A Two-level TDM Access Protocol for IP Support in WDM Optical Networks

Chi-Yuan Chang, Sy-Yen Kuo

Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
sykuo@cc.ee.ntu.edu.tw, chiyuan@lion.ee.ntu.edu.tw

Abstract

This paper presents a new access protocol for constructing IP over photonic systems. The protocol is based on a two-level TDMA structure with wavelength division multiplexing (WDM). Many IP-based network applications such as distributed VPN services, high-resolution image, distributed database, and real-time video/audio service, generally require high-speed transmissions and usually exhibit traffic locality in WAN/LAN. Consequently, based on the traffic parameters such as locality and loading, an architecture named a 2TDMA (2-level TDMA) network is proposed. An analytical model is also presented and evaluated. The results of the analytical evaluations show that 1) the performance of a network depends on the network loading, locality, and number of groups. 2) The increased traffic locality could increase the slot utilization and throughput. 3) If the traffic locality is larger than a threshold, a 2TDMA network can generate higher throughput than a network with TDMA. Hence, if a system can be operated with appropriate loading, group, and traffic locality, its utilization and throughput will be enhanced significantly.

Keywords: *Wavelength division multiplexing (WDM), Time Division Multiplexing Access (TDMA), Optical networks, IP over WDM*

1. Introduction

The diffusion of Internet traffic, IP-based networks, and applications is growing at an exponential rate in both private and public networks while IP is becoming the dominant protocol for information communication technology. Thus, IP over WDM has become a very important area of study. Recent developments in multiprotocol label switching (MPLS) open new possibilities to address some of the limitations of the traditional IP systems. MPLS switches use a simple label-swapping algorithm replacing the standard destination-based hop-by-hop forwarding paradigm to quickly forward packets [1], enabling easy scaling to terabit rates. The MPLS technique uses a label to save significant processing time by avoiding network layer label analysis at each hop. In addition, the MPLS can provide many of the same advantages of a connection-oriented network while still retaining the underlying efficiency and operation of a datagram network.

Current applications of WDM as a networking technique focus on relatively static utilization of individual wavelength channels. These wavelength routing approaches are still basically variants of the circuit-switching paradigm, resulting in very inefficient optical bandwidth usage. A technological breakthrough in this direction is represented by optical packet switching [2-4], enabling fast allocation of the WDM channels and their utilization as shared resources in an on-demand fashion with very fine granularities. The drawbacks of this approach mainly consist of the difficulties of implementing the optical synchronizer and of processing the packet headers in the electronic domain [5]. Recently, variable-length optical packet switching, such as optical burst switching (OBS) [6,7], has been proposed as an optical switching paradigm to combine the best of optical circuit and packet switching. All these researches focus on designing a technique to improve the parameters in terms of transmission scheduling, synchronization issues, contention resolution, and switching strategies, but never takes the traffic characteristic into consideration.

The (80%, 20%) rule, 80% income comes from 20% customer, can be found everywhere. Of course it is also adopted and named as traffic locality in the network environment [8-11]. Claffy and Polyzos [8] depicted the cumulative distribution of messages sent from and to the n busiest source and destination networks within the NSFNET. Over 50% of the traffic is generated by the busiest 31 of the 4254 site networks (0.7%), and over 50% travels to the 118 most popular (2.8%) destinations. And 46.9% of the total traffic on the backbone travels between 1500 (0.28%) of the 560,049 site-pairs. Quantitative measures of locality in communication over a local area network have been described in [10]. Traffic flows satisfying various temporal and spatial locality conditions are observed at internal points of the network in [11]. These researches show the traffic locality existence

Many network applications such as distributed VPN services, multimedia, video conferencing, etc., usually exhibit traffic locality (which means most traffic between the transmitter node and the receiver node is located at some specific areas). This type of traffic uses relatively high bandwidth on a continuous basis for a long period of time. Besides, for the traffic with higher priority, the network can also route the traffic as soon as possible by sacrificing the lower priority traffic. Hence, we propose and analyze a new routing architecture for Internet protocol (IP) operating on a WDM backbone network. We also investigate the relationships between the traffic characteristic and performance. The proposed routing architecture, named as 2TDMA (2-level TDMA), is based on the TDM multiplexing approach, with WDM addressing the number of regional exchanges (REXs) and time-division switches communicating among the backbone-hubs which is essentially an optical crossconnect featuring a classical time-division space switch. The 2TDMA consists of r different groups of nodes/REXs. All node/REXs in each group exhibit traffic locality and have the same control channel. The control channel, named as the interleaved control slot (ICS), is arranged by overlapping half cycle. By partitioning/reconfiguring a network as much as possible and using ICS as control protocol, the slot reuse can be easily achieved and the slot utilization of the network can be improved with the high traffic locality.

This paper is organized as follows. In section 2 we discuss the architecture and functional blocks for IP-based WDM networks. Section 3 describes the 2-level TDMA architecture with the interleaved control slot concept. Section 4 depicts the analytical model and results. In section 5 we conclude the paper.

2. Network Architecture

2.1 System level architecture

Recently, there has been an increasing interest in the implementation of IP over photonic networks by using optical networking techniques. Fig. 1 illustrates the general network architecture. The architecture consists of 1) an Internet access part (IAP), 2) a regional exchange (REX), and 3) an optical-hub switching backbone (OSB). The IAP, as shown in Fig. 1, consists of IP hosts connected to corporate with servers via a local-area network (LAN) for offices or connected to an Internet services provider (ISP) via phone-line modems, digital subscriber line (xDSL), cable modems, or wireless access. The regional exchange (REX) and optical-hub switching backbone (OSB) form the optical switching system for the IP network.

