METROPOLITAN AREA NETWORKS

Metropolitan area networks (MAN) have been studied, standardized, and constructed for less than 20 years. During that time the capabilities of the telecommunications network and the requirements of users have changed rapidly. What a MAN is supposed to do has changed as quickly as MANs are designed. Recent changes in user requirements, resulting from the growing use of the Internet at home, are likely to redefine MANs once again.

To understand the evolution and predict the future of MAN, one must consider the applications and alternative technologies. There are also inherent differences in the capabilities of local, metropolitan, regional, and wide area networks.

HISTORY

The first mention of MANs, that I am aware of, occurred at a workshop on local area networks (LAN) in North Carolina in the late 1970s. One session at the workshop was dedicated to customer experience with LANs. One of the customers, from a New York bank, described a successful application of LANs but complained about the difficulty he had transferring data between branches of the bank in the same city.

The feeling among the workshop participants was that we could do better connecting sites in the same city than using technology that was designed for a national network. In the 1970s telephone modems were expensive, about a buck a bit per second, and the highest rate modem that was generally available was 9.6 kbit/s. High-rate private lines, such as the current T-carrier system, were not widely deployed. Using the available technology was expensive and created a bottleneck between LANs. With the customer and application identified, work on MANs began.

The first MANs were designed to interconnect LANs. They evolved from LANs and looked very much like LANs. Two MAN standards that are clearly related to LANs, fiber distributed data interface (FDDI) and distributed queue, dual bus (DQDB), are described later in this article. A third network, which is based on a mesh structure, the Manhattan street network (MSN), is also described. The MSN is a network of two-by-two switches that operate on fixed-size cells. The MSN straddles the middle ground between a LAN and a centralized asynchronous transfer mode (ATM) switch (1).

Interest in MANs waned as the useful functions performed by MANs were subsumed by wide area networking (WAN) technology. WANs had a much larger customer base than MANs. It became more economical to interconnect LANs in a city with routers and private lines than to deploy new, special-purpose networks.

There is a resurgence in interest in MANs because of the World Wide Web (Web). The time required to download Web pages using WAN technologies is frustrating many users and constraining the growth of this service. As more and more individual homes are connected to the Internet, there is a rapidly growing demand for bursty, high-rate data to a large number of locations in a metropolitan area. Therefore, there is a renewed interest in MANs, although the requirements and customer set are completely different from those of the earlier MANs.

DIFFERENCES AMONG LANS, MANS, AND WANS

The distance spanned by WANs is greater than that by MANs, and the distance spanned by MANs is greater than that by LANs. It is useful to define the maximum distance spanned by the various network technologies as increases of an order of magnitude. LANs span distances up to 3 miles, and include most networks that are installed in a building or on a campus. MANs span distances up to 30 miles (50 km according to the standards committees) and can cover most cities. RANs (regional area networks) span distances up to 300 miles, the area serviced by the telephone operating companies in the United States. And WANs span distances up to 3000 miles, the distance across the United States. The next order of magnitude increase covers international networks. The distances spanned by networks affects the transmission costs, access protocols, ownership of the facilities, and the other users who share the network.

The transmission costs usually increase with distance. This cost affects both the applications that are economically viable and the protocols that are used to transfer data. For instance, access protocols have been designed for LANs that trade efficiency for processing complexity. Carrier sense multiple access/collision detection (CSMA/CD) protocols, in which many users share a channel by continuing to try until the data successfully get through, are used on LANs. WANs use reservation mechanisms that require more processing, but pack the transmission facility as fully as possible. With CSMA/CD protocols the propagation delay across the network must be much less than a packet transmission time, which precludes using these protocols in WANs.

Traditionally, LANs are networks that are owned and installed by a single company or organization. An organization can choose to try new technologies. MANs are less expensive to install than WANs and may not be interconnected. There is more freedom to experiment with new technologies on MANs than on WANs. The expense of installing MANs relative to installing LANs has resulted in far fewer experimental MANs than LANs.

In a LAN the other network users are generally more trusted than the users in a more open environment. Traditionally, MANs, such as CATV networks, and RANs and WANs, such as the telephone network, service an unrelated community of users. The users in these networks do not trust one another as much as the users on a LAN, and greater measures must be taken to protect data.

