SPACES: A Distributed Fairness Algorithm for Bus-Based Metropolitan Optical Network

Daniel POPA and Tülin ATMACA,

Institut National des Télécommunications, Dept. RST
9 rue Charles Fourier, 91011 Evry, France

Summary

In metropolitan optical networks, the fairness issue has mostly been analyzed with regard to synchronous transmission (i.e., slotted WDM ring). In this context, we focus on the performance of asynchronous CSMA/CA protocol with variable packet sizes. We examine the fairness issue arising between ring nodes sharing a common data channel and show that the asynchronous CSMA/CA protocol suffers from *bandwidth fragmentation* and *positional priority* fairness problems. The paper proves that the phenomenon of *bandwidth fragmentation* represents the most important issue in asynchronous CSMA/CA protocol because it leads to bandwidth waste and, thus, increases unfairness for certain ring nodes. In this regard, we explore limitations of a previous proposed solution, the so-called Traffic Control Architecture using Remote Descriptors (TCARD). Afterwards, we present an algorithm, called **S**mart s**PACE**d tran**S**mission (*SPACES*), which deals with the aforementioned fairness problems. The devised scheme is based on a statistical and fully distributed mechanism to grant fair access to shared resources. As illustrated in this article, *SPACES* algorithm alleviates performance degradation and resources underutilization while achieving fairness among ring nodes.

Key words:
Asynchronous optical CSMA/CA protocol, Variable length packet, Shared bus ring network, Distributed fairness algorithm, Bandwidth fragmentation, Performance.

1. Introduction

A crucial aspect of networking in metropolitan area is the design of a robust and simple access scheme for building bandwidth-efficient and flexible metropolitan optical networks [11][12]. The asynchronous CSMA/CA scheme was proposed in two recent experimental projects: Dual-Bus Optical Ring Network (DBORN) [7][10] and Hybrid Optical Ring NETwork (HORNET) [5], respectively. As it appears, the asynchronous CSMA/CA scheme is a good candidate for access protocol in next generation of metropolitan optical networks. Several arguments come in favor of this affirmation. First, it can be a fully-distributed asynchronous protocol not based on a centralized controller or separate control wavelength to synchronize the operation of nodes on the ring. This is an advantage in implementation compared to slotted optical MAC protocols. In WDM ring networks (e.g., [3]), slotted MAC protocols maintain synchronous slot boundaries over many wavelengths through dispersion compensation. Next, an access protocol based on asynchronous CSMA/CA technique shares efficiently the available capacity between access points at the time scale of individual packets. Finally, when the asynchronous CSMA/CA protocol is used in passive optical ring configuration costs decrease significantly (e.g., [10]).

In this article, we focus on the performance issue in DBORN. The DBORN architecture will be briefly described in this paper, but for more details the reader is invited to refer to [7]. The performance of asynchronous optical CSMA/CA technique relies mainly on how optical resource (i.e., channel/wavelength) is shared among different access nodes. Even though the MAC protocol enables collision-free transmission over the shared medium, it does not address the inherent fairness issue which is pronounced in the case of shared medium ring networks. Previous analyses on performance of shared medium networks with ring topology have polarized the research community to the major problem of such configuration: The performance of a node depends upon its physical location within the network. The relative performance between ring nodes was shown to be quite unfair.

Unfortunately, however, the fairness issue is more complex in ring configuration with CSMA/CA protocol. The fairness-related problems, as we shall show in this article, are *positional priority* and *bandwidth fragmentation* respectively. *Positional priority* is a fairness issue common to all ring architectures (i.e., slotted or unslotted, with or without spatial reuse of the bandwidth). Because the priority is given to upstream traffic in this access scheme, a problem known as "starvation" [14] can happen when downstream nodes are constantly covered by upstream traffic and not able to access network for a very long period of time. In CSMA/CA protocol, *positional priority* is more important because the spatial reuse of the bandwidth is not possible. *Bandwidth fragmentation* appears as a direct consequence of the random and asynchronous insertion of optical frames among ring nodes: Upstream nodes create useless

---

Manuscript received August 5, 2006.
Manuscript revised August 25, 2006.

idle periods that increase fairness problems at downstream nodes. There are only two papers available identifying the phenomenon of *bandwidth fragmentation* as a factor of unfairness in asynchronous CSMA/CA protocol. We highlight, however, that previous analytical works [1][2] do not provide any detail how this phenomenon influences the performance of the access protocol. The complexity of the developed mathematical models shows that this issue is complicated by the lack of possibility to analytically describe such process and, thus, of sufficiently accurate knowledge of distribution of voids among ring nodes. Therefore, we use simulation models to provide an intuitive explanation of the interaction between *bandwidth fragmentation* and the fairness issue in DBORN. Furthermore, we prove that *bandwidth fragmentation* represents the most important issue in such architecture.

Most of the schemes proposed to deal with the starvation problem in ring configurations use global fairness algorithms. Existing mechanisms proposed in MAGNET [17], Orwell [18] and ATMR [15] are all global fairness algorithms. Such algorithms operate using fairness cycles and allow nodes to enter the new cycle after an idle period of time between successive fairness cycles. The local fairness algorithms put a quota on the transmission of upstream nodes in order to preserve enough bandwidth for downstream nodes. Some of proposed schemes (e.g., MetaRing [14]) regulate access through a token-based-like approach, since others (i.e., Token Bucket) simply survey that access nodes do not violate the negotiated throughput. However, as we shall explain latter in this article, aforementioned fairness control mechanisms do not perform satisfactorily in ring configuration with CSMA/CA scheme.

Recently, Ciavaglia *et al.* [2][6] proposed the so-called Traffic Control Architecture using Remote Descriptors (TCARD) to alleviate the fairness issue in DBORN. Although this scheme deals with the starvation problem at light-to-moderate load, we present in this paper unexpected behavior of TCARD algorithm. For moderate-to-high load, starvation moves backward to upstream nodes.

The design of a scalable fairness algorithm is not only limited to the performance of the access protocol, but also to traffic dynamicity. The parameters setting of TCARD in any case relies on the information regarding the average of the offered traffic load in downstream nodes. In real network operation, due to traffic variation, is not always easy to predict precisely the load of the best effort traffic in edge nodes and, thus, to collect information for setting of the TCARD parameters. In this context, we introduce a statistical and fully-distributed traffic control mechanism to deal with the fairness issue in DBORN. Our solution, called **S**mart s**PACE**d tran**S**mission (*SPACES*), is based on a "smart" smoothing mechanism that aims dealing with both *bandwidth fragmentation* and *positional priority* fairness-related problems. *SPACES* regulates the asynchronous transmission of optical frames by using two timers to control the process of emission. The performance evaluation study shows that this algorithm deals efficiently with aforementioned fairness problems. As illustrated in this paper, the proposed solution achieves performance and fairness among ring nodes.

