MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Synchronous Bandwidth Allocation in FDDI Networks

Qin Zheng, Kang G. Shin

TR93-06 December 1993

Abstract

It is well-known that the FDDI guarantees a bounded access delay and an average bandwidth for synchronous traffic. However, this fact alone cannot effectively support many real-time applications that require the timely delivery of each critical message. We solve this problem by developing a synchronous bandwidth allocation (SBA) scheme which calculates the synchronous bandwidth necessary for each application to satisfy its message-delivery delay requirement. The result obtained in this paper is complementary to the SBA protocol in the FDDI station management standard SMT 7.2, and is essential for effective use of the FDDIś capacity of supporting synchronous traffic.

Mitsubishi Electric Research Laboratories
Cambridge Research Center

Technical Report 93-06 **May 28 1993**

SYNCHRONOUS BANDWIDTH ALLOCATION IN FDDI NETWORKS

by

Qin Zheng[†], and Kang G. Shin[‡]

[†]Mitsubishi Electric Research Laboratories, Inc.
Cambridge Research Center
Cambridge, MA 02139

[‡]Real-Time Computing Laboratory
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan 48109-2122

Abstract

It is well-known that the FDDI guarantees a bounded access delay and an average bandwidth for synchronous traffic. However, this fact alone cannot effectively support many real-time applications that require the timely delivery of each critical message. We solve this problem by developing a synchronous bandwidth allocation (SBA) scheme which calculates the synchronous bandwidth necessary for each application to satisfy its message-delivery delay requirement. The result obtained in this paper is complementary to the SBA protocol in the FDDI station management standard SMT 7.2, and is essential for effective use of the FDDI's capacity of supporting synchronous traffic.

To be presented at *ACM Multimedia'93.*

201 Broadway
Cambridge Massachusetts 02139

Publication History:-

1. First printing, TR 93-06, May 1993

1 Introduction

The Fiber Distributed Data Interface (FDDI) is a proposed ANSI standard for a 100 Mbps token ring network using a fiber-optic medium [1, 2, 3, 4]. Thanks to its high transmission speed, the FDDI alleviates the bandwidth saturation problem of the current 10 Mbps Ethernet and the 4 or 16 Mbps IEEE 802.5 Token Rings. The synchronous transmission capacity of the FDDI also makes it ideal for real-time applications like digital video/audio transmissions.

However, before using the FDDI effectively for real-time applications, one must develop a synchronous bandwidth allocation (SBA) scheme which determines how much of synchronous bandwidth should be allocated for each application. Assigning an application too much bandwidth reduces the network's ability of supporting other real-time traffic, and assigning too little may not satisfy the real-time requirements of the application.

FDDI networks guarantee a bounded access delay and a minimum average bandwidth for synchronous traffic. Specifically, if the target token rotation time of an FDDI network is set to $TTRT$ and the high-priority token holding time of node i is h_i, then the time node i needs to wait for a chance to transmit its synchronous packets is bounded by $2 \times TTRT$, and on the average, it is guaranteed to have a bandwidth of $h_i/TTRT \times 100$ Mbps to transmit its synchronous packets. These two properties make FDDI networks capable of supporting synchronous traffic, but they do not directly yield an SBA scheme.

Allocating an average synchronous bandwidth (i.e., $h_i/TTRT \times 100$ Mbps) equal to the average signal rate is obviously not enough since the peak traffic rate could be much higher than the average traffic rate. A more serious problem is that allocation of a synchronous bandwidth equal to the peak signal rate is still not enough as discussed below. Note that for most real-time applications, each critical message is required to be delivered to its destination within a pre-specified delay bound. Consider real-time video transmissions as an example. Suppose node i wants to establish a real-time video channel with the following features:

- The source of the channel generates a video frame every T units of time. For full motion video, $T = 33$ms.
- The time needed to transmit a maximum-size video frame at a 100 Mbps transmission rate is C_{max}.
- For a smooth real-time video at the destination, each frame is required to be delivered to the destination within d units of time after its generation.

Based on the well-known fact about the worst-case token rotation time of the FDDI, the target token rotation time (TTRT) must be set to no larger than $d/2$ in order to satisfy the frame-delay requirement. Suppose TTRT is set to $d/2$. Assigning a synchronous bandwidth equal to the peak signal rate means $h_i/TTRT = C_{max}/T$, thus resulting in $h_i = C_{max}d/(2T)$. Then, for applications which require $d < 2T$, we get $h_i < C_{max}$. From the MAC protocol of the FDDI, h_i is the maximum time node i is allowed to have for transmitting synchronous packets once it gets the token. Thus, $h_i < C_{max}$ implies that a maximum-size video frame would take more than one token's visit to get transmitted. If the TTRT is set to $d/2$, the token rotation time could be as large as d, so a maximum-size frame would not be transmitted within a delay bound d in the worst-case.

