PERFORMANCE OF MULTIPRIORITY-CYCLE AND TIMED-TOKEN RING PROTOCOLS WITH FULL-DUPLEX VOICE TRANSMISSION

D. Karvelas and A. Leon-Garcia

Department of Electrical Engineering
University of Toronto
Toronto, Ontario Canada M5S 1A4

ABSTRACT

In this paper we investigate the voice/data performance of two token ring protocols that can provide priority services to different types of traffic as well as guarantee bounded delays for real-time applications. The constraints that the system parameters must obey to achieve bounded delays for voice packets are derived under Full-Duplex voice transmission. Simulation is used to investigate the effect of various system parameters on performance.

1. Introduction

In the Integrated Services environment, networks must be capable of providing priority services to accommodate different types of traffic and to guarantee bounded delays for real-time applications. Token ring networks fit very well in this environment because they can fulfill these requirements.

In this paper we consider two types of Media Access Control (MAC) mechanisms that can provide these services. The first of these protocols has a centralized character and it uses different token cycles to provide access to designated traffic classes [1]. We have used the name Multiple Priority Cycles (MPC) protocol [2] for this method. The second has a distributed character and uses a target token rotation time to determine access to the ring. The name Timed Token Protocol (TTP) [3] has been used for this access method. The TTP protocol is the proposed MAC protocol for the Fiber Distributed Data Interface (FDDI) [4].

2. Access Mechanisms

We assume that voice and data sources are connected to the ring. Voice and data sources can be connected to the same or different stations. Data traffic is transmitted in fixed size packets. These can be independent messages or segments of longer messages. Voice traffic consists of packets which are formed by packetization of the talkspurt periods of the voice signal.

During the talkspurt period, packets arrive deterministically, one packet every packetization period \( P_y \). The access schemes provide a mechanism which guarantees that every packet will be transmitted before the next one is generated. In this way the maximum delay (waiting time + transmission time) of a voice packet at the transmitter is one packetization period \( P_y \).

We now present the two priority access mechanisms.

In the MPC protocol the operation of the network alternates between the Synchronous Mode and the Asynchronous Mode. Voice packets are transmitted during the Synchronous mode; data packets during the Asynchronous. This operation justifies the "Multiple Priority Cycle" (MPC) name given to this protocol. The transition from one mode to the other is decided by one station (station 1 in our case) called Synchronous Bandwidth Manager (SBM). The SBM decides to switch to Synchronous mode \( T_{MPC} \) msec after the last Asynchronous to Synchronous transition. The value of \( T_{MPC} \) parameter is chosen to guarantee the delay bound for the voice packets.

In the Timed Token Protocol (TTP), the operation has a distributed character and there is no central controller. Each station keeps the values of two parameters. The value of the first parameter \( T_y \) shows the maximum allowed transmission time for Synchronous traffic (voice packets) by station \( i \) per token arrival. The second parameter \( T_{TTP} \) has a common value to all stations and controls the available time for data

In section 2 we briefly present the two access mechanisms and we derive constraints that the system parameters must satisfy in order for the delay bounds to be met. In section 3 we present the modeling assumptions. In section 4 we present and discuss performance results and finally in section 5 we present conclusions.
transmission in such a way that the delay bound of the voice packets is guaranteed. Every time the token arrives at a station, voice packets are transmitted first until either the voice buffer becomes empty or the \( T_v \) time expires. Data packet transmission is initiated by a station only if the immediately preceding token rotation was less than \( T_{TTP} \) in duration. If \( C \) is the time of a complete token rotation which started and ended at the same station, then \( T_{TTP} - C \) is the maximum allowed time for data transmissions from this station.

### 2.1. Bounded Delay Constraints

In this section we derive the relation between the values of the operational time parameters \( T_{MPC}, T_{TTP} \) and the values of the required bound \( P_v \) for the delay of a voice packet at its transmitter, number of voice sources \( V_v \), voice packet transmission time \( h_v \), data packet transmission time \( h_d \), and walk time \( u_0 \). Walk time is the time the token takes to cycle around the ring when no station transmits. In [2] we derived these constraints under Half-Duplex (H-D) operation for voice traffic. We now compute these constraints under Full-Duplex (F-D) operation. We look at the maximum waiting time that a tagged voice packet can encounter, and we bound this time by \( P_v - h_v \) in order to guarantee that the last bit of the current voice packet has been transmitted before the next ones arrives.