The remaining parts of this article are organized as follows. In Section 2, we present the basic principles of the network architecture. The fairness issue is discussed in Section 3. Afterwards, we explain, in Section 4, why the traditional fairness algorithms cannot perform satisfactorily in ring topology with CSMA/CA scheme. Additionally, we explore limitations of the TCARD mechanism. The principles of the *SPACES* scheme are explained in Section 5. In Section 6, we show through simulation results that the *SPACES* algorithm achieves fairness and improves the performance of the access protocol. Finally, conclusions are provided in Section 7.

2. The Network Architecture

The network architecture is depicted in Fig. 1. The system is organized around a hub which connects the metro network to the backbone. Two counter rotating optical fibers (for protection reasons) with transparent and passive Optical Add Drop Multiplexers (OADMs) compose the network transmission medium.

The architecture is based on two buses: an upstream and a downstream bus, respectively. The upstream bus is composed of one or several wavelength(s) created in edge nodes and terminated at the hub node. Downstream wavelengths start from transmitters in the hub and are terminated in ring nodes. Several edge nodes share upstream and downstream channels respectively in asynchronous time division multiplexing. An access node can be attached to more than one upstream or downstream wavelength. A simple collision avoidance mechanism is implemented through photodiode power detection on each locally accessible upstream wavelength. Thanks to a sliding window realized by a photodiode and a Fiber Delay Line (FDL), ring nodes detect idle periods and capitalize adequate null periods. A size of FDL slightly larger than the network Maximum Transmission Unit (MTU) ensures packet insertion on upstream wavelengths without collision. When the null period is too small, the information remains electronically buffered in ring nodes until a sufficient void space is detected.

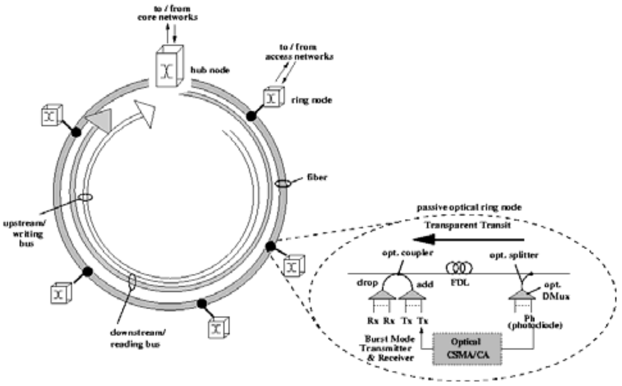

Fig.1: The network architecture

It is obvious that this architecture presents two major drawbacks with the regard to the network performance. First, there is no direct communication between the access points and hence all traffic - external and intra-ring - has to pass through the hub. Next, edge nodes cannot receive or remove transit frames on upstream channels (i.e., no spatial reuse of the bandwidth) or insert traffic on downstream channels.

This architecture implies multipoint-to-point communication on upstream bus and point-to-multipoint on downstream bus. We notice that the fairness issue arises only on upstream bus, where several nodes compete for the bus bandwidth. Since the hub node exclusively transmits on downstream wavelengths, traditional scheduling mechanisms can be applied here. Therefore, the access mechanism has to be controlled at the electronic edge on upstream bus and this is the topic of the following sections.

3. The Performance Issue

In this section, we firstly explain the *positional priority* and *bandwidth fragmentation* problems. Afterwards, we analyze the efficiency of the access protocol and highlight the impact of the aforementioned issues on the network performance.

To begin with, let us consider the situation when two nodes share the bus. The channel is always idle for the first node on upstream bus. The packet situated in the head-of-line of the second node can access the available bandwidth only if the transmission medium is free. Therefore, the transmission process of the second node depends on the activity of the first node. Consequently, the process of transmission at j^{th} node ($j = 2,..., N$) depends on the activity of previous $j-1$ nodes. When upstream nodes constantly occupy the transmission media, starvation takes place at downstream nodes. In addition, access protocol cannot provide concurrent transmission over distinct segments of the ring (i.e., spatial reuse of the bandwidth), because the hub is the only one which remove frames from the upstream channels. That is, in such architecture, starvation represents a fairness issue more important than in other ring configurations.

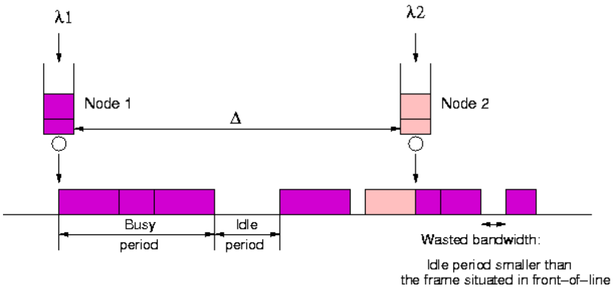

Fig.2: The phenomenon of bandwidth fragmentation

As illustrated in Fig.2, the data channel can be in one of two states: free (idle) or occupied (busy). The first node on upstream bus can fragment bandwidth in useless idle periods. When the size of voids created at the first node is not sufficiently large to allow the collision-free transmission of the item situated in front of the second queue, the transmission is blocked. Because the service discipline is based on FIFO scheduling, packets smaller than one situated in head of the queue cannot use these idle periods - the so-called Head of Line (HoL) blocking phenomenon and, thus, the bus bandwidth is wasted

3.1 Positional Priority and Bandwidth Fragmentation

In what follows, we analyze the performance of the asynchronous CSMA/CA protocol. We consider a ring with N nodes sharing a common medium (e.g., one optical channel) which channels traffic to the hub. The aim of this study is to present the average access delay (E[W_i}], $i = 1,..., N$) and the distribution of the size of idle periods among different access nodes. We define access delay as the time spent by a packet in queue i until its successful transmission.

In this study, simulations are carried using Network Simulator 2 (ns-2) [9]. In our simulation works, unless is otherwise specified, we assume that:

- All the ring nodes share a common upstream optical channel with a capacity of 10 Gbps.