The network can be divided into two independent levels, the switching-hub level and the REX level. The former performs IP packet exchange in the optical domain within hub, while the latter performs the same task in the electrical domain. The switching-hub level and the REX level are described below respectively.

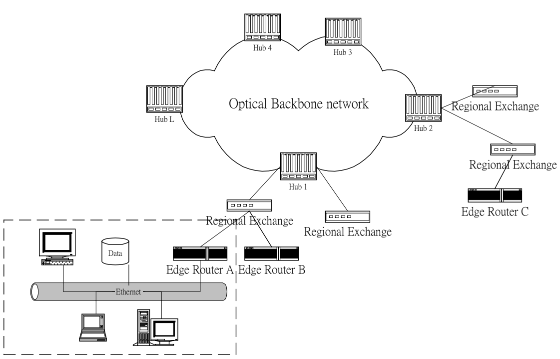


Figure 1. The general network architecture.

2.1.1. Switching-hub level: The connections between hubs can be a bus/dual bus, a ring/dual ring, a mesh (fully connected network), or other interconnection networks. In order to explain easily, we choose the fully connected network for its simplicity and its robustness. Specifically, N (the number of hubs) switching states can easily be generated, and there are $N-1$ alternate paths to use when the direct path is down. For non-fully connected interconnection between hubs, complicated switching patterns need to be designed so as to emulate a fully connected network with end-to-end path for all connections between REXs within the network. Let us begin by examining the proposed architecture at the hub level. The hub shown in Fig.1 is for the exchange of packets between regions within the same hub or across distributed hubs. For example, if a source transmits data to the destination in the same region, the data will not be transmitted to other hub region (ex: edge router A=>edge router B in Fig.1). On the contrast, the data will be transferred to distributed hubs of remote region if the source and destination pair is in different regions (ex: edge router A =>edge router C in Fig.1). The architecture of the hub switching network shown in Fig.2 is implemented by a time-division optical switch, as well as the switching states are cycled through in each period. The resulting transmission orders for the hubs are shown in the hub level of Fig. 3. In each time slot of transmission cycle, the hub transfers its packets to a predefined hub. For example, in the time slot 2 of transmission cycle, hub 1 to hub2, hub 2 to hub3, …, and hub L to hub1, etc.

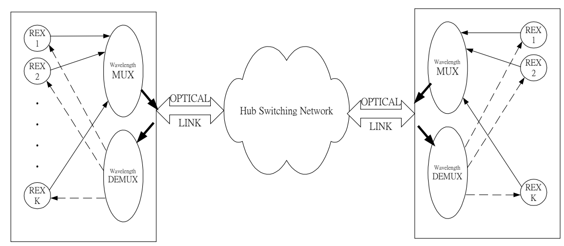


Figure 2. Network with K connected REXs.

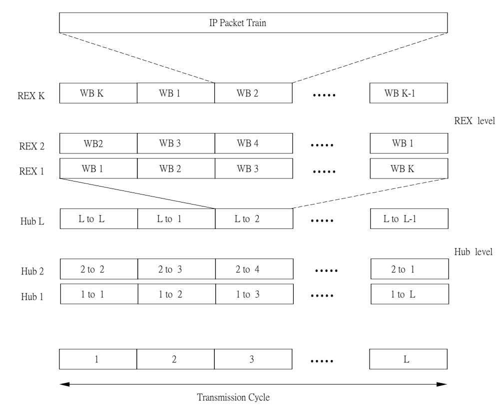


Figure 3. The transmission cycle.

2.1.2. REX level: At the REX level, let us first focus on intra-REX communication. The exchange of IP packets between edge routers in the same region is performed in the electrical domain with the local REX switch playing the role of a "switching" router. As we are concerned with the transport of IP packets, the label can be included as part of the header of Layer 3 (i.e., by using the Flow Label field in the IPv6 packet with appropriately modified semantics [12]). This label indexing is done at the edge routers to ease the processing load of the REX switches. Since a label is bound to an IP address prefix, IP packets heading to a group of destinations (i.e., to edge routers connected at a specific region on a specific hub) can share the same label. For IP packet routing within a region, the destination router addresses on the labels are resolved at the local REX switch. IP packets destined for edge routers at a remote region are assigned to different transmission queues according to their respective labels. The transmissions from each queue onto the optical backbone network are then organized into transmission cycles, with each cycle subdivided into wavelength burst (WB) periods using WDM. The WB periods consist of a series of wavelengths to different destinations. To illustrate, let us focus at Hub L in period 3, as shown in the REX Level of Fig. 3. Here, Hub L is connected to Hub 2. During this period, REX 1 of Hub L transmits WB $\lambda_1, \lambda_2, \cdots, \lambda_k$ to REX 1, 2, …, REX K in Hub 2 (assuming there are K REXs in both Hub L and Hub 2). In the same period, REX 2 of Hub L transmits WB $\lambda_2, \cdots, \lambda_k, \lambda_1$ to REX 2, 3,.., REX 1 in Hub 2. Other REXs use cyclic permutations in the wavelengths so that no two wavelengths are used at the same time. These transmissions are merged at a coupler in the hub before being sent out to another hub. Note also that the flow of intra-REX traffic is decoupled from the flow of inter-REX traffic. Fig. 3 shows the transmission cycle at different levels of switching hub level, REX level and IP packets level. Here we assume that a network has L hubs and each hub contains several REXs. There are K REXs in a network.