We assume that both speakers, A and B, taking parts to a conversation are on the same ring, and that there is at most one active speaker per station. In this way the number of voice sources \( V_v \) is equal to the number of voice stations \( N_v \). (This implies that for the TTP protocol the value of \( T_v \) is equal to \( h_v \)).

The operation of the protocols in F-D case is similar to their operation in H-D case. The only difference is the behavior of the destination station when a voice packet arrives. This station not only copies the incoming voice packet, but it also replaces it with its own packet if it has one. In this way when this packet returns to the source station, it contains the packet from the destination station.

It is clear that if both talkspurts and silent periods are packetized and transmitted, significant gains can be obtained by the use of F-D operation. For instance twice as many speakers can be supported by the network. However if only talkspurts are transmitted, as it is the case we examine here, the advantage of F-D operation is not so obvious. In this case, although it is possible for both partners A and B to speak simultaneously, most of the time when A speaks, B listens. Hence when a voice packet arrives from the station of A to the station of B, the voice buffer of B is empty in most cases and the savings in bandwidth from the use of F-D operation are rather minor. It is thus expected that the number of voice stations cannot increase significantly.

The data performance however is also affected, for both protocols, by the values of the operational time parameters \( T_{MPC} \) and \( T_{TTP} \). Larger values of these parameters imply longer Asynchronous modes for MPC protocol, and more data transmissions per cycle for TTP protocol, i.e. improved data performance. Therefore it is desirable to operate with the highest possible values of \( T_{MPC} \) and \( T_{TTP} \) that guarantee the required bound for the voice packets. In the sequel we show how the \( T_{MPC} \) and \( T_{TTP} \) values are affected by F-D operation. We first compute the maximum number of voice stations that the network can support.

### Maximum number of voice stations

We consider a ring with voice stations only, and compute for this case the maximum number of stations that can be supported with Full-Duplex operation. Under these conditions TTP and MPC protocols become similar to a simple token passing ring protocol without any priority mechanism.

There are two possibilities with respect to the maximum number of voice packets that can be transmitted during a cycle: a) \( N_v/2 \) packets, b) \( N_v \) packets.

The first case occurs when a station B is not allowed to transmit a voice packet when it receives the free token if: a) In the same cycle B had a voice packet transmission opportunity when a voice packet from its partner A arrived, b) B did not use this opportunity because its own packet arrived after the first bits of the packet from A had passed by. We have used the name Voice Packet Postponed Transmission (VPPT) for this procedure.

The second case occurs when the voice stations are allowed to transmit voice packets every time they receive a free token. We called this procedure Regular Voice Packet Transmission (RVPT).

#### VPPT case

In Fig.1, \( T_1 \) and \( T_3 \) are the time instants when two successive cycles (with respect to station 1) start. From Fig.1 it can be seen that the largest voice packet delay occurs when the tagged packet arrives as close as possible to \( T_1 \) and it is transmitted as far as possible from \( T_3 \). This case appears when:

a) The tagged packet belongs to the last voice station (\( N_v \)) on the ring with talking partner in the first station on the ring.

b) The maximum allowed number \( (N_v/2) \) of voice packets is transmitted during the second
cycle and the tagged packet is the last one among these packets.

c) The first station transmits a voice packet during the first cycle and a VPPT case appears.

d) The maximum allowed number of additional \( \left( \frac{N_v}{2} - 1 \right) \) voice packets are transmitted during the first cycle and VPPT cases appear for all of them.

From Fig.1 the maximum waiting time for the tagged packet is:

\[
\frac{N_v}{2} h_v + u_0 - u_i + (\frac{N_v}{2} - 1) h_v + u_i = (N_v - 1) h_v + u_0
\]

where \( u_i \) is the walk time from station 1 to station \( N_v \).

Since this time must be less than \( P_v - h_v \) the following condition for the maximum number of voice sources (voice stations) \( N_v^* \) must hold:

\[
N_v^* h_v < P_v - u_0 \quad (1)
\]

**RVPT case**

In this case all voice stations can transmit voice packets during a cycle. The longest waiting time for the tagged packet is shown in Fig.2. Again, the tagged packet belongs to a speaker in the last station with partner in the first station. We see that we have two cases to consider:

a) If \( h_v \geq u_0 \), the worst case appears when the tagged packet arrives after the first bits of a packet from station 1 have passed and all other stations transmit packets in this cycle.

b) If \( h_v < u_0 \), the worst case appears when the tagged packet arrives in the previous cycle immediately after the departure of the token; the first station does not transmit a packet in the current cycle; all other stations transmit packets during this cycle.

\[
\begin{align*}
1 & \quad 2 \quad \cdots \quad N_v \quad 1 \\
\text{cycle} & \quad \text{cycle} \\
\text{max. waiting time} & \quad \text{max. waiting time} \\
a) & \quad b)
\end{align*}
\]

Fig.2: Worst case for RVPT procedure.

