Performance Comparison of DQDB and FDDI for Integrated Networks

Farzad Tari, Member, IEEE
Comdisco Systems, Inc.
919 E. Hillsdale Blvd.
Suite 300
Foster City, CA 94404
Phone: 415-574-5800

Victor S. Frost, Senior Member, IEEE
University of Kansas
2291 Irving Hill Rd.
Lawrence, KS 66045
Phone: 913-864-4833
FAX: 913-864-7789

Abstract
The performance of two high speed network protocols, IEEE 802.6 Distributed Queue Dual Bus (DQDB) MAN and Fiber Distributed Data Interface (FDDI), are compared based on their ability to support integrated traffic. The comparison includes the asynchronous priority characteristics of the two protocols as well as the performance characteristics of their isochronous service classes. In the integrated network considered in this study, voice stations are assigned to the isochronous class of both protocols. Video sources, which are models of video CODEC systems for ATM environments, are given the highest priority of the asynchronous class for both protocols. The asynchronous data traffic is distributed equally among three asynchronous classes. The two protocols are then compared on the basis of their throughput, delay, network utilization, and fairness.

1. Introduction
By increasing the demand for connectivity of LANs with data intensive applications, the existing solutions for LAN interconnection are becoming obsolete. LAN interconnections are commonly done using a public or private packet-switching network like X.25 or by leasing circuit switching lines from phone companies. These solutions can accommodate intermittent or bursty traffic; however, they are restricted to speeds of 64 kbps which introduces a bottleneck for today's 4 Mbps, 10 Mbps, and 16 Mbps. Although leased lines are not restricted to allocation in increments of 64 kbps, the users pay whether the lines are used or not. Besides, these methods do not provide an efficient way for supporting integrated services. The required bandwidth for a typical video conference, for instance, is about 5 Mbps [16]. Such services cannot be supported within the bandwidth limitation of the existing interconnection schemes.

Metropolitan Area Networks (MANs) are being designed to accommodate LAN interconnections. They carry different types of communication traffic simultaneously, allowing more station to communicate over longer distances and at greater speeds. Among several schemes that have been presented for MANs during the recent years, two clear competitors are: the IEEE 802.6 Distributed Queue Dual Bus (DQDB) [1-7, 13] and the Fiber Distributed Data Interface (FDDI) proposed by the ANSI [8-13]. Although both protocols support integrated traffic, the access schemes and the performance characteristics of these protocols are fundamentally different.

M.N. Huber and W. Schodl [13], compared the delay performance of the asynchronous service of these protocols for one priority class. This performance comparison was done under different loads of isochronous traffic (i.e. traffic generated by time sensitive services like voice). In this paper, we present a more thorough performance comparison of the two protocols for integrated traffic. Using computer simulations, we compare the two protocols on the basis of their asynchronous priority characteristics, performance of the isochronous service class, and fairness in serving the stations on the network. We also investigate the feasibility of supporting video conferencing on the highest asynchronous class for both protocols. The DQDB performance measures presented in this paper are based on the 1988 version of the DQDB protocol [1] which specifies four asynchronous priority classes and does not support the bandwidth management mechanism. As explained in Section 2, the results presented here are applicable to the latest version of the DQDB standard [2] assuming that the bandwidth management mechanism is disabled. The results of this study are presented after a short overview of DQDB and FDDI (Section 2) and a brief description of the simulation model (Section 3).

2. Overview of DQDB and FDDI

2.1 Distributed Queue Dual Bus (DQDB)
The DQDB uses a pair of counter flowing unidirectional buses, known as a dual pair. One bus is denoted bus A and the other is denoted bus B. Because both buses are operational at all times, the capacity of the network is twice the capacity of a single bus (Figure 1).

The transmissions on each bus are formatted as fixed length entities called "slots". All slots are generated at
the head of the bus known as the "slot generator" and are terminated at the end of the bus. Figure 2 shows the format of a slot. Each slot consists of an Access Control Field (ACF) and a segment, where the segment is used to carry the protocol data units between peer entities. The ACF is used to control reading a segment from a slot and writing a segment into a slot. The BUSY bit indicates whether the slot contain data and the REQUEST bit is used for sending requests for future packet transmissions.

In DQDB, there are two types of slots. Queued Arbitrated (QA) slots are used to transfer asynchronous segments. Non-Arbitrated (NA) slots are used to carry isochronous traffic where isochronous refers to the timing characteristics of events or signals recurring at known, periodic intervals such as conventional digitized voice or video. The queued arbitrated access scheme is fundamentally different than the non-arbitrated access scheme. In the remainder of this section, we describe these access schemes.