One would then raise a question: "if allocation of a synchronous bandwidth equal to the peak traffic rate still cannot satisfy the requested delay bound, how much of synchronous bandwidth should we allocate for a given application? " The station management standard

SMT 7.2 of the FDDI describes a synchronous bandwidth allocation (SBA) protocol [4], which specifies *how* synchronous bandwidth is allocated to a node, but it does not indicate *how much* of synchronous bandwidth needs to be assigned for a specific application. Clearly, the FDDI's capacity of supporting synchronous traffic cannot be effectively used without a proper SBA scheme.

Agrawal *et al.* [5] proposed a normalized proportional SBA scheme which has the following features and/or problems.

1. The scheme can be used for applications where the requested message-delay bound d always equals the message generation period T. In other words, each synchronous message is required to be delivered to its destination before the generation of the next message.

2. Using this scheme, an FDDI network is proven to be able to support any set of synchronous channels with a total peak signal rate less than 33% of the ring bandwidth. This percentage was claimed to be the highest to date.

3. This scheme is *not* optimal in the sense that a set of synchronous channels which cannot be established with the normalized proportional SBA scheme may be established with some other scheme.

4. It is a *global* SBA scheme in that the allocation/deallocation of synchronous bandwidth to a node would require to change the synchronous bandwidths previously assigned to all other nodes.

The requirement of $d = T$ limits the type of applications that can be supported and the non-optimality of the scheme does not fully utilize the network's ability of accommodating synchronous traffic. Use of a global SBA scheme also complicates the SBA protocol, making it difficult to implement.

As an improvement of the scheme in [5], Chen *et al.* proposed an optimal SBA scheme in a recent paper [6]. However, it still suffers the limitation of $d = T$ and is a global scheme. Besides, it uses an iterative algorithm for the calculation of the optimal bandwidths which may, in theory, need an infinite number of steps to converge.

In this paper we propose an SBA scheme which does not require $d = T$ and is optimal when $d \leq T + TTRT$. The calculation of the optimal bandwidths can be done in just one step. Further, allocation/deallocation of synchronous bandwidth to one node does not require to change the synchronous bandwidths assigned to other nodes, thus making the SBA protocol easy to implement.

This paper is organized as follows. Section 2 reviews the MAC protocol of the FDDI and its relevant properties. A new SBA scheme is proposed in Section 3 and analyzed in Section 4. The paper concludes with Section 5. Due to space limit, proofs of some results are omitted. Readers are referred to [7] for more details.

2 Preliminaries

For convenience of discussion, we review the FDDI's MAC protocol and some of its properties in this section. The FDDI's MAC protocol [3] is summarized below.

Protocol 2.1 .

P1: Suppose there are N active nodes in a ring which are numbered from 0 to $N-1$. As part of the FDDI ring initialization process, each node declares a Target Token Rotation Time (TTRT). The smallest among them is selected as the ring's TTRT. Each node which supports synchronous traffic is then assigned a portion of the TTRT to transmit its synchronous packets. Let h_i, called the *high-priority token holding time*, denote the portion of TTRT that node i is assigned to transmit its synchronous packets.

P2: Each node has two internal timers: the token–rotation–timer (TRT) and the token–holding–timer (THT). The TRT always counts up and a node's THT counts up only when the node is transmitting asynchronous packets. If a node's TRT reaches the TTRT before the token arrives at the node, the TRT is reset to 0 and the token is marked as late by incrementing the node's late count L_c by one. To initialize the timers at different nodes, no packets are allowed to be transmitted during the first token rotation after the ring initialization and L_c's are set to 0.

P3: Only the node that has the token is eligible to transmit packets. The packet-transmission time is controlled by the node's timers, but an in-progress packet transmission will not be interrupted until its completion. When node i receives the token, it does the following:

P3.1: If $L_c > 0$, set $L_c := L_c - 1$ and $THT := TTRT$. Otherwise, set $THT := TRT$ and $TRT := 0$.

P3.2: If node i has synchronous packets to transmit, it transmits them for a time period up to h_i or until all the synchronous packets are transmitted, whichever occurs first.

P3.3: If node i has asynchronous packets to transmit, it transmits them until the THT counts up to the TTRT or all of its asynchronous packets are transmitted, whichever occurs first.

P3.4: Node i passes the token to the next node $(i+1)$ mod N.