2.2 Functional block of edge router and REX switch

Fig. 4 shows a functional diagram of the edge router and the REX switch. A station is equipped with an edge router and its local REX switch. In the edge router, the Input Dispatcher sorts IP packets from connected REXs. Those destined for REXs attached to the local router are buffered and sent to the Output Dispatcher. Those destined for remote routers are sent to the label-indexing buffer with the MPLS technology [1]. IP packets in the label-indexing buffer are labeled according to their destination-router addresses and placed onto the output buffer. These are then sent to the input buffer of the local REX switch where IP packets from individual edge routers are switched to output queues, according to their respective labels, to form IP packet trains for inter-REX routing. In other words, IP packets from the same train are all destined to the same remote region. The destination addresses of those IP packets labeled for the local region will be resolved, and the packets are rerouted to their respective destination edge routers.

On the input side of the optical links, optical signals carrying labeled IP packet trains from the hub are converted to electrical signals before breaking up into individual IP packets. The REX switch then resolves the IP destination addresses of these packets and passes them to their respective destination edge routers. For example, as shown in Fig. 1, a label is used to identify a traffic flow from local edge router A=> local REX switch => local edge router B or between local edge router A => local REX switch => hub 1 => hub 2 =>remote REX switch => remote edge router C. The network inspector shown in Fig. 4 monitors all the packets of the output queues in the REX switch. It monitors all the transmission conditions from the edge router and inspects the traffic from the remote REX switch. The network inspector can get information that can help the switch to understand the traffic bandwidth requirements in every REX switch. In this way, we can know the load distribution along the optical backbone and decide on the switching reconfiguration in the network. We will discuss the details next section.

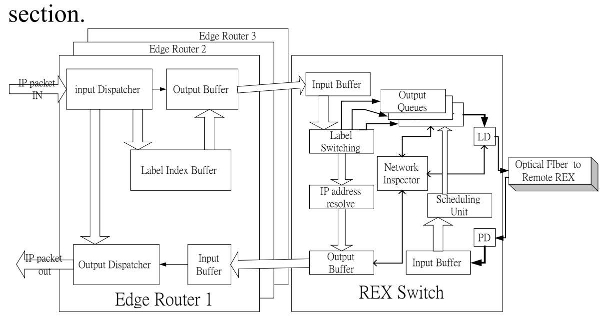


Figure 4. The Functional block of edge router and REX switch.

3. 2-Level TDMA Architecture

Before discussing 2-level TDMA structure, we first review the flows of basic TDM access protocol. The traditional approach uses a wavelength as the control channel to schedule a node to transmit data [13]. Hence, each node has one fixed receiver for receiving the TDM-based control channel. A pseudo cycle with a length of *N* control slots is generated on the control channel, and only one control slot is allocated for each node in a cycle. On the other hand, a tunable transmitter is used for transmitting data to destination and a fixed/tunable receiver is used for receiving data from source. When a node *i* wants to transmit data to node *j*, there are several setup steps before transmitting. First, the tunable transmitter *Tx* needs to tune its wavelength to λ_j and then wait for transmitting data until the *ith* control slot is accessed by the control channel. Node *i* can only transmit in the interval of the *ith* control slot. If node *i* needs to transmit data again, it must wait until the *ith* control slot of next control cycle. It is clear that each node must wait *n* control slots to transmit data again. If node *i* has a great deal of data to transmit to node *j*, it results in inefficiency.

3.1 2- Level TDMA Transmission

The transmission cycle shown in Fig. 3 lists the benefits of combining WDM and TDMA. The hub *L* can transmit to other *L-1* hubs simultaneously by different wavelengths of independent optical paths on the transmission slot *L*. Each REX switch transmits the corresponding Wavelength burst (WB) in its slot. The length of transmission trains for hub L in the *ith* control slot is *K*. Hence the length for a transmission cycle is *KL*.

Proceedings of the 2003 International Conference on Parallel Processing Workshops (ICPPW'03)
1530-2016/03 $ 17.00 © 2003 IEEE

If the traffic parameters in terms of traffic locality and specified loading condition exist in the network, the length for a transmission cycle can be reduced. It means that each REX in a hub needs not wait for a longer time to transmit. Hence, the network needs to reconfigure according to the traffic characteristics in terms of loading and traffic locality. In order to alleviate the drawback of long transmission length and improve performance, an architecture named as 2-level TDMA with ICS (Interleave Control Slot) is proposed to partition/reconfigure the network into several control groups. We denote r as control groups. Fig. 5 shows the transmission cycle for 2-level TDMA, *r=2,* at the switching hub level, the REX level and the IP packets level, respectively. The 2TDMA consists of *r* different groups of REXs. All the REXs in a group exhibit traffic locality and have the same control channel. Fig. 5 shows the slot-interleaved control slot for *r=2*, where REXs *1,3, …, k-1* are together as group *1* and REXs *2,4,…,k* are group *2*.

The 2TDMA is a network that consists of *r* groups of REXs. All REXs in a group have the same wavelength for control channel. For example, the channel assignment of REXs in a 2TDMA, *r=2*, in which REXs *1,3,…, k-1* and REXs *2,4,…,k* are grouped together. Since there are only *k/2* REXs in a group, the average waiting time of a REX for transmitting data in the 2TDMA is less than that in the TDMA as shown in Fig. 4. If we assume that the traffic in the same group exhibits locality, the slot utilization is almost twice than that of the original TDMA. However, this is an upper bound of slot utilization for 2TDMA with *r=2* if full traffic locality exists in each group.