**Queued Arbitrated Slot Access:** Queued Arbitrated slot access is controlled by a function called "Distributed Queue" which allows formation and operation of a queue of QA segments which is distributed across the network. For the distributed queue to operate properly, each station on the network needs to maintain two counters called request counter (RQ_CTR) and countdown counter (CD_CTR). In the following, we consider access to bus A (access to bus B is a mirror image of this case).

Each station on the bus is either idle, when there is nothing to transmit, or active. When it is idle, it counts, via the request counter, the number of outstanding requests from the downstream station that downstream is with respect to the direction of slot flow on bus A. The RQ_CTR increases by one for each request received on bus B and is decremented by one for each empty slot passing by on bus A. When the station becomes active, the content of the RQ_CTR is transferred to CD_CTR and the RQ_CTR is set to zero. This initializes the countdown counter to the number of outstanding requests from the downstream stations. The station then sends a request on bus B by setting the request bit of the first request free slot on that bus to one. This request reserves a place for the station in the distributed queue. The CD_CTR is decremented by one for each empty slot that is passed on bus A. When it reaches zero, the asynchronous segment is written to the next empty slot which passes on bus A. If the transmission has to be followed by another transmission by the same station, the station has to send another request and the above procedure starts over.

The distributed queue protocol can be extended to allow multiple priority levels. There are four priority levels (0-3) specified in the protocol. For each priority, a separated distributed queue is formed and a separated request bit is used in the ACF. The operation of the counters would also be slightly different. When a request of priority 1 arrives on bus B, all RQ_CTRs of priority 1 or lower of idle distributed queues within the station are incremented. In this way, the value of each RQ_CTR indicates the number of QA segments in downstream stations with priority equal to or higher than the counter's priority level. For active distributed queues, the arrival of such a request causes the CD_CTR of priorities lower than 1 to be decremented. This priority scheme results in immediate access of a higher priority segment irrespective of the length of the low priority queues.

**Non-Arbitrated Access Scheme:** The non-arbitrated access scheme is designed for the transfer of isochronous service octets. As mentioned before, the unit of transmission here is called an NA slot. Unlike QA slots, NA slots can be shared among a number of users.

The slot generator at the head of the bus takes responsibility for sending a sufficient number of NA slots to ensure that all isochronous service users have adequate bandwidth available. When the head of the bus generates an NA slot, it places a Virtual Channel Identifier (VCI) into the slot header. For each VCI value that the node must access, the node will maintain a table indicating which octet position(s) within the slot that it should use for reading and writing. Thus, the node will read from a particular location within the slot and write to another position within the slot. It ignores all other octets. This read and write operation must be done while the NA slot is passing the node. If the NA slot contains a VCI value which is not used by the node, the entire slot is ignored.

**2.2. Fiber Distributed Data Interface**

FDDI is a protocol designed for a 100 Mbps token passing ring using a fiber optic medium. Like the DQDB protocol, FDDI is designed to accommodate both asynchronous and synchronous traffic. There are also eight levels of asynchronous priorities specified by the standard. Access to the channel is controlled by passing a small token frame around the ring. The station in possession of the token may transmit frames for a period specified by the token holding rules. When the stations do not transmit, they simply repeat incoming frames from downstream on the ring. Each station is responsible for removing frames that it transmits from the ring. When the transmitting station stops transmitting, it passes the access permission to
the next station on the ring by issuing a new token. An FDDI frame consists of 180 bits of protocol overhead plus the protocol data units (several PDUs can be encapsulated into one FDDI frame).

The FDDI protocol calls for a number of timers in each station to control the length of time the token may be held for transmitting frames of a certain class. As a part of ring initialization process, all station negotiate a Target Token Rotation Time (TTRT). The long term average token rotation time cannot exceed TTRT. The protocol also guarantees that the maximum token rotation time will not exceed 2 x TTRT. At the end of the negotiation process, the smallest value of TTRT requested is determined and is used to set the variable T_Opr (operative TTRT) identically in each station.

Each station is equipped with a Token Rotation Timer (TRT) which measures the successive time between arrivals of the token at that station. Normally, the TRT is reset each time the station receives the token. If the TRT counts up to T_Opr, it expires but does not reset. When TRT expires, Late_Ct (Late Counter), which counts the number of TRT expirations, is incremented. When the token arrives late at a station (Late_Ct ≠ 0), the TRT is not reset; it continues timing to accumulate the lateness of the current token rotation into the next token rotation time. The result of this accumulation is that if in one token rotation, the value of TRT exceeds T_Opr by A, in the next token rotation, asynchronous transmission will be restricted until the TRT is less than T_Opr by a total of time A. This ensures that the average token rotation time never exceeds T_Opr. Note that Late_Ct is reset to zero each time the token is received and if it ever exceed one, the error recovery procedures are performed.