Let T_{ring} denote a ring's latency which is the time needed to circulate the token around the ring once without transmitting any packet, and T_p denote the time needed to transmit a maximum-size asynchronous packet. Then, the parameters of the FDDI's MAC protocol must satisfy the following protocol constraint:

$$\sum_{i=0}^{N-1} h_i \leq TTRT - T_{ring} - T_p. \tag{1}$$

The physical meaning of the above inequality is that the summation of the assigned synchronous bandwidths over the nodes in the network should not exceed the effective ring bandwidth. Violation of this constraint would make the ring unstable and oscillate between "claiming" and "operational" [4]. Under this protocol constraint, a well–known fact about the FDDI is that the worst-case token rotation time is bounded by $2 \times TTRT$, and the average token rotation time is bounded by TTRT [8]. A more general result was obtained by Agrawal *et al.* in a recent paper [5] as stated below.

Lemma 2.1 .
Under the protocol constraint (1), the time elapsed between any n consecutive token's visits to a node i is bounded by $n \times TTRT - h_i$.

Once node i gets the token, it is given up to h_i units of time to transmit its synchronous packets. The following lemma gives a lower bound of time that node i is allowed to transmit its synchronous packets during a time period t [5].

Lemma 2.2 .
Under the protocol constraint (1) of the FDDI, node i has at least $\lfloor t/TTRT - 1 \rfloor \cdot h_i$ *units of time to transmit its synchronous packets during a time period t. This lower bound is reached when* $\lceil t/TTRT - 1 \rceil \cdot TTRT - t \geq h_i$.

Lemmas 2.1 and 2.2 are the best published synchronous properties of the FDDI to date. We will improve Lemma 2.2 and derive a new SBA scheme in the next section.

3 A New SBA Scheme

As discussed in the Introduction, an important feature of real-time communication is that each message must be delivered to its destination within a pre-specified delay bound. Due to the limited network transmission bandwidth, this requirement cannot be satisfied without some information on message generation characteristics.

We use two parameters, T and C, to describe a message generation pattern, where T is the *minimum* message inter-generation time and C is the *maximum* message-transmission time (i.e., the time needed to transmit a maximum-size message). It is reasonable to assume prior knowledge of these parameters for many real-time applications, such as interactive voice/video transmission and real-time control/monitoring. For applications where the traffic pattern is less predictable, the estimated values of T and C could be used. A source node may exceed its pre-specified maximum message size and/or message generation rate at the risk that these messages may not be delivered within the pre-specified delay bound, but this particular node will not affect the guarantees of other applications.

Together with the requested message-delivery delay bound d and the address of the source node S, we use the concept of a *real-time channel* [9] for real-time communication. A real-time channel is described by a 4-tuple $\tau = (T, C, d, s)$ and guarantees each message generated at the source node s to be delivered sequentially to one or more destination nodes in a time period $\leq d$, as long as the message inter-generation time is $\geq T$ and the message transmission time is $\leq C$.

A real-time channel is a convenient way to achieve real-time communication. Users can set up channels with adequate bandwidths and delay bounds for their applications. This is in sharp contrast to the conventional circuit-switched transmission where users have few choices on the bandwidth and quality of the circuits. We will in this paper deal with the implementation of real-time channels in FDDI networks only. Readers are referred to [9, 10, 11, 12, 13, 14] for discussions on real-time channels in point–to–point networks.

A set of real-time channels is said to be *establishable* over an FDDI network if the requested message-delivery delay bound of each channel can be guaranteed by properly setting the parameters of the FDDI's MAC protocol. From Protocol 2.1, the user-adjustable parameters are the TTRT and the high-priority token holding time h_i's. The TTRT is usually set at the network initialization time and does not change frequently. It determines the minimum message-delay bound, $d_{min} = 2 \times TTRT$, that the network can guarantee. Any channel request with a delay bound smaller than d_{min} will be rejected. With a given TTRT, the synchronous bandwidth allocated to node i is determined by the value of h_i. Thus, an SBA scheme determines the