There are increasing numbers of wide area networks that are owned and controlled by a single organization. Corporate networks and intranets, which use Internet technology within a corporate network, are becoming common. These networks have trust structures and flexibility that is more closely related to LANs than to WANs. The differences between general WANs and intranets are reflected in the applications of the networks and are leading to different implementations.

Many of the economic tradeoffs that are related to the distance spanned by networks change with time. However, the difference in propagation delay can never change. As the size of the network increases, the maximum useful transmission rate that a user can access to transfer a particular size packet decreases. This phenomenon is demonstrated in Fig. 1. The lines in this figure show when the propagation delay and transmission time are equal in LANs, MANs, RANs and WANs. The calculations are performed assuming that the propagation delay in the medium is 80% of the speed of light in free space, which is common for optical fibers. To the right of these lines, the time it takes to get the message from the source to the destination is dominated by the propagation delay rather than the transmission time.

Increasing the transmission rate when operating to the right of the line does not bring a commensurate decrease in the time it takes to deliver a message. For instance, on a 3000 mile WAN the equilibrium point on a 1.5 megabits/s T1 circuit is about 30.3 kbit. For a message of this size, the propagation delay and transmission time are equal. If the user's rate is increased to 45 megabits/s, a T3 circuit, which is 30

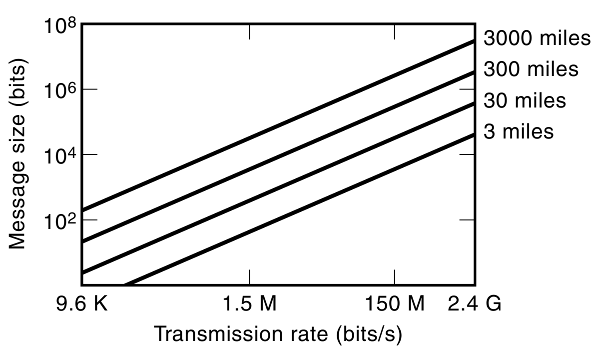

Figure 1. Equal delay lines when the distance between the source and destination is 3, 30, 300, and 3000 miles. Along the line the propagation delay for light and the transmission time of the message are equal.

times faster, the time to transmit the message decreases by a factor of 30, but the propagation time remains the same. The time it takes the message to get to the destination decreases by less than a factor of 2 as the transmission rate increases by a factor of 30. At T3 rates the message delivery time is almost entirely due to the propagation delay, and if the user's rate increases to 155 Mbits/s, an ATM circuit, there is virtually no decrease in the time it takes to receive the message. If the user sends the same message on a MAN and increases the rate from T1 to T3, the time to deliver the message decreases by almost a factor of 30, and increasing the ATM rates decreases the delivery time by another factor of 2. Therefore, ATM rates may be used to obtain faster delivery of this size message in a MAN, but not in a WAN.

THE MAN ANOMALY

Generally, transmission links cost more as the distance increases. At present, high-rate channels are readily available in LANs and WANs, but not in MANs. The use of computers in offices has become ubiquitous and has resulted in most office buildings being wired for high-speed communications. The backbone of the wide area telephone network is shared by a large number of users. Even though only a fraction of the users require high-rate facilities, the number is large enough to warrant providing those facilities between central offices.

Fiber to the curb and other methods to provide high data rates in a MAN exist, but most users do not require these rates. The lines running down a street in a MAN are not shared by as large a number of users as the lines between central offices. As a result, when high-rate channels are installed in an office, the line that spans the final mile or two from the central office to the office building is frequently an expensive, custom installation.

The current networks must be modified to provide high data rates to homes until the demand increases to the point where new facilities are justified. Two possible technologies are digital subscriber loops (DSL) and the CATV network. DSL and ADSL (asymmetric DSL) use adaptive equalizers to transmit between 1.5 and 6.3 Mbits/s over current local loops in the telephone network. ADSL provides higher rates in one direction than in the other.

To date, the only really successful MAN for distributing information to a large number of homes has been the CATV network. CATV networks are mainly used to distribute entertainment video; however, experimental networks are being deployed to deliver data to homes. Standards organizations and working groups are actively considering these networks. The multimedia requirements of the Web may well make CATV technology the correct solution for the next MAN. Several of the early proposals for MANs used CATV networks to deliver point-to-point voice and data services, as well as broadcast TV. Later we describe one of these techniques, which is still one of the most forward-looking CATV solutions.