- We assume a number of eight nodes (i.e., $N = 8$) competing for the same resource.
- The process of client arrivals is modeled by a Poisson process.
- The arrival rate (e.g., λ) of clients at each edge node is identical (i.e., balanced load: $\lambda_i = \lambda_j$, $i \neq j$, $i,j = 1 \ldots N$) in order to underline the fairness issue.
- The capacity of the electronic memory at each ring node is set to 5 Mbytes.
- We consider an optical overhead of 56 Bytes, as described in [8].

For the rest of this article, ρ_i denotes the average traffic load to be guided by the upstream bus from the i^{th} node to the hub. We call ρ the average traffic load of the upstream bus ($\rho = \sum_{i=1}^{N} \rho_i$). In addition, we consider electronic packets of variable size (i.e., 64, 576 and 1518 bytes) representative of the peaks in packet size distribution in Internet. The packet size distribution is as follows [8]: 10 percent of traffic volume is composed of packets of 64 bytes (i.e., TCP ACK), 40 percent of traffic volume is composed of packets of 576 bytes and the remaining 50 percent represents packets of 1518 bytes (i.e., Ethernet MTU).

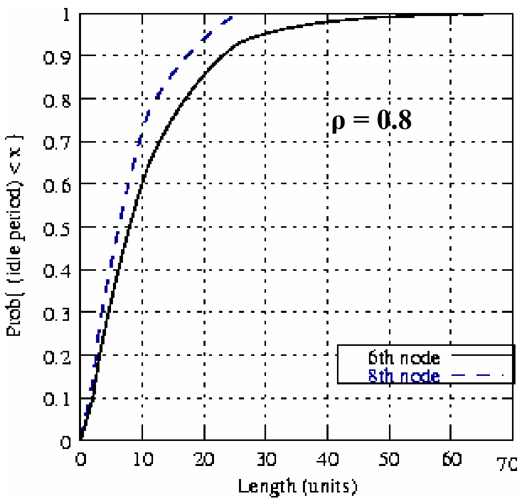

Fig. 3: Statistics on void lengths

Before presenting the simulation results, notice that 95 percent confidence intervals were computed. Because the obtained width of confidence intervals is of 10^{-3}, the confidence intervals are not drawn here in order to not obscure the data.

In Fig. 3, we present statistics on idle periods circulating on bus. The curves plotted here illustrate the void length distribution as measured at transmission tap of the node number six and eight. We assume an upstream bus loaded at 80 percent of its capacity. For this analysis, an unit represents 64 Bytes. Therefore, the transmission of the largest optical frame (i.e., network MTU) requires a void with a length of approximately 25 units (i.e., $\lceil (1518+56)/64 \rceil = 25$ units). From these results it is immediately evident that *bandwidth fragmentation* leads to fairness problems at downstream nodes: Bandwidth is arbitrarily fragmented and most of "sensed" idle periods are less than the largest optical frame.

The simulation results of average access delay are illustrated in Fig. 4. Results presented here are significant. We observe that for light loads (i.e., ρ = 0.3), the protocol performance is good. For moderate-to-heavy load (curves named "rho = 0.7" and "rho = 0.8"), the downstream node's performance is strongly affected. The poor performance of the access protocol is not due to the saturation of transmission media. The fairness problems are important even the channel occupancy is not high.

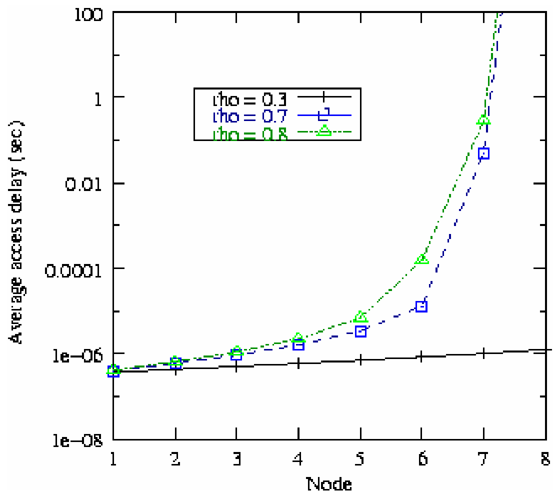

Fig. 4: Access delay versus position on ring

In Fig. 5, we depict the channel utilization as a function of the offered load (i.e., ρ). Note that the curve named "Ideal" represents the ideal ratio of bandwidth utilization when no bit loss occurs and no optical overhead is used to transport client layer information. On one hand, we observe that network remains stable for light loads of the bus; a threshold of $\rho = 0.6$ can be defined. Beyond this threshold the channel utilization becomes saturated. This has a simple explanation. Because most of the voids "sensed" at downstream nodes are less than network MTU, the process of transmission is blocked once the largest electronic packet arrives in head of downstream queues. Packets smaller than one situated in front-of-queue cannot use these idle periods and, thus, the buffer overflow occurs.

On the other hand, we remark that the channel utilization, as measured by simulation, increases faster than the "Ideal" case. This is due to the optical overhead used to transport electronic packets at optical level. For example, the transmission of a small electronic packet of 64 *Bytes* creates an overload (also called *tax*) of approximately [1-(64/(64+56))]*100 = 50 percent.

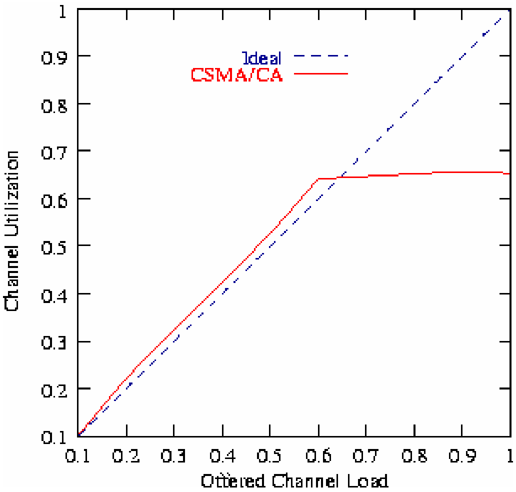

Fig. 5: Channel utilization vs. offered load (ρ)

The results presented in this section show clearly that the phenomenon of *bandwidth fragmentation* leads to significant performance degradation at downstream nodes. Here, the mismatch between the size of idle periods circulating on the transmission medium and the length of frames situated in head of downstream queues relies in bandwidth waste as well as fairness problems with regard to resource access.