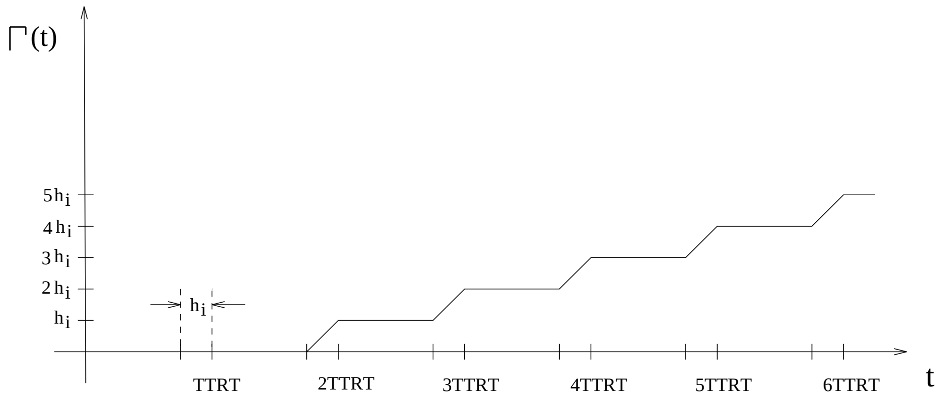

Figure 1: Worst-case synchronous transmission time , (t).

values of h_i's to accommodate real-time channels. An SBA scheme is said to be *feasible* with respect to a set of real-time channels if it can guarantee the requested delay bounds of all the channels. An SBA scheme is said to be *optimal* if it is always feasible whenever there exists a feasible SBA scheme. The advantage of an optimal SBA scheme is the full-utilization of the FDDI's synchronous transmission capacity since a set of real-time channels rejected by an optimal SBA scheme cannot be established with any other SBA schemes.

We derive in this section the conditions for establishing real-time channels over an FDDI network. From these conditions, a new SBA scheme will be developed which has many advantages over the SBA schemes of [5, 6].

Let , (t) denote the time that a node in the worst-case is allowed to transmit its synchronous packets during a time period t. Lemma 2.2 gives a lower bound of , (t) for node i. We improve Lemma 2.2 by calculating the exact value of , (t) as follows.

Lemma 3.1 .
Under the protocol constraint (1) of the FDDI, node i in the worst-case has

$$, (t) = \lfloor t/TTRT - 1 \rfloor h_i + \delta(t)$$

units of time to transmit its synchronous packets during a time period t, where $\delta(t)$ is calculated as $\delta(t) = 0$ if $\lceil t/TTRT \rceil TTRT - t \geq h_i$ or $t \leq TTRT$, and $\delta(t) = t - (\lceil t/TTRT \rceil TTRT - h_i)$ otherwise.

Proof: , (t) is plotted in Fig. 1. Its correctness can be seen from Lemma 2.1. In the worst-case, node i would first wait $2 \times TTRT - h_i$ units of time to get the token. Once it gets the token, it has h_i units of time to transmit its synchronous packets. This proves the correctness of , (t) for $t \leq 2 \times TTRT$. From Lemma 2.1, the following worst-case token inter-arrival time at node i would be $TTRT$. This proves the correctness of , (t) for $t > 2TTRT$.

Suppose no two real-time channels have the same source node and the synchronous transmission time of a node is used for real-time channel messages only. Then from Lemma 3.1, we have the following necessary and sufficient condition for the establishment of a real-time channel over an FDDI network.

Theorem 3.1 .
A real-time channel $\tau = (T, C, d, s)$ can be established over an FDDI network under the protocol constraint (1) if and only if

$$\forall t \geq 0, \quad \lceil (t-d)/T \rceil^{+} C \leq \, , (t), \tag{2}$$

where $, (t)$ *is calculated from Lemma 3.1 with* $i = s$, *and* $\lceil x \rceil^{+} = n$ *if* $n-1 \leq x < n$, $n = 1, 2, \cdots$, *and* $\lceil x \rceil^{+} = 0$ *for* $x < 0$.

Proof of the necessary condition: Suppose node s does not have any message of channel τ at time $t = 0$. Then, $\forall t > 0$ a necessary condition for no messages to miss their deadlines in $[0, t]$ is that the amount of time, $\tau(t)$, needed to transmit all those messages generated during $[0, t]$ by channel τ with deadlines $\leq t$ is not greater than $, (t)$, the time that node s in the worst-case is allowed to transmit its synchronous packets. Since the minimal message inter-generation time of channel τ is T, there are at most $\lceil (t-d)/T \rceil^{+}$ messages generated by channel τ during $[0, t]$ with deadlines $\leq t$, which need at most $\lceil (t-d)/T \rceil^{+} C$ units of time to transmit. Thus, the maximum value of $\tau(t)$ is $\lceil (t-d)/T \rceil^{+} C$. This proves the necessary condition.