THE FIBER DISTRIBUTED DATA INTERFACE

FDDI (2) is a token passing loop network that operates at 100 megabytes/s. It is the American National Standards Institute (ANSI) X3T9 standard and was initially proposed as the successor to an earlier generation of LANs. FDDI started as a LAN and has been primarily used as a LAN; however, it is capable of transmitting at the rates and spanning the distances required in a MAN. Therefore, it has become common to discuss FDDI in the context of MANs.

Baseband Transmission

FDDI uses a baseband transmission system. Baseband systems transmit symbols, ones and zeros, on the medium rather than modulating the symbols on a carrier, as in a radio network. Baseband systems are simpler to implement than carrier systems; however, the signal does not provide timing and there may be a dc component that is incompatible with some system components. For instance, a natural string of data may have a long sequence of ones or zeros. If the medium stays at the same level for a long period, it is difficult to decide how many ones or zeros were in the string and the dc level of the system will drift toward the value of that symbol. To tailor the signal to have desirable characteristics, the data are mapped into a longer sequence of bits.

A common code for transmitting baseband data on early twisted pair networks is a Manchester code. Each data bit is mapped into a 2 bit sequence, a one is mapped into $+1,-1$ and a zero into $-1,+1$. There is at least one transition per bit, which provides a strong timing signal. There is no dc component in this code. Twisted pairs are connected to receivers and transmitters by transformers to protect the electronics from energy picked up by the wires during lightning storms, and transformers do not pass dc. A framing signal is needed to identify bit boundaries and the beginning of a sequence of bits. Framing signals occur infrequently and are sequences that do not occur in the data. Framing can be obtained by alternately transmitting $-1,-1$ and $+1,+1$ every n bits. With a Manchester code the bit rate on the medium is twice the bit rate from the source; however, this is a very simple coding system to implement.

The FDDI standard uses a rate 4/5 code that maps 4 data bits into 5 transmitted bits. The constraints on codes used for fiber optics are different from those on twisted pairs. Fibers do not act as antenna during lightning storms, are not coupled to electronics through transformers, and can tolerate a dc component. In addition, logic costs have decreased since the early twisted pair networks were designed, so that it is now reasonable to implement more complex bit mappings and to design signal extraction circuits with fewer transitions. In the FDDI code there are 16 possible data patterns per symbol and 32 possible transmitted patterns. The 16 patterns are selected to guarantee at least one transition every 3 bits. Some of the remaining patterns serve control and framing functions. The use of the transmitted patterns is listed in Ref. 3.

Token Passing Protocol

A token is a unique sequence of bits following a framing sequence. When a station on the loop receives the token, it may transmit data after changing the token to a different pattern of bits. When the station has completed its data transmission, it transmits a framing sequence and the token so that the next station has a chance to transmit. When a station does not have the token, it forwards the data it receives on the loops. A station may remove the data destined for itself. When a station has the token and is transmitting, it discards any data it receives on the loop. The discarded data either passed this station prior to the token, and has circulated around the loop or was transmitted by this station after accepting the token. In either case, the data have circulated around the loop at least once and every station has had a chance to receive the data.

In FDDI there is a time for the token to circulate around the loop, a target token rotation time (TTRT). In a simple token passing protocol one station can hold the token for a very long period of time. That station can obtain a disproportionate fraction of the bandwidth and delay other stations for long periods. In the FDDI protocol, station i is entitled to send $S(i)$ bits each time it receives the token. The TTRT is set so that every station can send at least the bits it is entitled to send in a single rotation.

$$\sum_i S(i) + \Delta < \text{TTRT}$$

where Δ is the propagation delay around the loop plus the maximum time that is added by the transmission format. The TTRT provides guaranteed bandwidth and delay for each station.

After a station i transmits $S(I)$, it can continue to transmit data if the time since it last forwarded the token is less than TTRT, which indicates that the token is circulating more quickly than required. How long a station holds the token depends on the priority of the data it is transmitting. If the data are high priority, the station can hold the token until all of the surplus time in the token rotation has been used. If the data are lower priority, the station may leave surplus time on the token to give stations with higher-priority data a chance to acquire the surplus time.