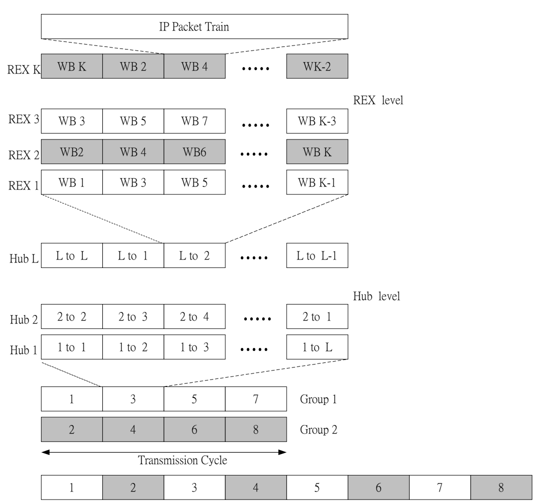


Figure 5. The modified transmission cycle for 2TDMA, *r=2*.

3.2 Interleaved control slot mechanism

In this section, the control slot arrangement is proposed in order to get good slot utilization in networks. There are many solutions to resolve the collision problem. For example, the subcarrier technique can be used to avoid the collisions. That is, the receivers of all REXs can do the channel inspection [14]. Apart from channel inspection, we propose a slot-interleaved mechanism as shown in Fig. 6 to prevent the collisions. Fig. 6 shows the slot-interleaved control slot for *r=2*, where REXs *1,3,5,7* are together as group *1* and REXs *2,4,6,8* are group *2*. In Fig. 6 each control slot consists of *2* parts: cross-group section (CGS) and non-cross-group section (NCGS), where the CGS allows a REX to transmit data to destinations in different groups and the NCGS can not transmit data across different groups. For example, for REXs of group *1*, if REX *3* wants to transmit data to REXs in the same group such as REX *1,5,or 7*, it can transmit the data during the full control slot containing the CGS and the NCGS. However, if REX *3* wants to transmit data to REXs in group *2,* such as REX *2,4,6 or 8*, it can only transmit the data during the CGS, that is a half of control slot. Owing to the traffic locality of the REXs in the same group, the cross-group traffic is less than the traffic within the same group. As a result, the probability of collisions will be very small and collisions should occur at most one half of each control slot. In order to detect the collisions, a special unit known as the collision manager will be described here. The collision manager that is a tunable receiver in each station monitors the transmitter wavelength to decide whether to transmit data or not. If the collision manager detects the wavelength that has been used by other REXs, then wait and try again in the next control cycle.

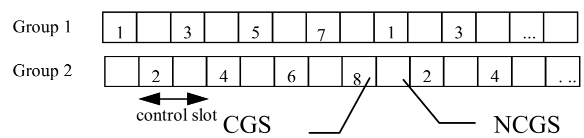


Figure 6. An example for slot-interleaved control

4. Performance Analysis

In the following, we will propose an analytical model and evaluate the network performance by comparing the relationship between the slot utilization and the number of groups. We assume each link has *k+r* wavelengths /channels ($\lambda_1, \lambda_2, \cdots, \lambda_{k+r}$) by employing the WDM technology, where channels $\lambda_1, \lambda_2, \cdots, \lambda_r$ are dedicated for control and others $\lambda_{r+1}, \lambda_{r+2}, \cdots, \lambda_{k+r}$ are for data. When several control groups wish to take place concurrently, a more general control pattern is formed. Assume there are *r* control groups $G_r = (g_1, g_2, ..., g_r)$, and g_i is the control group with D^i REXs where $1 \le D^i \le N$, $1 \le i \le r$, $\sum_{i=1}^{r} D^i = N$.

4.1 Analytical Evaluation

The network operates in a slotted mode, i.e., all

REXs are synchronized to the slot boundaries. We assume each REX has *N-1* single-packet buffers, one for the packets to each possible destination. Packets arriving at a full buffer are lost. However, packets that cannot be buffered at the MAC layer are not actually lost, but are typically buffered at a higher layer. Because the behavior of a channel in TDMA/WDMA is not independent of other channels, the assumption of single-packet buffer can make the analysis traceable. First, let ρ_i be the probability that a new packet arrives at REX *i* during a slot, P_{ij} the probability that a packet arriving at REX *i* is destined to REX *j*. Thus, $\rho_i P_{ij}$ characterizes the arrival process from the higher layer to the single-packet buffer for REX *j* when the buffer is empty. In addition the source-destination traffic distribution is derived from the following equality [13] that describes traffic locality:

$$P_{ij} = \begin{cases} 0 & j \le i \\ \dfrac{(1-p)^{j-i-1}p}{1-(1-p)^{N-i+1}} & j > i \end{cases} \qquad (1)$$

The Eq.(1) represents a normalized geometric distribution where $p(0 \le p \le 1)$ determines the level of traffic locality. [$\rho_i P_{ij}$] is the matrix of external traffic in units of packets per slot. Moreover we assume that the reconfiguration time of network is regarded as a uniform distribution (i.e., reconfiguration of a topology is completed within a specified time period). Furthermore, the case that $N = k$ is analyzed first to observe the effect of 2TDMA/WDMA among the slot utilization, the traffic locality and the network loading. The $N > k$ case or the unbalanced traffic can be easily extended.