Proof of the sufficient condition: We prove this by contradiction. Suppose a message misses its deadline at time t_1, meaning that at least one message with deadline $\leq t_1$ has *not* been transmitted by t_1. Then there must exist $t' < t_1$ such that during the time period $[t', t_1]$, node i uses *all* of its allowed synchronous transmission time for channel τ's packets. Let t_0 be the smallest such t', then there are no messages with deadlines $\leq t_1$ queued at the link at time t_0^-. Thus, in the time period $[t_0, t_1]$, node i uses all its synchronous transmission time transmitting only those packets of channel τ which are generated during $[t_0, t_1]$ with deadlines $\leq t_1$. Based on the same reasoning as the proof of the necessary condition, the maximum amount of time needed to transmit these messages is $\tau(t_1 - t_0) = \sum_{i=1}^{n} \lceil (t_1 - t_0 - d)/T \rceil^{+} C$. Since one message misses its deadline at t_1, this $\tau(t_1 - t_0)$ must be larger than $, (t_1 - t_0)$, that is,

$$\lceil (t_1 - t_0 - d)/T \rceil^{+} C > \, , (t_1 - t_0).$$

By letting $t = t_1 - t_0$, the above inequality contradicts the condition that $\forall t \geq 0, \quad \lceil (t - d)/T \rceil^{+} C \leq \, , (t)$.

Since the left-hand side of Eq. (2) changes only at points $t = d_i + kT$ with the value $\lceil ((d_i + kT) - d_i)/T \rceil^{+} C = (k+1)C$, we have the following corollary from Theorem 3.1.

Corollary 3.1 .
A real-time channel $\tau = (T, C, d, s)$ can be established over an FDDI ring under the protocol constraint (1) if and only if h_s is set such that

$$h_s \geq \alpha_s C \tag{3}$$

where $\alpha_s = \max\{(k+1)/(, (d + kT)/h_s) : \; k = 0, 1, \cdots\}$.

Then, we have the following SBA scheme for the establishment of a real-time channel.

Algorithm 3.1 .
Suppose $n-1$ real-time channels $\tau_i = (T_i, C_i, d_i, s_i)$, $i = 1, \cdots, n-1$ have already been established over an FDDI ring. Then a new channel $\tau_n = (T_n, C_n, d_n, s_n)$ can be established with the following steps.

Step 1: *Calculate* α_n *from Corollary 3.1 and* $h_{s_n} = \alpha_n C_n$.

Step 2: *If the protocol constraint (1) is satisfied, set the high-priority token holding time of* s_n *to be* h_{s_n} *and establish channel* τ_n. *Otherwise, the channel establishment request is rejected.*

Algorithm 3.1 gives an optimal SBA scheme since it uses the sufficient and necessary channel establishment condition of Corollary 3.1. In other words, if a real-time channel cannot be established with Algorithm 3.1, so cannot with any other SBA schemes. However, one problem with Algorithm 3.1 is the calculation of α_n. The definition of α_n in Corollary 3.1 is not given in closed-form. Thus, we need the following theorem for the calculation of α_n.

Theorem 3.2 .
Let $x = d_n/TTRT - 1$ *and* $y = T_n/TTRT$. *Then,*

$$\alpha_n = \begin{cases} 1/\lfloor x \rfloor & \text{if } y \geq \lfloor x \rfloor \geq 1 \\ 1/y & \text{if } y \leq 1 \text{ and } x \geq 2 \end{cases}$$

and

$$\alpha_n \leq \begin{cases} 1 + (2-x)/y & \text{if } y \leq 1 \text{ and } 1 \leq x < 2 \\ 1/\lfloor y \rfloor & \text{if } 1 < y < \lfloor x \rfloor. \end{cases}$$

The values of α_n in different regions of the x–y plane are plotted in Fig. 2. We need not consider the case of $x \leq 1$ since it means $d_n < 2 \times TTRT = d_{min}$ and the channel cannot be established. In most cases, the inequality $d_n \leq T_n + TTRT$ is satisfied, meaning that $y > \lfloor x \rfloor$. So, the exact value $\alpha_n = 1/\lfloor x \rfloor = 1/\lfloor d_n/TTRT - 1 \rfloor$ can be obtained and an optimal SBA scheme is realized via Algorithm 3.1. For regions on the x–y plane where the exact value of α_n cannot be obtained, one can use an upper bound of α_n instead, with little loss of accuracy, because the difference between the upper bound and the actual value of α_n is always smaller than 1.

As an application example of Algorithm 3.1, we calculate the synchronous bandwidth needed for establishing a video channel in an FDDI network. Suppose the video frame inter-generation period $T = 32$ ms, the frame-transmission time is C ms, and the requested frame delay bound $d = \lambda T$. Also suppose the network's TTRT is set to a typical value $TTRT = 8$ ms.