To give an end, we underline that fairness algorithms to be adopted in shared-bus architecture should treat equally both *bandwidth fragmentation* and *positional priority* fairness-related issues. As we shall explain in the next section, most of traditional fairness algorithms do not fit with DBORN architecture, since others cannot deal efficiently with the phenomenon of *bandwidth fragmentation*.

4. Related Works in Fairness Issue

Between the algorithms proposed to deal with fairness issue in ring topologies, many of them require operations that are not well suited to very high-speed optical networks (e.g., DQDB [16]) or are incompatible with such dual-bus ring architecture (e.g., S++ [13]} requires operation at bit-level on transit traffic).

Generally, fairness issue in ring architecture can be solved following two approaches. The former introduces waiting lines in ring nodes to intercept the traffic in transit (i.e., RPR [4]). This approach does not fit to the DBORN architecture because ring nodes are transparent and cannot process the transit traffic. Yet, introducing in-line buffers at each node increases the network costs and this is an important constraint in networking in metropolitan area [11]. As for the later, we can proceed as follows. The unfairness is dynamically detected and explicitly signaled to upstream nodes (e.g., ATMR [15], MetaRing [14]). Otherwise, ring nodes limit locally their activity using a transmission-based quota (i.e., Token Bucket - TB). Explicit signaling of unfairness problem relies on the detection of node(s) which cause(s) starvation. On one hand, such approach fails in DBORN because the network throughput can reduce significantly due to the important time of distribution of the control message between neighbor nodes (i.e., a message makes a complete tour of the ring to pass to both downstream and upstream stations). On the other hand, the access protocol detects only the presence or not of the light on transmission media and, thus, cannot extract information from the transit traffic.

Local fairness algorithms using transmission-based quota can share the bandwidth fairly among ring nodes. Yet, the bandwidth access at ring nodes remains arbitrarily. Bouabdallah *et al.* [2] have shown that transmission-based quota mechanisms (i.e., TB rate-based) do not ensure satisfying results in asynchronous-based architecture like DBORN. In addition, they analyzed the behavior of the TCARD scheme and showed that TCARD provides a somewhat higher throughput to the access protocol. However, TCARD leads to unexpected behavior at heavy traffic loads, as we shall see next.

4.1 TCARD Limitations

In what follows, we provide a short description of TCARD mechanism. For further details, the reader is invited to refer to [6].The basic operation of such scheme is as follows. For each ring node i ($i = 1,..., N-1$), TCARD maintains an anti-token pool. A TCARD anti-token prevents the node i from transmission, and hence reserve bandwidth, for a fixed amount of time d, on a detected idle period. The size of d is upper bounded by the length of the FDL (i.e., the length of network MTU). Idle periods "sensed" on the bus can be used by the node for its own usage purpose as long as the anti-token pool is empty. The arrival of an anti-token to the pool during the transmission of a frame does not preempt the emission process of a local frame. The anti-token has to wait the end of packet transmission in order to be served. The rate of anti-tokens (i.e., *B*) and the size of the anti-token are two parameters directly related to the TCARD scheme. The anti-token rate B_i, at node *i*, corresponds to the total amount of average bandwidth to be reserved for its (*N-i*) downstream partners, and can be computed according to the following formula:

$$B_i = \frac{C * \sum_{j=i+1}^{N} \rho_j}{d} \qquad (1)$$

Here, *C* denotes the channel capacity, ρ_j is the portion of bandwidth reserved for the node *j*, and *d* is the size of the anti-token.

Fig. 6 plots the void length distribution as "seen" by the fourth node, as TCARD is enabled and disabled respectively. For this study we consider a bus loaded at 70 percent (i.e., ρ = 0.7) of its capacity. We assume *d* = network MTU = (1518+56) = 1574 Bytes. As for parameter B_i, if the first node must reserve 6.125 Gb/s on average of available bandwidth for downstream nodes, anti-tokens are generated periodically at rate equal to (6.125 Gb/s)/(1574*8 b/anti-token) anti-tokens/s. Similar to the first node, the second node reserves 5.25 Gbp/s for its downstream partners, and so on. The curves show clearly that the phenomenon of *bandwidth fragmentation* is not significantly alleviated.

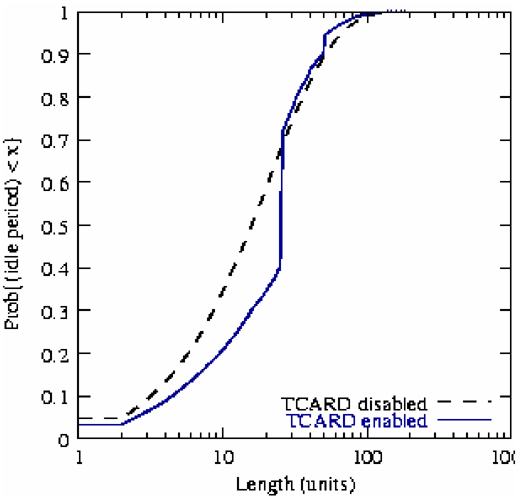

Fig. 6: Statistics on void length: The cdf measured at transmission tap of the 4th node; ρ = 0.7

TCARD has a major drawback. The manner of bandwidth reservation cannot substantially avoid the phenomenon of *bandwidth fragmentation*. On one hand, downstream nodes can fragment arbitrarily the idle periods reserved by their upstream partners. On the other hand, the first node on upstream bus fragments arbitrarily the bus bandwidth between two anti-tokens.

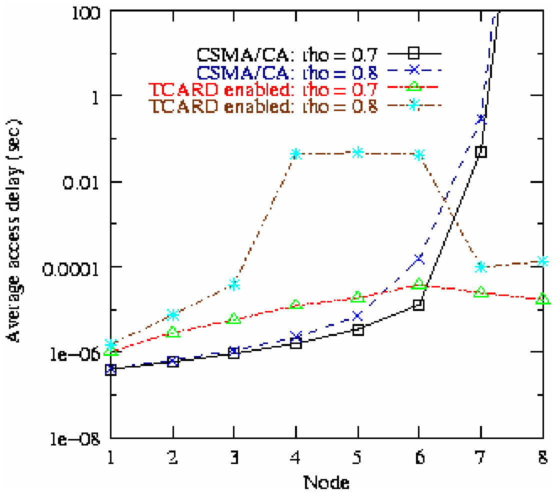

Fig. 7: Access delay: CSMA/CA vs. TCARD

Even if TCARD preserves a certain quantity of voids equal to network MTU, the starvation can still take place at relatively heavy traffic loads. The curves depicted in Fig. 7 come in favor of the previous affirmation. The results show the average access delay for an upstream bus loaded at 70 and 80 percent respectively of its capacity. When ρ = 0.7, the mean waiting time grows fast in the beginning and then tends to be constant or even decreases. This implies a better fairness than the case when TCARD is disabled (curve called "CSMA/CA: rho = 0.7"). With ρ = 0.8 and TCARD enabled, the curve "explodes", then tends to be constant and sinks for the downstream nodes. This leads to unfairness: The starvation moves backward to upstream nodes. Such unexpected behavior can be explained by two aspects: *i)* the amount of large free periods sensed by an edge node is not enough to consume immediately anti-tokens and, *ii)* frames transmission is stopped until the anti-token pool is emptied.