A Token Holding Timer (THT) is used by each station to control the amount of time a station may transmit asynchronous frames. If the token is not late (Late_Ct = 0), the value of TRT is loaded into THT. However, the THT does not start timing till the transmission of synchronous frames has been completed. Each asynchronous priority class has a threshold value associated with it (T_Pri(i); i = 1,8). When THT reaches the threshold of a particular priority, frames of that priority may no longer be transmitted. However, transmission already in progress when THT expires are completed. The maximum threshold value for a priority is T_Opr.

The transmission of isochronous frames in each station is controlled by a synchronous holding timer. The amount of isochronous bandwidth allocation might vary from one station to another; however, the total isochronous allocation must meet the following equation:

\[ \sum_{i} S_{A_i} + D_{MAX} + F_{MAX} + \text{Token-Time} \leq T_{Opr} \]

\[ S_{A_i} = \text{Synchronous allocation for station } i \]
\[ D_{MAX} = \text{Propagation time for one complete circuit on the ring} \]
\[ F_{MAX} = \text{Time to transmit a maximum length frame (4500 octets)} \]

Token-Time = time required to transmit a token

Also note that the transmission of the isochronous class is done each time the station receives the token whether or not the token is on time.

3. The Network Model

To compare the performance characteristics of the 802.6 and the FDDI protocols, an integrated network consisting of the components shown in Figure 3 was considered [3]. As shown in this figure, the network is consisted of three types of components: voice multiplexers, video sources and data stations. For both DQDB and FDDI, the line rate and the network length were assumed to be 100 Mbps and 200 km, respectively. Note that these figures are the maximum allowable settings for FDDI. It was also assumed that the stations are equally distanced and that the velocity of media is 2/3 of the speed of light. The three types of stations shown in Figure 3 were placed on the network in such a way that the network would be homogeneous. Figure 4 shows the location of each station on the network.

In this section, the significance of each type of network component as well as the modeling perspectives of the network for both protocols are explained.

Voice Multiplexers: A voice station in this study refers to as a voice multiplexer that multiplexes the traffic generated by 48 PCM channels used in digital telephony. That is, it multiplexes two T1 lines [14]. To provide each PCM channel with 64 kbps of bandwidth, each channel is serviced by generating a packet voice consisting of 512 bits every 8 milliseconds. Thus, the traffic generated by 6 voice multiplexers of this kind consumes 18432 Mbps of the total network bandwidth. In this study, voice traffic was assigned to the isochronous service class of DQDB and the synchronous class of FDDI.

Video Sources: The Asynchronous Transfer Mode (ATM) has been approved to be the mode of transportation for Broadband ISDN networks [15]. The concept of ATM is based on the idea of dynamic allocation of bandwidth to a particular service based on the burstiness of the traffic generated by that service. Verbiest and Pinnoo developed a CODEC system which uses variable bit rate coding to transmit video over the ATM networks [16]. In this scheme, the required bandwidth depends on the type of the video service. For instance, cable TV requires an average of 16.8 Mbps whereas a video conference can be supported with an average bit rate of 5 Mbps. Kishimoto and Ogata [17] showed that, in ATM networks, the distribution of
mean interarrival time for packet video depends not only the type of the video service but also on the length of the video packets. According to this research, slow moving pictures, like video conferences, with small packet lengths (512 bits) have exponential mean interarrival distributions.

The four video sources considered in this study are modeled based on the assumptions explained above. It is assumed that each video packet contains 512 bits and that the video packets have exponential interarrival time distributions. The mean interarrival time is set to generate an average bit rate of 5 Mbps per video source. To minimize the packet loss rate caused by the network overload, the video service was assigned to the highest asynchronous priority in both protocols (i.e., priority 3 in 802.6 and class 7 in FDDI). Note that the four video sources support two active video conferences between nodes 3 and 18 and nodes 33 and 48 (Figure 3).

Data Stations: There are forty data stations on the network model. The load for these stations is equally distributed among three priorities (Priorities 0-3 for 802.6 and 4-6 for FDDI). It is assumed that the data packets have exponential message length distributions with the mean of 1536 bits. The mean interarrival time of the packets is assumed to be exponential. The destination addresses of the packets are set randomly with equal probabilities. Note that the parameter to be varied here was the total data rate of each station.