Since d must be no smaller than $2 \times TTRT$, the above video channel cannot be established if $\lambda < 1/2$. For $\lambda \geq 1/2$, α is calculated from Theorem 3.2:

$$\alpha = \begin{cases} 1/\lfloor 4\lambda - 1 \rfloor & \text{if } 1/2 \leq \lambda < 5/4 \\ 1/4 & \text{if } \lambda \geq 5/4 \end{cases}$$

Then, using Algorithm 3.1, the high-priority token holding time at the source node s of the video channel should be set to $h_s := \alpha C$, and the video channel can be established if the protocol constraint (1) is satisfied.

As discussed in Section 1, the synchronous bandwidth assigned to the video channel equals $B_c = (h_s/TTRT) \times 100$ Mbps. After normalizing with the signal rate of the channel $B_s = (C/T) \times 100$ Mbps, we have

$$B_c/B_s = \begin{cases} 4/\lfloor 4\lambda - 1 \rfloor & \text{if } 1/2 \leq \lambda < 5/4 \\ 1 & \text{if } \lambda \geq 5/4. \end{cases}$$

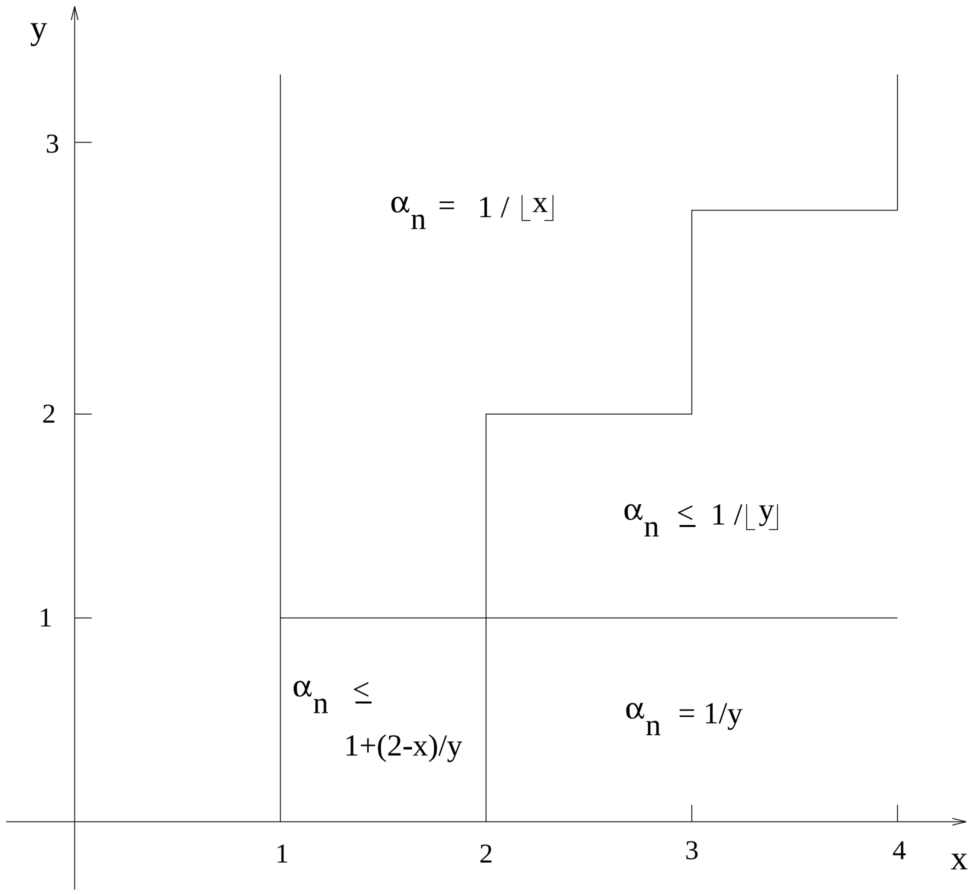

Figure 2: Calculation of α_n.

From this result, one can see that the video channel needs four times as much synchronous bandwidth as its signal bandwidth in order to guarantee a frame delay bound $d = T/2 = 16$ ms. However, if the delay bound can be relaxed to $d = 5T/4 = 40$ ms, we need only the signal bandwidth to establish the channel. Further increasing of d will not help reduce the required bandwidth since the assigned synchronous bandwidth must be at least as large as the signal bandwidth.

From the above example, we can also see that the requested frame-delay bound has a significant impact on the amount of synchronous bandwidth needed to establish a video channel. Users should try their best to avoid using small delay bounds. However, this is not always possible for interactive video applications and/or in cases where video frames have to traverse several LANs to reach their destinations.