Certainly, when the TCARD algorithm is enabled, upstream nodes reserve bandwidth for their downstream partners. However, the size of voids reserved for downstream nodes is not sufficiently large to avoid starvation for moderate-to-heavy load. In this context, we introduce a fairness algorithm that aims controlling packets transmission smarter than TCARD.

5. The Proposed Algorithm: SPACES

Let us consider *N* nodes sharing a common unidirectional channel. For each ring node *i* (*i = 1*, ..., *N*), *SPACES* maintains two timers (i.e., reservation T_{reserv} and transmission T_{transm} timers respectively). These timers are used to constraint the access to the transmission medium of the i^{th} node.

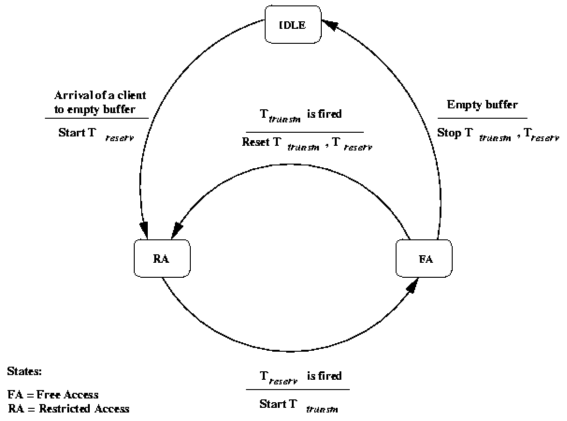

Fig. 8: *SPACES* algorithm – State transition diagram

The basic operation of such a scheme is simple. As shown in Fig. 8, our fairness algorithm has three states: *Idle*, *Restricted Access (RA)* and *Free Access (FA)*, respectively. Initially, the SPACES algorithm is in *Idle* state. Packets arriving to an empty buffer trigger the reservation timer and determine the transition to the *RA* state. In this state, nodes cannot transmit data as long as the reservation timer is not fired. The motivation for buffering packets during T_{reserv} time units is to minimize the phenomenon of *bandwidth fragmentation* and to create idle periods of T_{reserv} time units. When the reservation timer is fired, the protocol switches from the state *RA* to the state *FA*, and the transmission is released for a period of T_{transm} time units. In this state, nodes can transmit only a quota of data units before they transit to either *Idle* or *RA* state. If the buffer is emptied during the *FA* state, timers are stopped and the protocol transits to *Idle* state. When the end of T_{transm} occurs and the buffer is not empty, the protocol transits back to the *RA* state. The end of T_{transm} does not preempt the transmission of a frame in order to keep the complexity low and to avoid wasting bandwidth with packets partially transmitted. Yet, the current packet in emission must be buffered for a latter transmission. Therefore, the SPACES mechanism has to wait the completion of the service of current frame to stop transmission.

The *SPACES* algorithm has two key features. First, it alleviates the positional priority problem by forcing ring nodes to smooth smartly the transmission of local traffic: Each ring node controls its transmission by a series of busy and idle periods of activity. Next, unlike TCARD which reserves voids upper bounded to the size of the FDL, the *SPACES* algorithm can create idle periods with lengths several degrees of magnitude bigger than network MTU.

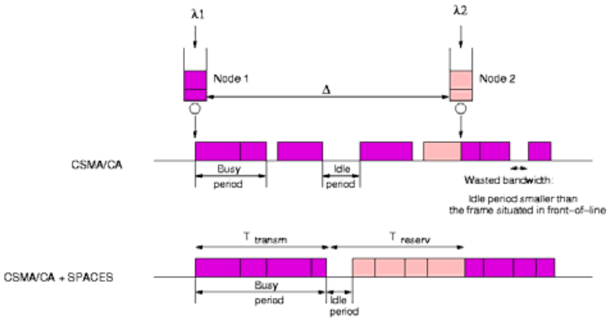

Fig. 9: Bandwidth fragmentation: CSMA/CA vs. SPACES

To illustrate the efficiency of the *SPACES* scheme, we present in Fig. 9 a simple example with two-nodes sharing a common bus. This picture shows immediately the main advantage of the *SPACES* algorithm. *Bandwidth fragmentation* is significantly alleviated. When *SPACES* is enabled, the first node creates only large idle periods. These idle periods can be exploited by a downstream node to further preserve bandwidth for its downstream partners, transmit a burst of a size of T_{transm} or emit a burst of smaller size.

Unfortunately, however, the phenomenon of *bandwidth fragmentation* cannot be completely avoided. Because the process of transmission remains asynchronous, bandwidth is fragmented when a node on the ring does not exploit the totality of an idle period. Therefore, the mismatch between the fixed-size of reserved idle periods and the variable length of optical frames will always lead with *bandwidth fragmentation*.

The proposed method has two obvious advantages. First, *SPACES* can be a distributed fairness algorithm using transmission-based quota: The average rate of traffic transmitted by a ring node is equal to its sustainable rate specified in Service Level Specification (SLS) (e.g., r_i). Each ring node limits its transmission by setting properly the values of T_{reserv} and T_{transm}. For example, a node i access a ratio of channel capacity described by the following formula:

$$r_i = \frac{T^i_{transm}}{T^i_{transm} + T^i_{reserv}} = \frac{1}{1 + \dfrac{T^i_{reserv}}{T^i_{transm}}} = \frac{1}{1 + a_i} \qquad (2)$$

The flexibility of the scheme permits to strictly limit the average rate of traffic transmission at upstream nodes while creating large idle periods. *SPACES* can also be a statistical and fully distributed fairness algorithm. In this case, the average rate of traffic transmitted by each ring node is no more limited by the sustainable rate specified in SLS. This is a significant difference of *SPACES* with respect to TB and TCARD respectively. For example, if the first node uses only a small portion of available bandwidth, the second node can benefit from the unused part of the first node. The third node can access the unused bandwidth of the previous two nodes and so on. Conversely, TB and TCARD limit the transmission of each ring node to its sustainable bit rate independent of the current state of the resources usage.