The corresponding maximum waiting times are \( (N_v - 1) h_v \) and \( (N_v - 2) h_v + u_0 \) respectively. Since these waiting times must also be less than \( P_v - h_v \), the maximum number of voice stations \( N_v^* \) must satisfy the following inequality:

\[
(N_v^* - 1) h_v < P_v - \max(h_v, u_0) \quad (2)
\]

If the number of voice sources is less than \( N_v^* \), then data transmissions are also allowed. We now derive the conditions that \( T_{MPC} \) and \( T_{TTT} \) must satisfy.

**MPC Protocol**

We now consider a ring with both voice and data sources and derive the condition that \( T_{MPC} \) must satisfy in this case. The operation now consists of both Synchronous and Asynchronous modes.

**VPPT case**

In Fig.3 \( T_1 \) is the time instant the SBM starts the Synchronous mode for first time. \( T_2 \) is the time instant SBM decides to switch to Synchronous mode for second time i.e.

\[
T_2 = T_1 + T_{MPC}, \quad T_3 \text{ is the time instant this second Synchronous mode actually starts.}
\]

The contribution of the voice stations to the maximum waiting time is the same with that of Fig.1. The contribution of the data stations is the additional delay introduced by the presence of the Asynchronous mode which, in the worst case, delays the beginning of the Synchronous mode by \( h_d + \max(h_d, u_0) \) [2]. From Fig.3 we see that the maximum waiting time is:

\[
T_{MPC} - u_i + h_d + \max(h_d, u_0) + (\frac{N_v}{2} - 1) h_v + u_i
\]

from which the following condition for \( T_{MPC} \) is directly derived:

\[
T_{MPC} < P_v - \frac{N_v}{2} h_v - h_d - \max(h_d, u_0) \quad (3)
\]

**RVPT case**

The longest delay case is shown in Fig.4. With a similar reasoning to that of the VPPT
case, we can easily see that the maximum waiting time is:

\[ T_{\text{MPC}} - u_t + h_d + \max(h_d, u_0) + (N_v - 2) h_v + u_t \]

from which the following condition for \( T_{\text{MPC}} \) results:

\[ T_{\text{MPC}} < P_v - (N_v - 1) h_v - h_d - \max(h_d, u_0) \]  (4)

We also repeat the inequality which must hold under Half-Duplex operation in order to clarify the effect of the type of operation (H-D, F-D) on \( T_{\text{MPC}} \) values.

**H-D operation**

\[ T_{\text{MPC}} < P_v - N_v h_v - h_d - \max(h_d, u_0) \]  (5)

If data packets of variable length are transmitted, then in the above inequalities \( h_d \) has to be replaced by \( h_{d_{\text{max}}} \), where \( h_{d_{\text{max}}} \) is the maximum length of a data packet.

Comparison of inequalities (3) and (4) shows that VPPT allows a value for the parameter \( T_{\text{MPC}} \) almost twice as large as the one allowed by RVPT. This implies a better data performance. We can also see from inequalities (4) and (5) that if RVPT is used only small improvement is expected as we go from H-D to F-D operation.

We also point out that inequality (3) alone, might give the wrong impression about the maximum number of voice sources that can be supported by the network. From (3) it appears that we can trade number of voice stations with values of \( T_{\text{MPC}} \) and that if we allow \( T_{\text{MPC}} \) to approach 0, the maximum number of voice stations can be given by the following inequality:

\[ \frac{N_v}{2} h_v < P_v - h_d - \max(h_d, u_0) \]

We see that the above inequality allows almost twice as many voice stations as those allowed by inequality (1). This contradiction in the allowed number of voice stations is however resolved if we notice that in the derivation of (3) we made the assumption that an Asynchronous mode of operation starts which tacitly assumes that the \( T_{\text{MPC}} \) value is higher than \( \frac{N_v}{2} h_v + u_0 \).