Compared with the SBA schemes of [5] and [6], Algorithm 3.1 has the following advantages.

Generality: The SBA schemes of [5, 6] can establish real-time channels with $d = T$ only, while Algorithm 3.1 can establish channels of arbitrary parameters, i.e., $d \leq T$ or $d > T$. This extension is very important in practice since for many applications, especially those in real-time control/monitoring systems, the required delay bound d is usually smaller than the message inter-generation period T. Real-time channels with $d > T$ are also useful for multimedia applications. Thus restricting $d = T$ would greatly limit a network's ability and effectiveness of supporting real-time communications.

Optimality: The SBA scheme of [5] is not optimal, even under the restrictive assumption $d = T$. Thus a real-time channel establishment request may be rejected even if it can be established using another scheme. The SBA scheme of [6] is optimal under the restrictive assumption of $d = T$ and requires complex computations. By contrast, Algorithm 3.1 is optimal for $d \leq T + TTRT$ (which subsumes the special case $d = T$ of [5, 6]) as well as for some other cases when the exact value of α_n can be calculated (see Algorithm 3.2), because it is based on the necessary and sufficient conditions of Theorem 3.1, and the computation of the optimal bandwidths is simple and straightforward. Rejection of a channel establishment request by Algorithm 3.1 means the violation of the necessary conditions, implying that the channel cannot be established with any other scheme.

Simplicity: The SBA schemes of [5, 6] are *global* schemes in the sense that the addition/removal of a channel or change of the parameters of a channel would require adjustment of the high-priority token holding times of all nodes in the network. This requires a complex SBA protocol. By contrast, Algorithm 3.1 needs only *local* parameter adjustment, thereby making it far easier to implement than those in [5, 6].

Algorithm 3.1 can also be used for real-time channels with a common source node. Specifically, if two channels τ_1 and τ_2 have the same source node s, and τ_1 requires $h_s = t_1$ and τ_2 requires $h_s = t_2$. Then setting $h_s := t_1 + t_2$ will satisfy the requirements of both channels provided there is a mechanism at the source node to regulate the transmission times of the packets of τ_1 and τ_2 so as not to exceed t_1 and t_2 during each token's visit, respectively.

4 Analysis

In this section, we analyze the SBA scheme derived in the last section. Specifically, we want to calculate:

1. *At least* how much of synchronous traffic can be supported in an FDDI network using the proposed SBA scheme?

2. *At most* how much of synchronous traffic can be supported in an FDDI network using the proposed SBA scheme?

The answer to the first question gives a "safe" region of the FDDI's synchronous capacity. The network is guaranteed to support any synchronous traffic within this region. The answer to the second question gives an upper bound of the synchronous traffic that a network can accommodate with the proposed SBA scheme. Since this SBA scheme is optimal in most cases, answers to the above two questions give a useful measure of the FDDI's ability of accommodating synchronous traffic.

We now state the above questions more precisely as follows. Given a set of real-time channels $\tau_i = (T_i, C_i, d_i, s_i)$, $i = 1, \cdots, n$, the network utilization by these channels is defined as $U = \sum_{i=1}^{n} C_i/T_i$. U_w is said to be the *worst-case achievable utilization* of a network if it is the largest value such that the network can accommodate every set of real-time channels with utilization $\leq U_w$. The *best-case achievable utilization* U_b is defined as the highest utilization of a set of real-time channels that a network can support. Our problem is then to calculate U_w and U_b of an FDDI network using the SBA scheme given by Algorithm 3.1.

The following theorem calculates U_w.

Theorem 4.1 .
Let $\lambda = \min\{d_i/T_i : \ i = 1, \cdots, n\}$*. Under the condition that* $2 \times TTRT \leq d_i \leq T_i + TTRT$ *and ignoring* T_{ring} *and* T_p*, the worst-case achievable utilization of an FDDI network using the proposed SBA scheme is* $U_w = \lambda/3$*.*

Since under the condition that $d_i \leq T_i + TTRT$ the proposed SBA scheme is optimal, the U_w given in Theorem 4.1 is the worst-case achievable utilization of an FDDI network. In other words, no other SBA scheme can guarantee the establishment of a set of real-time channels with utilization $> U_w = \lambda/3$. Agrawal *et al.* [5] proved that their normalized proportional SBA scheme has a worst-case achievable utilization of 33% when $d_i = T_i$. Thus, their scheme, albeit not optimal, reaches the highest worst-case achievable utilization when $d_i = T_i$.