4. Comparison of Network Performance

The network model described in the previous section was used to compare the performance measures of the two protocols [3]. A primary consideration in FDDI was the adjustment of the Token Rotation Timers (TRTs) to provide service for the voice and video users irrespective of the load generated by the data stations. By trial and error, it was found that a reasonable setting for the TRT values of class 6, 5, and 4 could be 5.4 msec, 5.3 msec and 5.3 msec, respectively. The TRT for class 7 was set to 7 msec. The TTRT was set to 8 msec to meet the timing requirements of the voice packets.

The two protocols were compared on the basis of their throughput, network utilization, delay, and fairness characteristics. To make the terminology simpler, we refer to priority 2 (or class 6), priority 1 (or class 5), and priority 0 (or class 4) as third, second, and first priorities, respectively.

**Throughput Comparison:** The throughput characteristics of the two protocols for voice and video stations are shown in Figure 5. The horizontal axis in this figure is the load per priority for the data stations which is offered to the MAC layer from the higher OSI layer. This load which does not include any MAC layer protocol overhead is referred to as the information load. The figure shows that for all values of information load, the voice traffic has total media throughput of 0.184 which is the same as the offered voice traffic. The video traffic, however, which requires 20 Mbps (i.e., total media throughput of 0.2) gets better service from the 802.6 protocol. The 802.6 supports video at almost 20 Mbps up to the information load of 0.371. Then, by increasing the information load, it gradually decreases to about 19 Mbps. This drop in throughput occurs at the information load of almost 0.0895 when FDDI is used.

Figures 6 and 7 show the total media throughput versus information load per priority for the three asynchronous priorities used in data stations. In both protocols, all three priority classes are supported at small loads. As the load increases, the throughput of the first, the second, and the third order priorities, in that order, decreases. Also, after the load increases beyond a certain point (almost 0.15 for FDDI and 0.3 for 802.6), the total media throughput for a particular priority becomes less than the offered load at that priority. This is because in both protocols the highest asynchronous priority is assigned to the video sources. The data stations can use only the portion of the bandwidth which is not used by the voice and video stations.

Figures 8–10 provide comparisons between throughput characteristics of the two protocols. According to these figures, all priority classes of the 802.6 protocol have higher throughput than their corresponding priority classes of FDDI. Two possible reasons for lower throughput figures in FDDI are the followings: TRT values of FDDI are small in comparison with the network size and the FDDI protocol adds more overhead to the packets than the 802.6 protocol does. To an average packet size of 1536 bits, FDDI and 802.6 add 180 and 120 bits of overhead, respectively. The former reason will be elaborated on in the next subsection.

**Comparison of Network Utilization:** We define utilization at a particular load as the sum of the total media throughputs for all synchronous and asynchronous classes at that load. A comparison between network utilization of the two protocols is provided in Figure 11. Note that the horizontal axis is the information load per priority class of data stations.

According to the this figure FDDI and 802.6 have almost the same utilization for small loads (loads of less than 0.1). Also, note that at the data load of zero, the utilization is the sum of the throughputs of voice and video stations. As the load increases, however, the 802.6 provides better network utilization than FDDI. This happens for the same two reasons mentioned in the last part of the previous subsection. First, FDDI adds more overhead to packets. To the 512-bit voice and video packets, FDDI adds 180 bits of overhead whereas the overhead for the 802.6 is only 40 bits. Also, the
overhead for a data packet with average length of 1536 bits is 180 and 120 for FDDI and 802.6, respectively. The extra overhead of FDDI causes lower utilization because it wastes the given bandwidth by sending overhead instead of information.

The second reason is that the TTRT and the TRT values of FDDI are small in comparison with the network size. As mentioned before, the reason for selecting small TRT values was to make sure that the token rotates around the network fast enough so that the time sensitive services (i.e. packet voice and packet video) could get network access on time. The effect of this fast rotation of the token on the data stations, at higher loads, is that they will not be able to hold the token long enough to transmit all their packets. The utilization of an FDDI network can be increased by choosing larger values of TTRT.

Delay Comparison: The delay characteristics of both protocols are shown in Figures 12 and 13. These figures indicate that, for both protocols, the average access delay increases as the information load is increased. Also, the delays increase very sharply as the load per priority approaches the peak throughput at that priority. Because voice and video priorities have small packet loss rates, their delay graphs never increases sharply. Also, as expected in both protocols the higher priorities experience less delay than lower priorities.