The TTRT is set individually for each FDDI system depending on the requirements of the stations on that system. The maximum delay that can occur is 2*TTRT, and the average token rotation time is less than TTRT (4). Therefore, TTRT can be set to provide the guaranteed bits, $S(i)$, and an upper bound on delay for synchronous applications. When best effort traffic, rather than traffic that requires service guarantees, is the dominant traffic type, then TTRT is set to trade access delay and efficiency. As TTRT is made smaller, the time delay until a token arrives decreases. As TTRT is made larger, the amount of data transmitted before passing the token increases, less time is spent passing the token, and the efficiency of the system increases. The increase in efficiency is greatest when there is only one active station.

The operation of the token protocol is depicted in Fig. 2. When a station transfers the token, it sets the local counter TRT(i) to TTRT. The station decrements TRT(i) every unit of time, whether or not it has the token. If TRT(i) reaches $-$TTRT before the token is received, then the token has not visited this station in 2*TTRT and is presumed to be lost.

When station i receives the token it has $W_j(i) \geq 0$ units of data waiting to be transmitted at priority level j, where $j_{\max}$ is the highest level and $W_{j_{\max}}(i) \leq S(i)$. A station transmits $W_{j_{\max}}$ no matter how much time is left on TRT(i). For the priority levels $j < j_{\max}$, a station transmits a data unit as long as TRT(i) is greater than the threshold T_j. $T_{j-1} < T_j$ so that we never transmit data at level $j - 1$ if we cannot transmit data at level j. We process priority $j_{\max}$ data in the same structure as the other levels by setting $T_{j_{\max}} < -(2^*\text{TTRT} + S(i))$, so that data at that level are never inhibited by the threshold.

Tokens that circulate the loop and can be used by any station are called unrestricted tokens. The FDDI standard also supports a mode of operation in which all of the capacity that is not being used by synchronous traffic is assigned to a single station, possibly for a large file transfer. To support this mode of operation, a restricted token is defined. A station that enters this mode forwards the restricted token rather than the standard token. Another station that receives a restricted token transmits its synchronous traffic, but does not transmit traffic at level $j < j_{\max}$. Therefore, all of the capacity available for asynchronous traffic is given to a single station until that station forwards an unrestricted token.

Isochronous Traffic

FDDI-II adds the ability to send isochronous traffic to FDDI. An isochronous channel provides a regularly occurring slot. The channel is assigned to a specific station on a circuit switched basis.

FDDI-II is implemented by periodically switching the network between a circuit switched and packet switched mode. There is a central station that sends out a framing signal every 125 μs. A portion of the interval following a frame signal is assigned to circuit switched traffic, the isochronous mode, and the remaining time in the frame is assigned to the token passing protocol. An isochronous station that is assigned one byte per frame has a 64 kb/s channel. This channel is adequate for telephone quality voice.

In FDDI-II the stations that implement the token passing protocol must switch between the two modes of operation when they receive framing signals. A station that enters the circuit switched mode must forward whatever bits it receives. When the circuit switched mode ends, the station must resume the token passing protocol where it left off.

Architecture

A single failure of a node or a link disconnects the stations on the loop. In a LAN, loops are made more reliable by using normally closed relays to bypass individual stations that lose power or fail in an obvious manner. It is also common to arrange stations in subloops that are chained together at a central location. Subloops with failures are removed from the network (5), so that the stations on the other subloops can continue to communicate.

Poor reliability prevents loop networks from spanning the distances and connecting the number of users associated with MANs. In FDDI the reliability is improved with a second loop. The second loop does not carry data during normal operation, but is available when failures occur.

