SUPPORTING MULTIMEDIA SERVICES OVER A TOKEN-PASSING BUS LOCAL AREA NETWORK

Abstract

Besides data transactions such as web surfing and e-commerce, the Internet today has been experiencing growth in audio and video streaming with the emergence of multimedia tools. However, the existing network infrastructure (e.g., Ethernet) only offers a single best-effort class of service without any guarantees on delay variation, packet losses or available bandwidth. This can adversely impact the end-to-end performance of multimedia applications. Accordingly, a different class of Media Access Control (MAC) protocols must be used to support multimedia services over the existing Local Area Networks (LANs) without seeking to change it [3].

The token-passing bus was introduced to avoid the non-determinism of Carrier Sense Multiple Access with Collision Detection (CSMA/CD) bus protocol as well as the complexities of active repeaters in the token ring [8]. This makes, token-passing bus MAC protocol is the best candidate to support multimedia traffic over Ethernet LANs [8].

In this paper, the capability of the token-passing bus LAN to provide multimedia services is explored and its performance under different multimedia traffic loads is evaluated using a simulation program.

I. Introduction

The basis for office information systems and factory automation is the widespread deployment of Local Area Networks (LANs). In such local environments, a LAN plays the key role of interconnecting user workstations and sharing of expensive devices and resources. Since LANs are usually constrained by their mode of operation to a geographically limited area, they can convey a high volume of data with rapid response time. Such performance is crucial for most office applications. The availability of the LAN in the mass market provides a base upon which other enhanced services can be implemented. The expected information to be exchanged between users on a LAN comprises different types of data other than simple characters or numerical information. In addition, conventional methods of interpersonal communications are acquiring more functionality and new methods are evolving. The emergence of multimedia computers and software which use all sorts of different audio, data, still image, and video signals simultaneously has heralded a new generation of computer applications that spurred the need for a powerful LAN capable of carrying different types of services and bit rates [3,4,12,13]. Obviously, it will be advantageous to use the existing network infrastructure,
originally designed for data applications, to support multimedia services without seeking to change it [3]. Exploring the feasibility of delivering multimedia applications over the already implemented networks is of a great importance from the technical and economical points of view. Migrating to a new technology may be a costly and not easy process that leads definitely to a delay in deploying multimedia services [3]. Instead, if the current technology is viable in front of the ever-growing user's needs, only purchasing the hardware and software supporting tools for video and audio on workstations are necessary. These tools are increasingly common with an ever-decreasing cost.

Technically, there are three standards, known collectively as IEEE 802 that define the operation of the most installed LANs. The IEEE 802.3 defines standards for the media access control (MAC) and physical layers for a carrier sense multiple access with collision detection (CSMA/CD) based bus network. The IEEE 802.4 defines those standards for a token-passing bus network, while the IEEE 802.5 defines standards for the MAC and physical layers for a baseband token-passing ring network [9,11,12]. The choice of the access protocol is the major determining factor in LAN performance [7,8]. However, the nature of the traffic offered by the user devices determines which protocol is more suitable for guaranteed quality of service (QoS). Multimedia applications integrate audio, images, video, graphics, text and traditional data. These services have different bit rates, delay and loss requirements. While data applications are bursty in nature, the audio and video traffic are stream-oriented that have to be delivered in a timely fashion [3,4]. Streams can either be constant bit rate or variable bit rate, but they are much less bursty than data. Although audio applications are time critical, they can sustain some loss. A loss rate of about 1% is tolerable and a transmission delay in the order of 50 msec is acceptable [2]. The required bit rate for data services are application-dependent. Electronic mail application for example generates sporadic single packet data traffic while a file transfer usually generates bursts of packets. Strict error control is a necessity so that data can be received inact and no loss is acceptable. The video applications, on the other hand, are delay and loss sensitive stream traffic that use a relatively high bit rate on a continuous basis for long periods of time. A high quality compressed video stream can use 1.5 Mbps [2].