The statistical version of the *SPACES* scheme must deal with the problem of finding an optimum solution for setting of the values of T^i_{reserv} and T^i_{transm} respectively. When $a_i < 1$ (i.e., $T^i_{reserv} < T^i_{transm}$) , a node grabs more bandwidth than it reserves for its downstream partners. This situation leads to unfairness with the regard to access bandwidth at downstream nodes. When $a_i > 1$, bandwidth is wasted if upstream nodes reserve a ratio of bus capacity larger than the load of the traffic to be transmitted at downstream nodes. The case $a_i = 1$ leads to what we call *local fairness* with the regard to downstream partners: A ring node transmits T^i_{transm} time units and, afterwards, stops its transmission for an identical period of time. Notice that at channel fully loaded, the upstream nodes create a series of identical busy and free periods.

We highlight that finding an optimum solution for setting of values of T^i_{reserv} and T^i_{transm} respectively is not an easy task. However, when *local fairness* version of the *SPACES* algorithm is used, the network performance increases significantly, as we shall show in the next section.

6. Performance Evaluation of SPACES

In this section, we evaluate the performance of the *SPACES* scheme and compare it with the one achieved by TCARD. We notice that we keep similar assumptions used within the simulation works presented earlier. For SPACES, we consider only the case of *local fairness* (i.e. $a = 1$) for a main reason: The parameters setting of *SPACES* is independent of the physical location of a node within a network. The parameters setting in TCARD is as defined in the previous section.

We first compare the performance of each ring node when the *SPACES* algorithm is enabled, when the TCARD algorithm is enabled, and when both *SPACES* and TCARD algorithms are disabled. Each node receives traffic from the access networks to be forwarded towards the hub at a mean rate of 1 Gb/s. Thus, the upstream bus is loaded at 80 percent of its capacity.

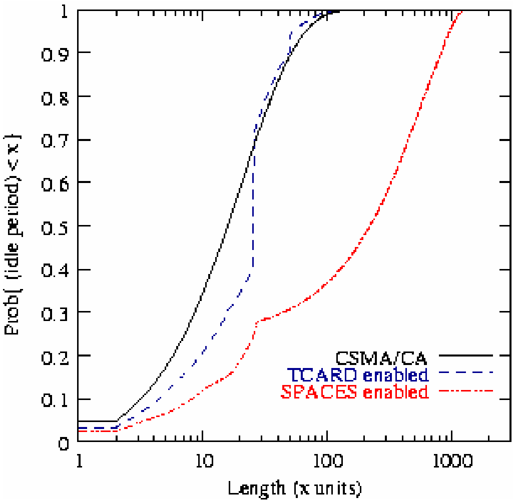

Fig. 10: Bandwidth fragmentation as "seen" by the fourth node: SPACES vs. TCARD & CSMA/CA

Fig. 10 depicts the interaction between the aforementioned schemes and the phenomenon of *bandwidth fragmentation*. We assume $T^i_{reserv} = T^i_{transm} = 50$ µs, $i = 1 \ldots N$. The results confirm the efficiency of the *SPACES* algorithm: We observe that the ratio of large-size idle periods increases significantly.

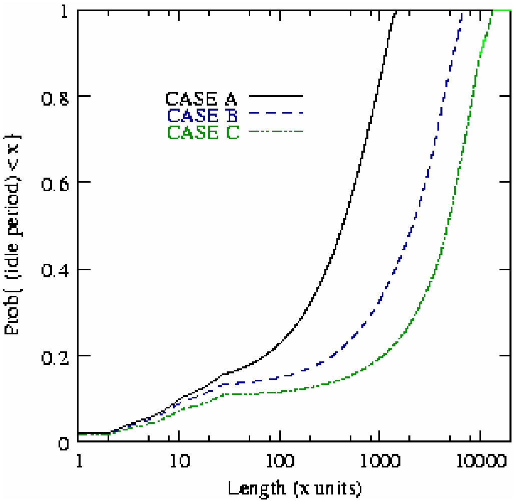

Fig. 11: Interaction between the SPACES scheme and *bandwidth fragmentation*

The probability of "sensing" large free periods at end of the bus increases as the size of the reservation timer is increasing. To get a more detailed insight into the correlation of the size of reservation and transmission timers and the protocol performance, on the other hand, we consider three cases: A) $T^{i}_{reserv} = T^{i}_{transm} = 50$ µs, B) $T^{i}_{reserv} = T^{i}_{transm} = 250$ µs, and C) $T^{i}_{reserv} = T^{i}_{transm} = 500$ µs, $i = 1 \ldots N$.

The curves plotted in Fig. 11 show clearly that an increase in reservation timer leads to an increase in length of voids "sensed" at downstream stations. Here, we present the distribution of the size of idle periods, as "seen" at the transmission tap of the fourth node.

SPACES adds an additional delay of T_{reserv} time units at upstream queueing times. Results, in terms of average access delay, depicted in Fig. 12 show that the fairness issue is alleviated at downstream nodes at expense of media access delay increasing at upstream nodes. We observe that in all three cases the relative performance between the ring nodes remains fair. It is clear that the efficiency of the *SPACES* algorithm in achieving fairness is independent of its parameters setting. We notice that no data loss is observed at ring nodes when the *SPACES* algorithm is enabled. We point out, however, that larger the reservation window larger access delays. The delay recorded for upstream nodes is increased, compared to the case where no fairness control mechanism is applied or TCARD is enabled, but still remains in acceptable limits.