Figure 3 shows three components that are used in FDDI networks:

A. Units that implement the token rotation protocol
B. Units that manage the reliability
C. A unit that is responsible for signal timing and framing

Type A units connect user devices to the primary loop. There can be more than one type A unit attached to a type B unit. The type A units do not have to be collocated with the type B

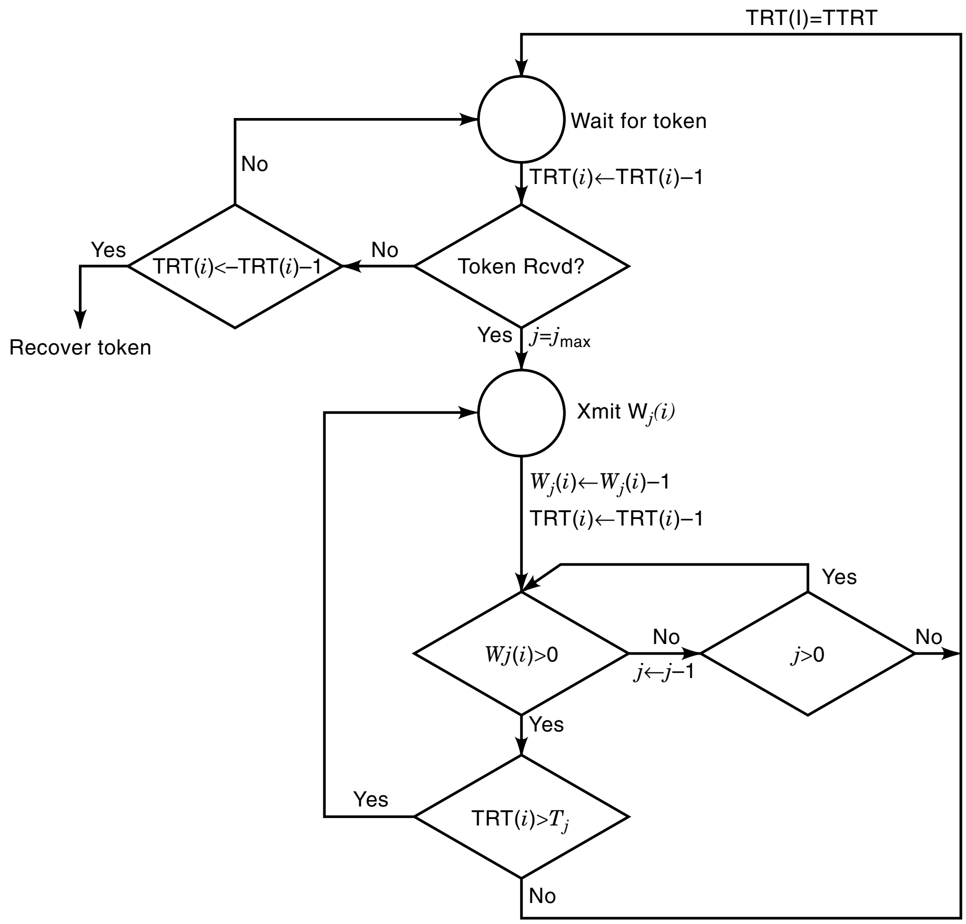

Figure 2. The flow diagram for the timed token rotation protocol in FDDI. The diagonal boxes are decision points where a terminal decides whether or not to transmit when a token is received dependent upon a local timer.