Accommodating this variety of applications on a single network is a challenge as the provision of real-time properties on a LAN is not so easy to assure. The use of CSMA/CD is not suitable for transmitting time-constrained messages [1] because it cannot bound the transmission delay of messages and cannot distinguish between different traffic types [4]. This is due to the random nature of the access protocol that may result in continued collisions between contending nodes especially at the high traffic levels. In contrary, token-passing ring is implemented with a priority traffic handling mechanism [4]. However, the complexities of active repeater elements in the token ring physical system are inherent to the ring operation. The token passing bus is introduced to avoid the non-determinism of CSMA/CD bus as well as the complexities of active repeaters in the ring. The standard is designed primarily to be applicable in office, military, factory and other industrial environments where real-time applications are common [9,11,12]. The focus of the current paper is to explore the capability of the token passing bus LAN to provide multimedia services and evaluate its performance.

II. The token passing bus LAN configuration

A token passing bus LAN is an implementation of a logical ring on the physically connected bus. Each station has a unique address and the addresses are arranged in a logical ordering such that each station has a known up-station and
down-station on the logical ring. The physical ordering of the stations on the bus is irrelevant of the logical ordering. As the bus is made to look like a logical ring, the bus can support the token-passing protocol as a dynamic means of sharing the common communication medium [9,10,11]. The token is passed explicitly from one station to the next using the logical ordering as shown in Sketch 1 [2,3,7,13]. There are many advocates for adopting a bus topology instead of a ring for the token passing protocol.

These include reliability and simplicity. Reliability is a measure of how vulnerable the network is to failures or improper operations. Stations on a token bus are passive rather than active and thus there is no station latency or delay. In contrast, the ring topology incorporates repeaters in series. Repeaters cause unavoidable delay and the failure of any node can be catastrophic that leads to network failure unless recovery techniques are adopted (e.g. self-healing approach) [9]. The fact that stations on the bus are purely passive has a positive implication for reliability in the sense that the failure of one or more nodes does not cause the failure of the network in total. Additional nodes can be incorporated in the bus network without disrupting the network operation. Furthermore, as the token passing bus network uses a broadcast transmission, it is possible for stations on the logical ring to receive messages while they are not members of the logical ring. On the opposite, adding new nodes to the ring requires the network to be taken out of service in order to insert new cable segments and repeaters. This simplicity and flexibility property of the bus dictates the choice of the bus. The token passing bus LAN can, therefore, be viewed as combining the good features of the bus networks (i.e., reliability and flexibility) as well as the token ring (i.e., its deterministic token access protocol). Consequently, this right compromise can be suitable for industrial and other real-time applications [11,13].

Under steady state conditions, the operation of a token passing bus consists of alternating data transfer and token transfer phases. In effect, each node waits for a token to be passed to it. When a node captures the token and it has data for transmission, this data is transmitted in a packetized form. If a node has no data, the token is simply passed to the next node. When a station finishes transmitting its data, or when its access time limit is reached (whichever happens first), it sends the control token to the next station in the logical ring. During its access time, the token-holder may transmit one or more packets and may poll other stations and receive responses. The control in the token bus is, therefore, completely distributed and once its allocated access time is expired, the station must relinquish control of the token. The IEEE 802.4 token-passing bus

Sketch 1. A token passing bus LAN configuration.
MAC protocol standards support a priority mechanism that can be used to regulate access to the logical ring. The priority mechanisms are designed specifically in order to reduce the access delay for real-time traffic to the logical ring. Four classes of services are defined in a descending order of priority; namely 6, 4, 2 and 0. Any station may have data to send that belongs to one or more of these classes. High-priority frames are first allocated the bus capacity and lower-priority frames are only sent when there is a sufficient capacity. A station having a frame to transmit of equal or higher priority than the token level can only capture that token [9,11-13].

III. The Simulation Model

The complex operation of the token-passing bus LAN when supporting multimedia traffic precludes the derivation of an accurate analytical model. Consequently, a computer simulation model is developed for the network in order to investigate its capability to deliver multimedia applications. The simulation model is intelligently built in order to reflect state changes in the network behavior at the required level of details. This gives the freedom and ease to capture the token-passing protocol details and modify the simulation model as needed.

In our simulation, a number of stations are assumed to be attached to a transmission medium of 10 Mbps transmission speed. The transmission on the network is controlled by a token-passing mechanism. Once a station captures the token, it transmits its packet to the destination and broadcasts the token to the next station. We consider an ideal medium with no transmission errors so that no retransmission is required. The simulation model supports the two algorithms of token utilization, namely, the limited and exhaustive disciplines [7]. In the limited approach, an upper limit to the token holding time is imposed. Once expired, the station should release the token. In the exhaustive mode, on contrary, the station holds the token until all queued packets in its buffer have been transmitted. All stations in the network are active with one or more traffic types (e.g., audio, video or data). No priority levels are invoked in the simulation model which is illustrated by the flowchart Sketch 2.