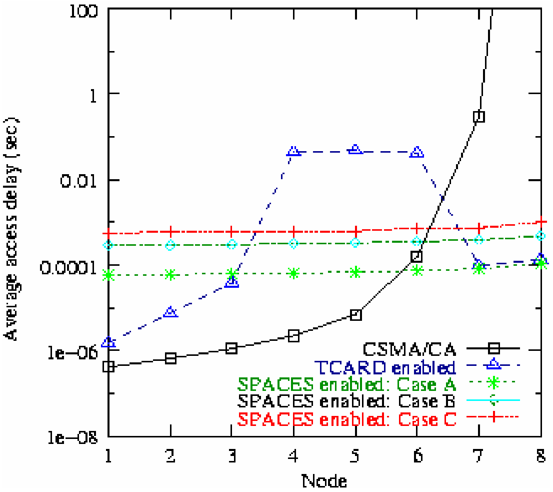

Fig. 12: Access delay as a function of position on bus: SPACES versus TCARD & CSMA/CA

In Fig. 13 and 14, we show *SPACES* behavior, in terms of average access delay and node throughput respectively, when the bus is heavily loaded (i.e., $\rho = 0.9$ and 1.0). We observe that the *SPACES* algorithm may be unfair in the sense that the senders situated at the end of the bus are saturated when the system is 100 percent loaded. This is due to the phenomenon of *bandwidth fragmentation* and the low efficiency of transporting electronic packets at optical/physical level, as it was explained in Section 3. For loads less than 1.0, the efficiency of the *SPACES* algorithm is evident: The system remains stable up to a load of $\rho = 0.9$. For $\rho = 0.9$, we notice a light performance degradation at the last node on the upstream bus (i.e., eight node).

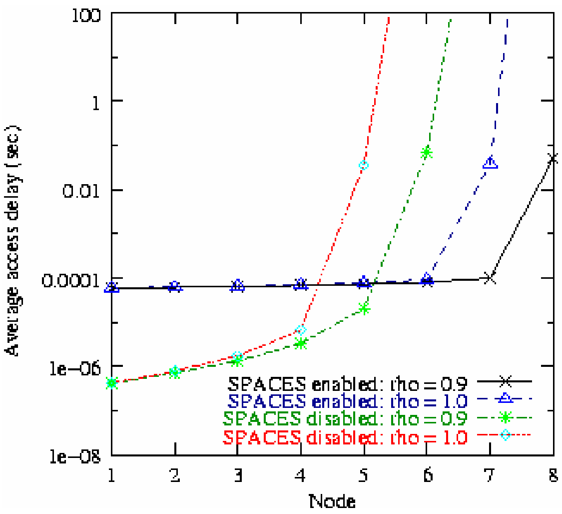

Fig. 13: Access delay for heavily loaded channel

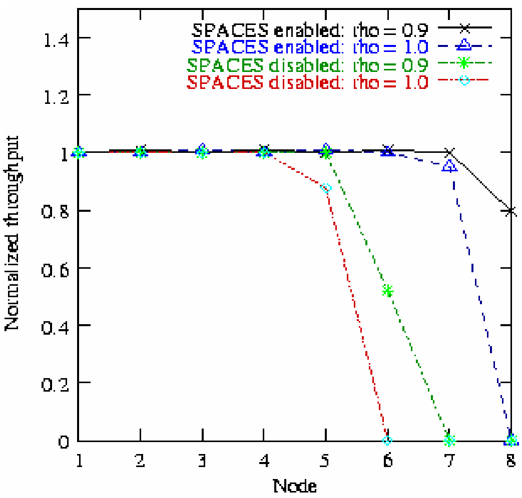

Fig. 14: Node throughput for heavily loaded channel

Of course, the system performance is nothing but an answer to a dedicated load. Therefore, results would differ from the curves shown in this paper if different arrival process would be assumed. However, our goal was to prove that the distributed version of the *SPACES* algorithm avoids efficiently the phenomenon of *bandwidth fragmentation* while achieving fairness among ring nodes. In addition, simulation results depicted in Fig. 15 show that *SPACES* improves significantly the resource utilization: The network remains stable for loads up to 0.8. Thus, the ratio of bandwidth utilization increases from 60 to 80 percent.

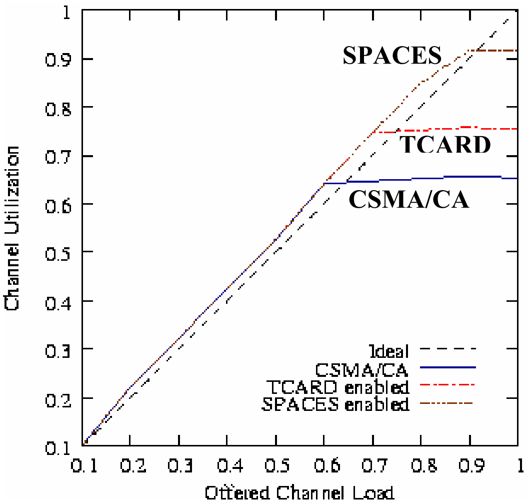

Fig. 15: Channel utilization versus offered load (ρ)

7. Conclusions

In this paper, we have studied the fairness issue of the asynchronous optical CSMA/CA protocol with variable packet sizes. For light-to-moderate load, the access protocol provides good performance while achieving fairness among ring nodes. The fairness discussion for heavy load has shown that the unfairness, in regard with the access of the shared bandwidth, is pronounced between upstream and downstream nodes.

The asynchronous CSMA/CA scheme suffers from *bandwidth fragmentation* and *positional priority* fairness-related problems. First, we proved the former represents the most important issue, with the regard to the access protocol performance, because it increases significantly the unfairness for downstream nodes. Next, we explained why the traditional fairness algorithms fail in achieving satisfactory performance in such shared-bus configuration. Finally, we exhibited the limitations of a previous proposed fairness algorithm, the so-called TCARD. For moderate-to-heavy load, we observed that TCARD cannot deal efficiently with the fairness issue in DBORN.

Consequently, we considered an additional flow control mechanism not only to limit the node transmission but also to regulate the emission process on shared bandwidth. We presented a statistical fairness algorithm, called *SPACES*, which contains a major advantage over TCARD and traditional fairness algorithms: It is fully distributed. It means that the parameters setting of *SPACES* is independent of the physical location of a node within the network. Simulation results showed that the proposed solution alleviates substantially the performance degradation at heavy bus loads while achieving fairness among ring nodes.

The results on which the comparison was based provide many useful insights. For moderate-to-high load, the phenomenon of *bandwidth fragmentation* can be substantially alleviated if the upstream nodes create large fixed-size idle periods. However, the mismatch between the fixed-size of the reserved voids and the variable-length of frames will always lead to *bandwidth fragmentation* in asynchronous CSMA/CA scheme. This means that little can be done to achieve throughput and access delay fairness when the traffic load is heavy.