As seen before, time slots are cycled in frames of N slots for $N = k$. The distance in slots between the beginning of the slot that REX *i* is pre-allocated to transmit to REX *j* and the beginning of such slot in the next or previous frame is denoted as $D^r{}_{ij}$. Let $R_{ij}(r)$ represent the probability of each corresponding case for a source-destination pair *(i,j),* $R^k_{ij}(r)$ represent the probability of each corresponding case for a source-destination pair *(i,j)* in the same group *k* where $i, j \in k$, and $R^{kl}_{ij}(r)$ represent the probability of each corresponding case for a source-destination pair *(i,j)* in different groups *k,l* where $i \in k, j \in l$. Therefore, $R_{ij}(1) = (1-(1-\rho_i P_{ij})^{N-1})$ because the probability that REX *i* has a packet for REX *j*, at the beginning of the slot is equal to the probability that at least one packet for REX *j* arrives at REX *i* during the previous $D^1{}_{ij} = N$ slots. For $R_{ij}(2)$, we have

$$R_{ij}(2) = R^1{}_{ij}(1) + R^2{}_{ij}(1) + R^{12}{}_{ij}(1) + R^{21}{}_{ij}(1) \ ; \quad D^1_{ij} + D^2_{ij} = N \qquad (2)$$

The first two terms of $R_{ij}(2)$ describe the probability that there are packets destined to group 1 and 2, respectively. The last two terms consider the situation that the packets are destined to the REXs in different groups. The $R^{12}{}_{ij}(1)$ or $R^{21}{}_{ij}(1)$ represents the probability for a source-destination pair *(i,j)* where the source *i* and the destination *j* are located in group 1 and 2, respectively . In general, we have

$$\begin{aligned} R_{ij}(r) &= R^1{}_{ij}(1) + R^2{}_{ij}(1) + \cdots + R^r{}_{ij}(1) + R^{12}{}_{ij}(1) + \cdots + R^{1r}{}_{ij}(1) + \\ &\quad R^{21}{}_{ij}(1) + \cdots + R^{2r}{}_{ij}(1) + \cdots + R^{r1}{}_{ij}(1) + \cdots + R^{r(r-1)}{}_{ij}(1) \\ &= \sum_{l=1}^{r} R^l{}_{ij}(1) + + \sum_{m=1}^{r}\sum_{n=1}^{r} R^{mn}_{ij}(1); \end{aligned} \qquad (3)$$

$$where \quad \sum_{k=1}^{r} D^r_{ij} = N, \ R^l{}_{ij}(1) = \begin{cases} R_{ij}; & if \ i \in l, j \in l \\ 0; & otherwise \end{cases} \quad and$$

$$R^{mn}{}_{ij}(1) = \begin{cases} R_{ij}; & if \ i \in m, j \in n, m \ne n \\ 0; & otherwise \end{cases}$$

$R_{ij}(r)$ consists of two terms: 1) the probability that there are the packets destined to REXs in the same group; 2) the probability that the packet destined to REXs in different groups. With $R_{ij}(1)$ and $R_{ij}(r)$ the slot utilization (*U*) and the system throughput (*T*) in a data slot can be obtained. *T* is obtained by the summation of utilization in each channel. Utilization in each channel *i* is determined by the successful packet transmission of station *i* to all possible destinations per data slot. The successful packet transmission of station *i* to station *j* is $R_{ij}(1)$ or the summation of $R_{ij}(r)$ for TDMA or 2TDMA, respectively. *U* is the average utilization of channels and can be obtained by averaging over *T*.

$$T = \begin{cases} \displaystyle\sum_{i=1}^{N} \frac{1}{N-1} \sum_{j=1}^{N} R_{ij}(1) & for \ TDMA \\ \displaystyle\sum_{k=1}^{r} \frac{1}{D^k_{ij}} \sum_{i \in k}\sum_{j \in k} R^k_{ij}(r) + \sum_{m=1}^{r}\sum_{n=1}^{r} \frac{1}{D^m_{ij}} \sum_{i \in m}\sum_{j \in n} R^{mn}_{ij}(r) & for \ 2TDMA \end{cases} \qquad (4)$$

$$U = \begin{cases} \displaystyle\frac{1}{N}\sum_{i=1}^{N} \frac{1}{N-1} \sum_{j=1}^{N} R_{ij}(1) & for \ TDMA \\ \displaystyle\frac{1}{N}\Big(\sum_{k=1}^{r} \frac{1}{D^k_{ij}} \sum_{i \in k}\sum_{j \in k} R^k_{ij}(r) + \sum_{m=1}^{r}\sum_{n=1}^{r} \frac{1}{D^m_{ij}} \sum_{i \in m}\sum_{j \in n} R^{mn}_{ij}(r)\Big) & for \ 2TDMA \end{cases} \qquad (5)$$

4. 2 Upper bound on the slot utilization (*U*) and the system throughput (*T*)

The slot utilization in each group is determined by the collision probability. If the source and the destination REXs of the traffic are located in the same group, i.e. the traffic is fully local in control group, the collision will not occur and the slot utilization is improved. As a result, the slot utilization is proportional to the traffic locality and the number of groups. In order to derive the upper bound of slot utilization, the collision is assumed to be non-existent. It means that the second term of the equations (4) and (5) is zero, $\frac{1}{D^m_{ij}} R^{mn}_{ij}(r) = 0, m \ne n$. Hence, we can express the slot utilization (*U*) and the system throughput (*T*) as

$$T = \begin{cases} \sum_{i=1}^{N} \frac{1}{N-1} \sum_{j=1}^{N} R_{ij}(1) & for \quad TDMA \\ \sum_{k=1}^{r} \frac{1}{D_{ij}^{k}} \sum_{i \in k} \sum_{j \in k} R_{ij}^{k}(r) & for \quad 2TDMA \end{cases} \qquad (6)$$

$$U = \begin{cases} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N-1} \sum_{j=1}^{N} R_{ij}(1) & for \quad TDMA \\ \frac{1}{N} \sum_{k=1}^{r} \frac{1}{D_{ij}^{k}} \sum_{i \in k} \sum_{j \in k} R_{ij}^{k}(r) & for \quad 2TDMA \end{cases} \qquad (7)$$

Fig. 7 shows the upper bound of slot utilization (U) and the system throughput (T) in different groups, in which the number of REXs in each group is assumed equal. For example, when the number of groups is 2, the number of REX in each group is *N/2* and the average waiting time of a REX for transmitting data needs only *N/2* time slots. As a result, we can find that the upper bound of slot utilization (U) and the system throughput (T), compared with one control channel, is r times if the network is partitioned into r groups.