Since increasing d_i will not affect the establishment of a real-time channel, an FDDI network guarantees the successful establishment of any set of real-time channels with utilization $< 33\%$ for $d_i \geq T_i$. U_w decreases linearly with the decrease of $\lambda < 1$. This means that the smaller the requested delay bounds, the more difficult to establish the real-time channels.

The following theorem calculates U_b.

Theorem 4.2 .
Under the condition that $2 \times TTRT \leq d_i \leq T_i + TTRT$ *and ignoring* T_{ring} *and* T_p*, the best-case achievable utilization of an FDDI network using the proposed SBA scheme is* $U_b = \max\{\lfloor d_i/TTRT - 1 \rfloor / (T_i/TTRT) : \ i = 1, \cdots, n\}$*.*

To see how restrictive U_b is, let us consider a special case when all channels are identical with $d_i = T_i = T, i = 1, \cdots, n$ and $TTRT = T/2$. Then, from Theorem 4.2, we get $U_b = 1/2$, meaning that in this case an FDDI network can use *at most* one half of its transmission bandwidth for real-time channels.

From Theorems 4.1 and 4.2, one can see that the FDDI's MAC protocol is not very efficient in supporting real-time communication. The readers are referred to [15, 7] for a simple modification to the MAC protocol which can significantly improve FDDI's ability of supporting real-time traffic.

5 Conclusion

This paper addresses the problem of allocating synchronous bandwidths in FDDI networks. We developed a general, optimal, and simple SBA scheme that can support a large variety of real-time applications, can fully utilize the network-transmission bandwidth, and is easy to implement. We also analyzed the FDDI's capacity of supporting synchronous traffic using the proposed SBA scheme.

References

[1] F. E. Ross, "An overview of FDDI: The fiber distributed data interface," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 7, pp. 1043 – 1051, September 1989.

[2] A. C. Weaver, "Local area networks and busses – an analysis," Technical report, Flight Data Systems, NASA–Johnson Space Center, 1986.

[3] *Fiber Distributed Data Interface (FDDI) – Token Ring Media Access Control (MAC).* American National Standard, ANSI X3.139, 1987.

[4] *FDDI Station Management (SMT) – draft proposed.* American National Standard, ANSI X3T9/92-067, June 25, 1992.

[5] G. Agrawal, B. Chen, W. Zhao, and S. Davari, "Guaranteeing synchronous message deadlines with the timed token protocol," in *Proceedings of IEEE International Conference on Distributed Computing Systems*. IEEE, June 1992.

[6] B. Chen, G. Agrawal, and W. Zhao, "Optimal synchronous capacity allocation for hard real–time communications with the timed token protocol," in *Proc. Real-Time Systems Symposium*. IEEE, December 1992.

[7] Q. Zheng, *Real-time Fault-tolerant Communication in Computer Networks*, PhD thesis, University of Michigan, 1993. PostScript version of the thesis is available via anonymous FTP from ftp.eecs.umich.edu in directory outgoing/zheng.

[8] K. C. Sevcik and M. J. Johnson, "Cycle time properties of the FDDI token ring protocol," *Technical Report CSRI-179, Computer Science Research Institute, University of Toronto*, 1986.

[9] D. Ferrari and D. C. Verma, "A scheme for real-time channel establishment in wide-area networks," *IEEE Journal on Selected Areas in Communications*, vol. SAC-8, no. 3, pp. 368–379, April 1990.

[10] D. D. Kandlur, K. G. Shin, and D. Ferrari, "Real-time communication in multi-hop networks," in *Proc. 11th Int. Conf. on Distributed Computer Systems*, pp. 300–307. IEEE, May 1991.

[11] Q. Zheng and K. G. Shin, "On the ability of establishing real–time channels in point–to–point packet–switched networks," *IEEE Transactions on Communication* (in press), 1993.

[12] Q. Zheng and K. G. Shin, "Real–time communication in local area ring networks," in *Conference on Local Computer Networks*, pp. 416–425, September 1992.

[13] Q. Zheng and K. G. Shin, "Fault–tolerant real–time communication in distributed computing systems," in *in* Proc. 22nd Annual International Symposium on Fault–tolerant Computing, pp. 86 – 93, 1992.

[14] Q. Zheng, K. G. Shin, and E. Abram-Profeta, "Transmission of compressed digital motion video over computer networks," in *Digest of COMPCON Spring'93*, pp. 37–46, February 1993.

[15] K. G. Shin and Q. Zheng, "FDDI-M: a scheme to double FDDI's ability of supporting synchronous traffic," submitted for publication, 1992.