Thus we cannot make the value of \( T_{\text{MPC}} \) very small and increase the maximum number of voice stations to values larger than those computed by (1).

**TTP protocol**

We now consider a ring with both voice and data sources, and derive the conditions that \( T_{\text{TTP}} \) must satisfy in both VPPT and RVPT cases.

**VPPT case**

The tagged packet longest waiting time is shown in Fig.5. The contribution of the voice stations on this maximum waiting time is again as in the case of Fig.1. The effect of the data stations is the additional delay introduced by the two maximum duration (\( T_{\text{TTP}} - u_t \)) data transmissions, allowed by the VPPT procedure, between the arrival and departure of the tagged voice packet. From Fig.5 this maximum waiting time is:

\[ 2(T_{\text{TTP}} - u_0) + \frac{N_v}{2} h_v + u_0 - u_t + \left( \frac{N_v}{2} - 1 \right) h_v + u_t \]

from which the following constraint for the parameter \( T_{\text{TTP}} \) is derived.

\[ T_{\text{TTP}} < \frac{P_v - N_v h_v + u_0}{2} \]  (6)

**RVPT case**

The longest delay case appears in Fig.6 where two cases, a) \( h_v \geq u_0 \) and b) \( h_v < u_0 \), are also considered as in the case of Fig.2. With similar reasoning we can show that the following condition must hold in this case:

\[ T_{\text{TTP}} < P_v - (N_v - 1) h_v + u_0 - \max(h_v, u_0) \]  (7)

For comparison we repeat the condition for \( T_{\text{TTP}} \) when Half-Duplex operation is used:

**H-D operation**

\[ T_{\text{TTP}} < P_v - N_v h_v \]  (8)

From (6) and (7) we see that in TTP protocol the values for \( T_{\text{TTP}} \) that the RVPT procedure
provides, are almost twice as large as those provided by the VPPT procedure. This is contrary to what happens when MPC protocol is used. Comparison of inequalities (7) and (8) shows that if RVPT is used, then the effect of the type of operation, F-D or H-D, on the value of $T_{TPP}$ is rather minor. However comparison of (6) and (8) shows that if VPPT is used, then better performance is expected with H-D operation.

### 3. Modeling

We consider a $C_{ap} = 10$ Mbps channel capacity ring with $N_v$ voice/data stations attached to it. We assume PCM encoding for speech with rate $R = 64$ Kbps. The packetization interval is $P_v$. Only talkspurts are packetized and transmitted. During talkspurts voice packets arrive deterministically, one packet per $P_v$.

The values of $P_v$ and $R$ determine the voice packet length $l_v$ through the relation: $l_v = P_v R$. We assume fixed size data packets of length $l_d$ bits. An overhead $O_v$ of 100 bits is added to both voice and data packets. The voice and data packet transmission times are given by:

$$h_v = \frac{l_v + O_v}{C_{ap}}$$

and $h_d = \frac{l_d + O_v}{C_{ap}}$, respectively. The total carried data information load (without overhead) is given by $D_{thv}$. We assume that the data packet arrivals in each station follow a Poisson process. The bit latency per station is $L_d$. The cable length is 5 Km and the propagation delay 5 $\mu$sec/Km. For the MPC protocol the maximum number of data packets allowed for transmission per token arrival is given by the parameter $L$.

In our simulations we compare the performance of MPC and TTP protocols, under the VPPT and RVPT procedures respectively. We have chosen these procedures because they provide higher values for the operational time parameters $T_{MPC}$ and $T_{TPP}$. Table VII in [5] is used in simulating the voice conversations.

### 4. Performance Results

The stopping condition in our simulations was the number of voice packets transmitted by station 2. In most of the cases the value of $15 \times 10^3$ packets has been chosen. In addition, for some cases we have run longer simulations with stopping condition $40 \times 10^3$ voice packets. These simulations have showed that the $15 \times 10^3$ packets condition is a safe one for performance results throughout the whole data load range.

In Fig. 7 we have plotted, for both voice and data, the delay-throughput characteristics of TTP and MPC protocols under F-D and H-D operation. For MPC protocol these characteristics are, for the voice packets, straight lines almost independent of the data load. They are strongly affected by the value of the parameter $T_{MPC}$ which determines the frequency of the Synchronous modes and consequently the voice performance.