TYPE	TDMA	2TDMA
Average number of REX in group	*N*	*N/2*
Upper bound of T (U)	X	2X

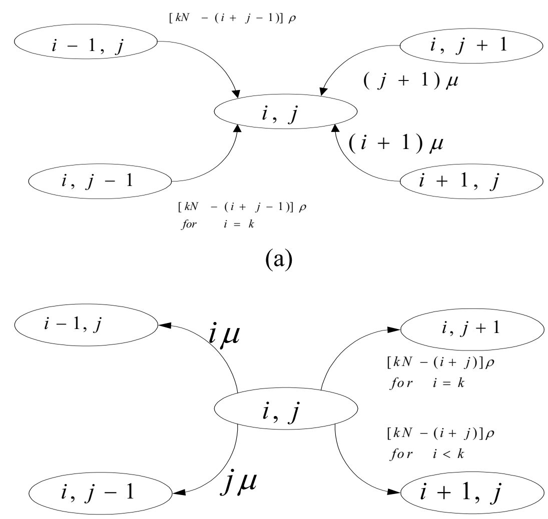
Figure 7. The upper bound of slot utilization vs. the number of REXs in different control groups.

4. 3 Packet Loss Probability

In order to evaluate the effectiveness of the ICS technique, we introduce an analytical model allowing us to calculate the IP packet loss probability $P_{loss}(r)$. We evaluate this performance index by assuming that the traffic offered by each input wavelength channel is modeled as an ON – OFF process where an ON period corresponds to the duration of an optical burst, and the cross group traffic scenario is symmetric, meaning that the input processes of cross group have the same statistic and an arriving burst has the same probability *1/N* to be directed to any REX. Due to the traffic increase in the same group is larger than that in the cross group, the same probability *1/N* is an over-estimated value for cross group traffic. We denote by L and S the random variables characterizing the n-th ON and OFF periods, and $E(L)$ and $E(S)$ denote the expected values, respectively.

We model the ICS in Fig.8 by means of a Markovian process whose state is constituted by two discrete variables (i,j), where i denotes the number of entering bursts directed to CGS and obtaining an wavelength (accept cross group transmission), whereas j denotes the total number of entering bursts that have been rejected due to the absence of available wavelengths(un-accept cross group transmission).

In order to evaluate the transition rate of the introduced Markov chain, as shown in Fig. 8(a), transitions entering a state *(i,j)* can occur from the states: $(i\text{-}1,j)$, when a new burst arrives and there is at least one available wavelength; $(i,j\text{-}1)$, when a new burst arrives and all of the wavelengths are busy; $(i,j+1)$, at the end of a previously rejected burst; $(i+1,j)$, at the end of a previously accepted burst.



(a)

(b)

Figure 8. (a) Entering and (b) outgoing state transition diagram.

The rates relative to the mentioned transitions are illustrated in Fig. 8(b) where $\rho = 1/N \cdot E(L)$ and $\mu = 1/E(L)$ .The transitions outgoing from the state *(i,j)* are, as illustrated in Fig. 8(b), toward the following states:
$(i,j+1)$, when a new burst arrives and all of the wavelengths are busy; $(i+1,j)$, when a new burst arrives and there is at least one available wavelength; $(i,j\text{-}1)$, at the end of a previously rejected burst.; $(i\text{-}1,j)$, at the end of a previously accepted burst.

Once evaluated, the limit state probability P_{ij} of the introduced chain we can determine $P_{loss}(r)$,

$$P_{loss}(r) = 1 - \frac{A_{success}}{A_{Full}} = 1 - \frac{\sum_{j=0}^{kN-k} \sum_{i=0}^{k} iP_{i,j}}{k \cdot \frac{E(L)}{E(L)+E(S)}} \qquad (8)$$

In order to solve (8) easily, an over-estimated value of $P_{loss}(r)$ can be obtained by the Engset formula [16]. We assume a new burst will arrive according to an exponentially distributed interval with $\rho = 1/N \cdot E(S)$ transition rate if a rejected burst engages its input wavelength channel for the duration. Hence, the offered traffic is more peaked than the real one. Under this consideration, we can express the packet loss probability as follows:

$$P_{loss}(r) \approx \frac{\binom{kN-1}{k}\left(\frac{\rho}{\mu}\right)^k}{\sum_{j=0}^{k}\binom{kN-1}{j}\left(\frac{\rho}{\mu}\right)^j} \quad (9)$$

4.4 Numerical Results

In Fig. 9 the analytical values of $P_{loss}(r)$ regarding 2TDMA adopting the ICS technique have been plotted as a function of the number of wavelengths for a number of REX *N=16* and the arrival rate ρ varying from 0.1 to 1. The results show that a reasonable number of wavelengths (*k*=12,16) can achieve acceptable packet loss probability ($10^{-4} \sim 10^{-7}$) if the arrival rate on each wavelength is not greater than 0.4. Of course, if the number of wavelengths is increased, the packet loss probability will be reduced as well.

