. 6 , pp. 514-515, March 1994.

[11] J. A. Salehi, "Code division multiple-access techniques in optical fiber networks- Part I: Fundamental principles," *IEEE Trans. Commun.* vol. 37, no. 8, pp. 824-833, Aug. 1989.

[12] W. C. Kwong, G. C. Yang, and J. G. Zhang, " 2^n prime-sequence codes and coding architecture for optical code-division multiple-access," *IEEE Trans.*

Commun. vol. 44, no. 9, pp. 1152-1162, Sept. 1996.

[13] H. M. H. Shalaby, "Synchronous fiber-optic CDMA systems with interference estimators," *J. Lightwave Technol.*, vol. 12, no. 11, pp. 2268-2275, Nov. 1999.

[14] W. C. Kwong and P. R. Prucnal, "All-serial coding architecture for ultrafast optical code-division multiple-access," in *Proc. ICC '93*, vol. 1, May 1993, pp. 552–556.

[15] J. Singer, "A theorem in finite projective geometry and some applications to number theory," *Trans. Amer. Math. Soc.*, vol. 43, pp. 377–385, 1938.

[16] J. Schönheim, "On maximal systems of k-tuples," *Studia Sci. Math Hungar.*, vol. 1, pp. 363–368, 1966.