The load on the network is described by specifying the number of distinctive classes of services and identifying the number of active stations for each class as well as the distribution of the arrival process of the related packets. Three categories of services are considered, namely, audio, video and data. For the audio source, the emerged analog signal is digitized by one of the commonly used speech coders. The bit rate of the audio source of 64 kbps is adopted [2]. Employing a speech activity detector, each audio source is characterized by on state (corresponding to the talkspurt period) and off state (corresponding to the silent duration) which appear in turn. During talkspurt, audio packets are generated periodically. During silent periods, on contrary, no packets are generated. The number of audio stations is designated by $N_a$. The talkspurt and silent periods are negatively exponential distributed random variables with expected values of 170 and 410 msec, respectively [2].

Video traffic requires larger data rate. If simple coding schemes are used, video can run up to 472 Mbps [9]. Fortunately, many compressing techniques can be used to reduce the video bit rate. A model for a compressed video source of 1.5 Mbps is considered [3]. We assume that the packets are reshaped such that a uniform packet stream is generated. The data sources generate fixed size messages with uniformly distributed inter-arrival times. A data source is characterized by its message size ($p_a$) and its average message inter-arrival time ($1/\lambda$, where $\lambda$ is the arrival rate of data packets). The uniform distribution was chosen over the exponential distribution for its lower complexity and faster convergence properties. We consider
the data source generate non-bursty data traffic, we let \( p_a = 1.024 \) bits, in which case each message fits in a single packet [3].

IV. Performance Measures

The token passing bus LAN must support diversity of service and performance requirements. Some traffic (e.g., data) is loss-sensitive and thus must be received without any errors or loss, whereas the inherent structure of audio allows for some loss of information without significant quality degradation. On the other hand, excessive delays have seriously disruptive effects on real-time traffic such as human conversation and video watching. Voice and video packets must be received at the destination station within a fixed amount of time (considered in this study as 50 msec and 20 msec for voice and video, respectively [2,3]) after their generation at the source station. Those packets that violate this strict delay limit are considered lost and should be discarded.

The components of delay suffered by a packet can be given by the following formula:

\[
D_o = W_p + W_a + T_s + R_{a,d}
\]  

(1)

where:

- \( W_p \) is the packetization delay during which the packet is generated (depending on the source rate and packet size), \( W_p \) is given by:

\[
W_p = P_a / A
\]  

(2)

where: \( P_a \) is the packet length excluding header (bits), and \( A \) is the source rate (bits/s).

- \( T_s \) is the transmission delay (which is a function of the transmission rate and packet size), \( T_s \) is given by:

\[
T_s = (P_a + H) / R_c
\]  

(3)

where: \( H \) represents the packet header length (bits), and \( R_c \) is the channel rate (bits/s).

- \( R_{a,d} \) is the propagation delay (which is constrained by the physical length of the medium and the propagation speed within

iv) and finally, \( W_a \) is the queuing delay at the source station. The queuing delay is the most significant participant in the packet delay [2,3]. The transmission delay of a data packet is defined in a similar manner except that there is no packetization delay.

It can serve as the major criterion for studying LANs, as other types of delays tend to be independent of the network. For real time traffics, if this delay exceeds the maximum tolerable limit, the packet is dropped. It should be pointed out that the Figures considered (1% for voice [2], less than 1% for video) have little or negligible effects on intelligibility of the human speech and video watching.

Consequently, we considered two-performance metrics for investigating the capability of the network to support multimedia services. These are the packet loss probability of real-time stream traffic and the average delay of data traffic.