Fig. 10 shows the utilization in the 2TDMA network with high locality p=0.8 under different group *r*=1 and *r*=2. Note the traditional TDMA is a special case of 2TDMA with *r=1*. The increased traffic arrival rate could increase the utilization. In Fig.10, we can observe that the 2TDMA technique allows improvement of the utilization almost *2* times in the case of high traffic locality in a group. Fig. 11 shows the utilization in the 2TDMA network under the different traffic locality. The increased traffic locality could increase the slot utilization. Beside, 2TDMA can make the slot utilization higher as the traffic locality increases. However, the increase in the slot utilization by 2TDMA is not proportional to the increase in the traffic locality. In fact, the increase in slot utilization saturates when the traffic locality or traffic arrival rate increases over some threshold.

In order to observe the relationship between the throughout and the number of groups under different traffic locality and loading, we compare the cases of TDMA and 2TDMA under different network loadings. Fig. 12 and 13 show the relationship among throughout, locality and groups with different loading (ρ = 0.1, 0.9), respectively. The traditional TDMA is a special case of 2TDMA with *r=1*. In Fig.12, if the locality is larger than the threshold 0.3, a network with 2TDMA can have a higher throughout than a network with TDMA. On the other hand, if the system is operated with locality below the threshold 0.3, it is possible that TDMA can generate higher throughput than 2TDMA. The poor performance of 2TDMA at low traffic locality is due to the fact that the increasing cross group traffic will result in collision probability increase. If the traffic locality is equal to 1,i.e. full locality, the slot utilization almost reaches its upper bound. Hence, the results in Fig. 12 show that the traditional TDMA technique does not allow a REX to grab more than the portion of the channel bandwidth assigned to it even if no other REXs are transmitting on the same channels. However, under high traffic locality, the throughput of the TDMA is very poor. The 2TDMA can be adopted to improve the slot utilization and throughput.

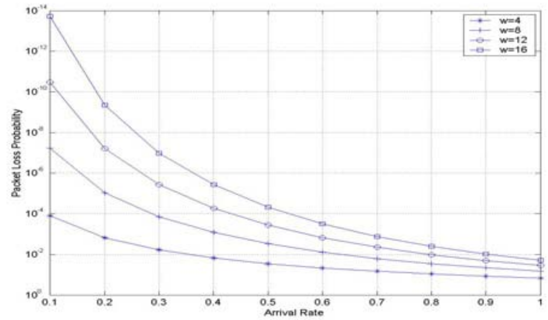


Figure 9. Comparison of analytical model of ICS technique with N=16 and the arrival rate in the range [0.1~1].

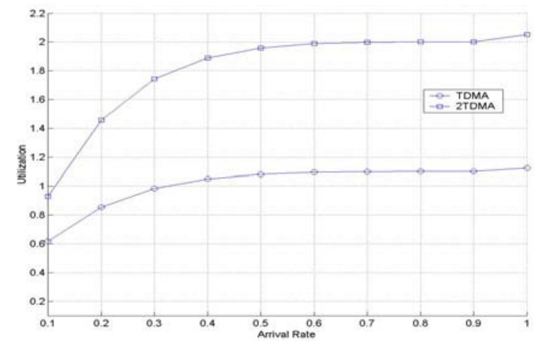


Figure 10. Utilization vs. traffic rate for TDMA and 2TDMA.

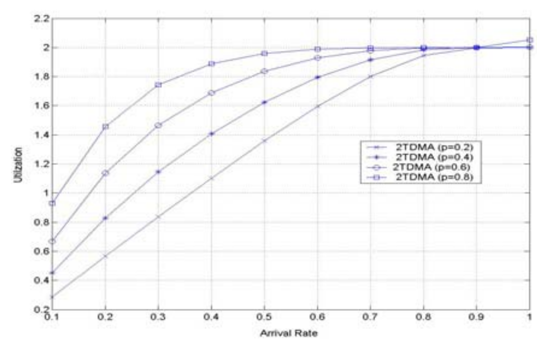


Figure 11. The slot utilization for 2TDMA.

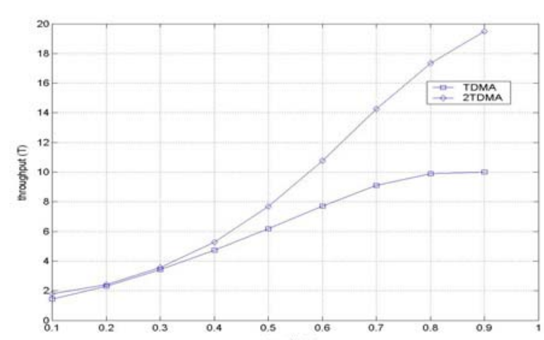


Figure 12. Throughput vs. traffic locality with *r=1* and *2* (ρ =0.1).

1530-2016/03 $ 17.00 © 2003 IEEE

Fig. 13 shows the throughput as a function of different groups under heavy network loading (ρ =0.9). The result shows that the case of 2TDMA with *r=1* has the best throughput under heavy network loading. The increased number of groups has no benefit in throughout. It is because of the high traffic arriving rate results in high traffic in each groups. Hence, the traffic in each group always has collisions. The throughput of 2TDMA is not significant. From the results of Fig. 10 and 13, it can be concluded that a network with high utilization is not necessary a network with high throughput. Although 2TDMA can always make the slot utilization higher than TDMA, it cannot fully show its advantage for higher arrival rate (loading). Hence, if a system can be operated with appropriate loading, group, and traffic locality, its utilization and throughput will be high. Therefore, we can conclude that TDMA is a better choice if the network is without traffic locality, otherwise 2TDMA is better in traffic with high locality.