For TTP protocol both the offered data load and the value of $T_{TPP}$ affect the voice packet delay. Fig. 7 also shows that changes of operation (from H-D to F-D and vice versa) only slightly affect the voice performance. The voice packet delays encountered with TTP protocol are much lower than those encountered with MPC protocol. The reason is that in the MPC protocol case, voice packets cannot be transmitted during any cycle but they have to wait for a Synchronous mode of operation.

The effect of the type of operation (F-D or H-D) on data performance is minor for TTP protocol and significant for MPC protocol. This was predicted by inequalities (7), (8) for TTP protocol and (3), (5) for MPC protocol.

Comparison of the data performance of the two protocols shows that the MPC protocol under H-D operation has the worst performance. A rather strange result is the slightly higher delays that data packets encounter with F-D operation over those with H-D operation for the MPC protocol, in low up to medium ring utilizations. One expects that since the F-D operation provides a much higher value for the $T_{MPC}$ parameter (6.08 msec over 2.38 msec), the data performance with F-D operation to be always better than H-D operation. The reason for the worse data performance under F-D operation is the following: with higher values of $T_{MPC}$ parameter more voice packets are accumulated during the Asynchronous mode. This results in longer Synchronous
modes which increase the cycle time variance. It appears that the negative effect of this variance on performance, in low up to medium ring utilization, is stronger than the positive effect that higher values of $T_{MPC}$ have which reduce the walk time overhead by reducing the frequency of the Synchronous modes of operation.

The performance of the MPC protocol under F-D operation is worse than the performance of TTP protocol, under both F-D and H-D operation, over almost the entire range of the offered data load. However under very heavy traffic conditions the MPC protocol under F-D operation performs better even than TTP under F-D. However the region at which this occurs is not an operational one, since it is very close to saturation and at these loads the system becomes unfair.

We consider a fair system to be one in which stations with similar arrival patterns and service disciplines encounter similar delays regardess their position on the ring. TTP and MPC under F-D, TTP under H-D and MPC under H-D remain fair for ring utilizations up to .95, .93 and .9 respectively. For higher utilizations some stations are favored over others but there is no specific pattern followed.

We have also investigated the effect of other system parameters on performance under F-D operation and we summarize the results. For both protocols small values of their time parameters drastically affect the performance of voice and data packets. As these values increase their effect on performance decreases. The bit latency per station has a rather strong effect on the data performance for both protocols. However its effect on voice performance is minor.

The data packet length affects both voice and data performance of the two protocols. The transmission of short packets, due to the increased amount of carried overhead, causes system saturation at much lower values than the transmission of long messages. However under light to medium ring utilization the delay encountered with long packets is higher, due again to increased cycle time variance, than the delay encountered with short packets.

Finally all other system parameters that affect the values of $T_{MPC}$ and $T_{TTP}$ parameters drastically, also affect the performance of the system drastically. Such parameters are the voice packetization interval $P_v$ or, equivalently, the voice packet size $l_v$, and the total number of active voice sources $V_v$ or, equivalently for the F-D operation case we examine here, the number of voice stations $N_v$ connected to the ring.

5. Conclusions

We have investigated the performance of two Medium Access Control protocols which are appropriate for Integrated Services token ring networks. These protocols not only can provide priority services to different types of traffic, but they can also guarantee bounded delays for real time applications.

For both protocols we have derived under Full-Duplex voice transmission the constraints that their operational time parameters $T_{MPC}$ and $T_{TTP}$ must obey in order for the imposed bounds to be met.

Through simulation we have investigated the performance of both protocols under F-D operation and we have compared it to that of H-D operation. This comparison showed that the effect of the type of operation (H-D or F-D) on both voice and data performance of TTP protocol is minor and that use of the more complicate Full-Duplex operation over the simpler H-D is not justified in this case. However F-D operation drastically improves the performance of MPC protocol.

The TTP protocol with either type of operation has better performance than MPC protocol with F-D operation almost throughout the entire offered data load range. All three protocols however provide low delays for data packets even under heavy traffic conditions. Their delay-throughput characteristics have small slopes, throughout a wide region of offered data load and they rise sharply very close to their saturation point. These protocols are fair even under high ring utilizations and only under very heavy traffic conditions unfair behavior is observed.

Finally from all system parameters, those which appear in the derived inequalities and affect the values of the operational time parameters drastically, are the ones which also affect the performance of the system drastically.

References