Figures 14-17 provide comparisons between each priority class of the two protocols. Figures 15-17 indicate that all asynchronous classes of data stations experience less delay when the 802.6 protocol is used. At small loads, the larger delay of FDDI is caused by the ring latency. That is, if the station which has a packet to transmit is not holding the token, it has to wait until the token arrives at the station. The 802.6, however, provides almost immediate access at small loads. The inferiority of delay characteristics of FDDI at larger loads is due to the small values of TRTs used in the FDDI network.

Figure 14 shows the comparison between the delay characteristics of the video sources. The 802.6 provides the video sources with less delay for loads of up to 0.55. Also, note that if the video receivers require a delay of less than 1.5 msec (minimum delay provided by FDDI), FDDI cannot be used to support the video service.

Fairness Comparison: The comparisons given in the previous sections were based on the global performance measures of the network. In this section, the fairness of the two protocols are compared by considering the per-node throughput of the highest priorities of data stations. Figures 18 and 19 show throughput versus order of data stations for the two protocols. Throughput is the ratio of the number of packets serviced at a station to the number of packets arrived at a station. The horizontal axis in these figures is the order in which the data stations appear on the network. For example, the data station with the order of three is the third data station on the network which, according to Figure 4, is station 4.

Figure 18 shows the fairness characteristics of the 802.6 protocol for three different loads. At small loads (load of 0.1) all the data stations get the throughput of 1. However, as the load increases, the stations that are farther from the slot generators, particularly the ones in the middle of the bus, get smaller throughputs. As shown in [4,7], this behavior is caused by the fact that the propagation delay of the request bits is larger for the stations that are located farther from the slot generators. This longer propagation delay also causes skewing in the queueing order of the distributed queue. Note that we have assumed that the bandwidth balancing mechanism, which could be used to alleviate this unfairness, is disabled.

Figure 19 shows that FDDI treats the data stations more equally than the 802.6 protocol. At small loads (load of 0.1) all stations get throughputs of almost one. As the load increases, however, the throughputs of all stations decrease. This is because as the load increases, the stations have more packets to send. However, due to the increased load of each station, the token does not rotate fast enough so that the stations can transmit all their packets. Consequently, the rate of packet loss increases which causes the per-node throughput defined in this section to decrease. This figure also illustrates a very interesting behavior of the network. The four data stations that are located right before the video sources have higher throughputs than the rest of the data stations. This effect which is noticeable at higher loads can be interpreted in the following way. The FDDI network has been tuned to allow the video sources to transmit all their packets even when the offered load from the data stations is high. That is, the amount of time that each data station gets to hold the token depends on the portion of the TTRT that is left to serve video and voice packets. Figure 19 suggests that, on average, by the time the token reaches the data station located right before a video source, there is enough time left to not only serve the video source but also to transmit more packets of the data station.

5. Conclusions

The results obtained in this study indicated that the DQDB protocol had higher network utilization and throughput for all priorities of data stations. Also, all priorities of data stations experienced less delays when the DQDB protocol was used. As far as the video stations are concerned, it turned out that for loads of up to 0.55, the DQDB protocol had superior delay and throughput performance. Although both protocols support the video service without significant loss of throughput, if the maximum delay requirements of the video service is less than minimum delay provided by
FDDI, this protocol cannot be used to support video. As far as the fairness characteristics are concerned, it was shown that almost all the stations get the same throughput when FDDI is used. The data stations that are located right before the video stations, according to the simulation results, get slightly higher throughputs than the rest of the data stations. In the DQDB model, however, the stations that are farther from the slot generator get lower throughput. This effect is even more pronounced at higher network loads.

Open issues for further study include studying the impact of using the bandwidth management mechanism of DQDB on the fairness characteristics of the protocol, studying the effects of replacing the data stations considered in this study by LANs, and replacing the video stations by video multiplexers. Also, the consequences of using different network topologies on the fairness characteristics of the two protocols need to be further studied.

References
Fig. 3: The Integrated Network Components

Data Stations:
- Packet Size: 1366 bits
- Exponential Arrival Distribution

Voice Multiplexers:
- Multiplex 2 T1 Base
- Packet Size: 312 bits

Video Sources:
- Packet Size: 512 bits
- Arrival Rate: 3 MPEG Sources
- Exponential Arrival Distribution
- Two Active Video Conferencing

Fig. 4: The network arrangement used for the simulation model
Figure 11. Utilization Versus Info. Load Per Priority

Figure 18. Throughput Per Node Versus Order of Data Station 802.6; Priority 2

Figure 19. Throughput Per Node Versus Order of Data Station FDDI; Class 6