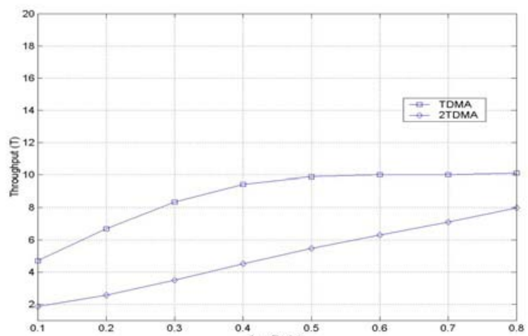


Figure 13. Throughput vs. traffic locality with *r=1* and *2* (ρ =0.9).

5. Conclusions

In this paper we have proposed a new technique and architecture for IP support in WDM optical networks, which is based on a two-level TDMA structure with wavelength division multiplexing (WDM). This solution integrates three technologies: IP addressing, label switching, and WDM network based on TDMA technology. The 2TDMA has also been adopted to improve the performance. Each control group in the 2TDMA has the same wavelength for the control channel based on the slot-interleaved structure.

The virtual topology of the networks, i.e. control groups, is reconfigured by the traffic locality in the networks. Due to the traffic locality of the REXs in the same group, the cross-group traffic is less than the traffic within the same group. As a result, the probability of collision is very small and collision will occur at most one half of each control slot.

We also have illustrated and analyzed the 2TDMA technique, which allows a significant increase in the performance of the network, in terms of slot utilization and throughput, with respect to the case in which only the traditional TDMA (a special case of 2TDMA, *r=1*) is adopted. An analytical model has been presented. The results of analytical model show that 1) the performance of network depends on the network loading, locality, and number of groups, 2) the increased traffic locality could increase the slot utilization and throughput. If traffic locality is larger than a threshold, a network with 2TDMA can generate higher throughout than a network with TDMA, 3) the traditional TDMA has better performance than 2TDMA under the heavy load. Hence, if a system can be operated with appropriate loading, grouping, and traffic locality, its utilization and throughput will be expanded. Therefore, the TDMA is the option in heavily loaded networks without traffic locality. Otherwise the 2TDMA is a better option in networks with high traffic locality.

References

[1] E. C. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture", *IETF draft-ietf-mpls-arch-06.txt*, Aug. 27, 1999.

[2] P. Cadro, A. Fravey, and C. Guillemot, "Performance evaluation of the KEOPS wavelength routing optical packet switch", *ETT*, vol. 11, no. 1, pp. 125 – 132, Jan. – Feb. 2000.

[3] F. Callegati, M. Casoni, C. Raffaelli, and B. Bostica, " Packet optical networks for high-speed TCP-IP backbones," *IEEE Commun. Mag.*, vol. 37, pp. 124 – 129, Jan. 1999.

[4] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, and I. Andonovic, "SLOB: A switch with large optical buffers for packet", *IEEE J. of Lightwave Tech.*, vol. 16, no. 2, pp. 1725-1736, OCT, 1998.

[5] S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: An overview," *IEEE Commun. Mag.*, pp. 84 – 94, Feb. 2000.

[6] C. Qiao and M. Yoo, "Optical burst switching (OBS): A new paradigm for and optical Internet", *J. of High Speed Networks (Special Issue on Optical Networking)*, vol. 8, no. 1, pp. 69–84, Jan. 1999.

[7] Andrea Detti, V. Eramo, and M. Listanti, " Performance Evaluation of a New Technique for IP Support in a WDM Optical Network: Optical Composite Burst Switching (OCBS)", *IEEE J. of Lightwave Tech.*,vol. 20, no. 2, pp. 154–165, Feb. 2002.

[8] K. Claffy, H.-W. Braun, and G. Polyzos, "Traffic characteristics of the T1 NSFNET backbone", *Proc. IEEE INFOCOM'93* , Jan 1993.

[9] N. Gulati, C. Williamson, and R. Bunt, "*LAN traffic locality: Characterization and application*", Proceedings of the First International Conference, October 1993.

[10] K. C. Claffy, H.W. Braun, G. C. Polyzos, "A parameterizable methodology for Internet traffic flow profiling", *IEEE Journal on Selected Areas in communications*, vol.13, no. 8, pp. 1481-1494, October 1995.

[11]http://traffic.caida.org/TrafficAnalysis/traffic.measure/mgp00001.html

[12] S. A. Thomas, "*The TCP/IP Protocols: Implementing the Next Generation Internet",* Wiley, New York, 1996.

[13] N. F. Huang and S. T. Sheu, "An efficient wavelength reusing/migrating/sharing protocol for dual bus lightwave networks", *IEEE J. of Lightwave Tech.*, vol. 15, no. 1, pp. 62-75, January, 1997.

[14] M. Ajmoni Marsan, A. Fumagallim E. Leonardi, F. Neri, "R-Daisy: an all-optical packet network", *EUROPTO European Symposium on Advanced Networks and Services*, Amsterdam, Holland, March 1995.

[15] Chen-Ken Ko, and Sy-Yen Kuo, "Multiaccess processor interconnection using subcarrier and wavelength division multiplexing", *IEEE J. of Lightwave Tech.*, vol. 15, no. 2, pp. 228-241, FEB, 1997.

[16] L. Kleinrock, "*Queueing Systems",* New York: Wiley, 1975, vol. 1, pp.164–169.