There are other practical aspects that have to be studied. In this paper, we have not addressed the ability of the *SPACES* scheme to deal with traffic with different QoS requirements. As described in this article, *SPACES* uses two timers (i.e., reservation and transmission timer respectively) to regulate the process of transmission. The reservation timer is a key parameter for the size of idle periods "sensed" at downstream nodes, and hence for downstream access delays, as well as for access delays at upstream nodes. A "target" access delay is strictly guaranteed at upstream nodes and only "statistically" achieved at downstream nodes. However, it is expected that queueing times at downstream nodes will decrease as the size of the spectral dimension (i.e., the number of upstream wavelengths) is increasing.

References

[1] J. Tomasik and D. Popa, "On Markov Chain Modelling for Asynchronous Optical CSMA/CA Protocol", submitted to IEEE Communications Letters.

[2] N. Bouabdallah, A.-L. Beylot, E. Dotaro and G. Pujolle, "Resolving The Fairness Issue in Bus-Based Optical Access Networks", IEEE JSAC, vol. 23, no. 8, pp. 1444-57, Aug. 2005.

[3] A. Carena, V. De Feo, J.M. Gaudino, R. Neri, C. Piglione and P. Poggiolini, "RINGO: An Experimental WDM Optical Packet Network for Metro Applications", IEEE JSAC, vol. 22, no. 8, pp. 1561-71, Oct. 2004.

[4] F. Davik *et al.*, "IEEE 802.17 Resilient Packet Ring Tutorial", IEEE Commun. Mag", vol. 42, no. 3, pp. 112-18, Mar. 2004.

[5] I.M. White, M.S. Rogge, K. Shrikhande and L.G. Kazovsky, "A summary of the HORNET project: A next-generation metropolitan area network", IEEE JSAC, vol. 21, no. 9, pp. 1478-94, Nov. 2003.

[6] L.Ciavaglia, N.Bouabdallah, E. Dotaro and N. Le Sauze, “Matching fairness and performance by preventive traffic control in optical multiple access networks”, in Proc. Of OptiComm2003, pp. 424-429, Oct. 13-17 2003, Dallas, USA

[7] E. Dotaro *et al.*, “DBORN: A Shared WDM Ethernet Bus Architecture for Optical Packet Metropolitan Network”, in Proceedings of Photonics in Switching PS'02, July 2002.

[8] “IP Packet Length Distributions”, http://www.caida.org/analysis/AIX/plein_hist, Juin 2002.

[9] “Network Simulator", http://www.isi.edu/nsnam/ns, 2002.

[10] N. Le Sauze *et al.*, “A Novel, Low Cost Optical Packet Metropolitan Ring Architecture”, in Proceedings of ECOC’01, pp. 66-67, 30 Sept. - 4 Oct. 2001.

[11] R.H. Cardwell, O.J. Wasem and H. Kobrinski, “WDM Architecture and Economics in Metropolitan Area”, SPIE Optical Network Mag., July 2000.

[12] M.J. O'Mahony, “The Application of Optical Packet Switching in Future Communication Networks”, IEEE Comm. Mag. pp. 128-25, Mar. 2001.

[13] G.C. Watson and S. Tohme, “S++ - A New MAC protocol for Gb/s Local Area Networks”, IEEE JSAC, pp. 531-539, vol.41, no. 4, May 1993.

[14] I. Cidon and Y. Ofek, “MetaRing - A Full-Duplex Ring with Fairness and Spatial Reuse”, IEEE Trans. on Comm., vol. 41, no. 1, pp. 110-120, Jan. 1993.

[15] T. Okada, H. Ohnishi and N. Morita, “Traffic Control in Asynchronous Transfer Mode”, IEEE Comm. Mag., pp. 58-62, vol. 29, no. 9, Sept. 1991.

[16] "Distributed Queue Dual Bus (DOQB) Subnetwork of a Metropolitan Area Network", IEEE 802.6 WG, Ed. P802.6/D15, Oct. 1990.

[17] A.A. Lazar, A.T. Temple and R. Gidron, “MAGNET II: A metropolitan are network based on asynchronous time sharing”, IEEE JSAC, vol. 8, no. 8, pp. 1582-94, Oct. 1990.

[18] R.M. Falconer and J.L. Adams, “Orwell: A protocol for an integrated service local network”, British Telecom. Technol. J., vol. 3, pp. 27-35, no. 4, Oct. 1985

Daniel POPA received the M.Sc. degree in telecommunication engineering from the Polytechnic University of Bucharest, Romania, in 2001, and the PhD degree in Computer Science from National Institute of Telecommunications, France, in June 2005. He is currently working in the Optical Network Division of ALCATEL R&I, in France, where he is Research Engineer in the Packet Transport Infrastructure Laboratory.

He was a Visiting Researcher in the Computer Science Department of Academy of Mathematical Sciences, Poland, in July 2002 and June 2003, as well as the Electrical Engineering Department of Istanbul Technical University, Turkey, in October 2004. From June 2005 to August 2006, he held a Research Fellowship in the Computer Science Department of National Institute of Telecommunications, France.

He has done extensive research in areas of performance evaluation, traffic management, Quality of Services and fairness in access networks. His recent research interests include optical MAN/WAN networks and wireless access networks. He was involved in many French national research projects (ROM, ROM-EO, IPSAT), as well as the European Network of Excellence.

Tülin ATMACA received the PhD degree in Computer Science from The University of Paris XI (ORSAY), in France in 1987. From 1986 to 1988 she has taught in the Department of Computer Science at the University of Paris XI. From 1989 to 1991, she was Assistant Professor in the Computer Science Department at the University of Paris VI. Since January 1992, she has been Associate Professor at National Institute of Telecommunications (INT) in France, in the networking and service department.

She was a Visiting Professor in the Computer Science Department of the North Carolina State University, as well as the department of Industrial Engineering of the Rutgers University. Her research interests are in the areas of performance evaluation of telecommunication networks (ATM, FR, Optical Networks, TCP/IP,...), of traffic and congestion control, and Quality of Services aspects in these networks. She is involved in several national and international research projects in the field of the optical packet switching network and their performance (ROM, ROM-EO, EuroNGI, Industrial research projects). She served on the program committees and was co-chair of various well-known international conferences (AICT’06, ICDT’06). Her recent research interests include the all-optical multi-service networks: WAN and MAN and the performance of optical access networks.