V. Performance Characterization

1- Simulation Results for a Network with Voice Users Only

We first investigate the network performance when delivering audio service only. The relative performance is affected by the audio packet length. Fig.1 displays the average transfer delay of audio packets as well as its contributing components as functions of the audio packet length when the number of active stations on the network is 250 stations. The loss probability of audio packets for the same set of parameters is shown in Fig.2. For small packet lengths the queuing delay is the major contributor to the total transfer delay while the packet formation time is the only significant contributor for long data packets as observed in Fig.1. This is because small packet length leads to more number of packets waiting for a token. With the aid of Fig. 2 we observe that the loss probability is high for small packet length due to the previously mentioned long queuing delay.
We further investigate how to choose the optimal packet length that supports the maximum allowable number of stations while satisfied the service quality requirements of audio service in terms of keeping the probability of loss below 1% and concurrently preserve an average delay smaller than 50 msec. Fig.4 shows the maximum number of audio stations allowed on the network as a function of the audio packet length while attaining quality constraints. An audio packet of length $730 \leq P_a \leq 1250$ bits can achieve a number of 280 audio stations whereas a packet of $380 \leq P_a \leq 2050$ bits has the capability to achieve 250 stations. In the model, we assume that all stations are active all the time. In reality, however, stations are not always active. Consequently, a network supporting 250 concurrent conversations in the simulation study may be quite sufficient to accommodate 2500 stations in the real world. A packet length of 380 bits for a maximum of 250 stations achieves an optimum average delay while satisfying the loss constraint of 1% as shown in Fig.3. Consequently, we prefer to proceed in our study of 250 stations to consider an audio packet length of 380 bits in order to preserve the integrity of the voice conversation. This coincides with the prevailing opinion in the voice communication arena.

Although for a long packet length, the formation time is more significant, the loss probability is very small as the formation time is still below the maximum tolerable delay. The transmission delay is very small compared to the formation and queuing delays as a high-speed transmission media (i.e., 10 Mbps) is employed. Therefore, it has inferior effect on the total transfer delay. We study the preferred packet length that satisfies minimum delay under different traffic load by running the simulation program for different number of stations. A packet length between 310 and 400 bits results in a minimum delay for a number of stations between 250 and 300 as shown in Fig. 3.
2. Simulation Results for Network with Voice and Data

To study the effect of integrating data with voice on the token-passing bus LAN, we investigate the probability of audio loss as a function of audio packet length for various data traffic intensities ρ. The traffic intensity ρ is defined as (Nd × λ × ρd)/Rc, where Nd is the number of data sources on the LAN, λ the arrival rate of data packets, ρd the data packet length and Rc the channel rate. A data packet length of 380 bits is considered in the simulation, as this is a typical data packet length application. Fig. 5 displays the audio packet loss as a function of audio packet length for ρ=0.45, 0.6, 0.8. The number of audio and data stations is the same and equals 125 stations.

If can be seen that the data traffic intensity determines the window of the permissible audio packet length without violating the 1% audio loss constraint. An audio packet length of 380 bits can satisfy the constraint of less than 1% audio loss for ρ=0.45 as well as 0.6.

To check if this packet length is also optimum in terms of minimizing the audio and data delay we study the delay of audio and data packets and audio loss probability as a function of the number of audio stations for different audio packet length of 200, 380, 730 bits under a data bound ρ of 0.45. The number of data stations is fixed as 250 stations. Increasing the number of audio stations results in a rapid increase in the audio packet loss when working with small audio packet length (e.g., <380 bits).

On the other hand, when choosing to operate with a long audio packet length (e.g., >380 bits) the average delay data is significantly increased with increasing the number of audio stations as shown in Fig. 6. It can be concluded that the audio packet length of 380 bits is a compromise between the two performance measures.
3- Simulation Results for Network with Video, Voice and Data

Now, we proceed to study the capability of the token-passing bus LAN to support multimedia traffic that comprises audio, data and video. The stream sources are the audio and video sources, with the audio source being implemented while the video source is considered to deliver constant data rate of 1.5 Mbps. The audio packet length as previously concluded is considered as 380 bits while the data packet length is taken as 1024 bits. A data traffic intensity \( \rho \) of 0.45 is assumed. We investigate the network performance when no priority is introduced and the video is served on a exhaustive basis. Fig.s 7, 8 display the average transfer delay of media services and the probability of loss of audio and video services, respectively, as functions of the supported number of video sources.

The number of data sources is 50 stations while the number of audio sources is 25 stations. The figures are plotted for different video packet lengths of 800, 1500 and 2000 bits. The average delay of all media is tolerable up to 4 video sources. For 5 video sources, the average transfer delay of video sources ranges from 20-30 msec while the data and audio delays are rapidly increasing to unacceptable delay of about 22-30 sec.

Referring to Fig.8, we observe that the video and audio loss probabilities are exactly zero up to 4 video sources except when a video packet length of 800 bits is used wherein the audio loss probability is about 5.5%.

We increase the number of audio stations to 50 and investigate the network performance for the same set of parameters considered in Figs.7 and 8. The delay and loss probability are shown in Fig.9 and 10, respectively. The audio and data packets suffer excessive delay that may be unbearable when 4 video stations are supported especially with a video packet length of 800 bits.
Duplicating the number of audio sources significantly affects the loss ratio of audio packets when 4 video stations are considered for all video packet lengths. Furthermore, the video loss probability increases dramatically when 800 bits video packet length is applied.

In order to check the effect of service policy on the network performance, we study the case of Fig.'s 7 and 8 but with limited service discipline for video traffic. Figs.11 and 12 with limited discipline are corresponding to Fig.'s 7 and 8 with exhaustive discipline, respectively. A major difference between Fig.'s 7 and 11 is the small delay encountered by audio and data users even when 6 video sources are employed for the limited discipline case.

On the other hand, the video packets suffer more delay when 4 video stations are supported with a video packet length of 800 bits. In addition, the video delay, when 5 video sources being applied increases to unbearable limit. The other face of the picture is the increased loss probability of video packets shown in Fig. 12 whereas the audio loss probability is marginally zerned.

We further investigate the effect of a broad range of video packet length by examining the average packet delay of all media and the probability of loss of audio and video sources against various video packet lengths. Fig. 13 shows the average packet media delay as a function of the video packet length whereas Fig.14 displays the loss probability of audio and video packets for different data traffic load. A video packet length between 800-2000 results in minimizing the average delay and concurrently satisfies almost zero loss except when data intensity is chosen as 0.6 whereas the upper limit of the range of packet length is preferred.
4- Simulation Results for Network with Video, Voice, and Data (Compound Traffic)

Till now we have discussed different cases of network configuration and service policies with a fundamental premise, that is, each active station has only one type of traffic, e.g., audio, video or data. However, it is anticipated that some stations may have multiple traffic.

We proceed in evaluating the network capability of delivering multimedia services by considering the cases of some stations with multiple services. Fig.15 displays the average packet delay of audio, video and data traffic as functions of the number of stations with the three services. A total of 50 stations are assumed to be active on the network. Out of these 50, a number of 25 stations are considered to deliver data services only while the remaining stations (i.e., 25-number of multimedia stations) deliver audio and data services. For those stations with multiple-traffic, three queues are used. Within the station, the arrival times of packets in the multiple queues are compared and the packets are served on a FIFO basis. No priority is provided on the bus and exhaustive discipline is adopted. The proportion of lost packets is shown in Fig.16.

From the figures we observe that only 4 stations with three types of traffic can be supported. This gives impression that the
video traffic is the most significant factor in determining the network performance.

VI. Conclusion

Extensive simulations were performed to examine the effectiveness of a token passing bus local area network with video, audio and data traffic. The following observations summarizes the simulation results discussed in previous sections.

1) **Audio packet length can have a significant effect on network performance.** Selection of an optimal audio packet length is dependent on the performance criteria used. For example, when maximizing the number of audio users with a 50 ms control time and an audio packet loss probability below 1 percent, our simulation results yielded a maximum of 250 audio users and an optimal audio packet length of 380 bits.

2) **The data traffic intensity and control time.** (i.e., the time limit for determining which audio and video packets are lost) required for acceptable network performance in real-world environments can be satisfied by the token passing bus local area network. We consider 50 ms as a control time for audio packets as considered in [2], and 20 ms for video packets as considered in [3]. Data traffic intensity of $p=0.45$ determines the window of the permissible audio packet length without violating the 1% audio loss constraint.

3) **Video packet length can have a significant effect on network performance.** Small video packet length (i.e., <800 bits) have a bad effect on the average transfer delay of data packets and loss ratio of audio packets.

4) **A network with a limited service discipline yields slightly better performance than a network with an exhaustive service discipline.** We achieve a better performance for audio and data, in the range of loss probability and average transfer delay, when an exhaustive service discipline is applied on the video service.

References


Sketch 2. Simple flowchart of the simulation model.