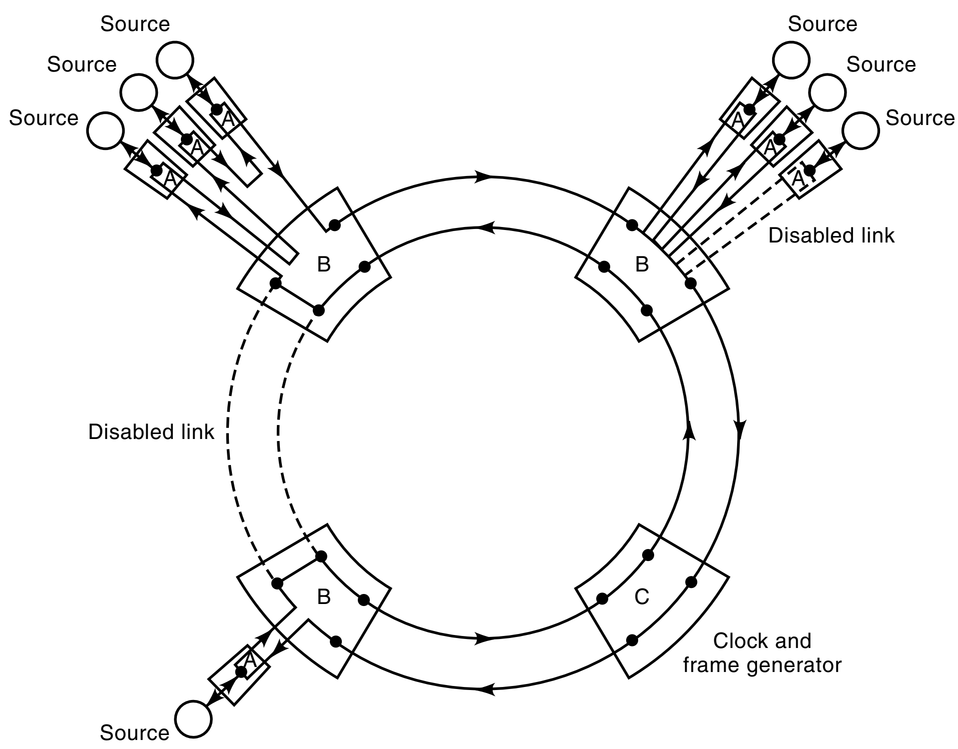

Figure 3. The topology of an FDDI loop. This shows the application of the three types of units that are defined for an FDDI in unidirectional and bidirectional loops. The dashed lines show how the loops are reconfigured after failures.

unit, and they can be daisy-chained together to form a subloop.

Type B units are connected to both loops. These units monitor the signal returning from type A units and bypass type A units that have stopped operating. Type B units also monitor the signal on the two loops and bypass failed loop components.

The secondary loop is used to bypass failed links, or failed type B units. The signal on the secondary loop is transmitted in the opposite direction from the primary loop. Normally type B units patch through the signal that they receive on the secondary loop. When a primary loop failure is detected, by a loss of received signal on that loop, a type B unit replaces the lost signal with the signal it receives from the secondary loop and stops transmitting on the secondary loop. When a type B unit stops receiving the signal on the secondary loop, it replaces that signal with the signal it would have transmitted on the primary loop and stops transmitting on the primary loop. As an example, in Fig. 3 the X signifies a link failure and the — (horizontal bar) indicates the link on which the type B unit has stopped transmitting. The entire secondary loop replaces the single failed link on the primary loop.

The configuration of type A and type B units reflects the way loop networks are installed. Loop networks are installed by running wires or fibers from an office to a wiring cabinet. The type A units are located in offices and the type B units are located in the wiring cabinets. Physically, the topology is a star, but the connection of wires in the wiring cabinet form a logical loop. The star topology makes it possible for stations to be added to a loop, or moved from one loop to another, without rewiring a building.

THE DISTRIBUTED QUEUE, DUAL BUS PROTOCOL

DQDB (6,7) is the IEEE 802.6 standard for MANs. It uses two buses that pass each station. The stations use directional taps to read and write data on a bus without breaking the bus. Directional taps transmit or receive data from one, rather than both, directions at the point of connection, and are common components in both CATV and fiber optic networks.

DQDB transmits information in fixed-size slots and uses the distributed queue protocol to provide fair access to all of the stations. Signals on the two buses propagate in opposite directions. A station selects one bus to communicate with another specific station and uses the other to place reservations for that bus.

The DQDB standard was preceded by two earlier protocols, Express-Net (8) and Fasnet (9), that used directional taps. In the two earlier systems there was a single bus that passed each station twice. On the first pass each station could insert signals and on the second pass a station could receive signals from all other stations. Both of the earlier protocols provided fair access by guaranteeing that every station had the opportunity to transmit one slot before any station could transmit a second slot.

Passive Taps

Passive taps distinguish directional buses from loop networks, which use signal regenerators. In loop networks, there is a point-to-point transmission link between each station. Each station receives the signal on one link and transmits on the next link. A station can add or remove the signal on the loop. A failure in the electronics in a station breaks the communications path. By contrast, the stations on a directional bus network do not interrupt the signal flow. A passive tap reads the signal as it passes the station, and another tap adds signal to the bus. If a station with passive taps stops working, the rest of the network is not affected.

The protocols that can be used on bus networks are a subset of the protocols that can be used on loop networks, since the stations can add but not remove signals. The inability to remove signals makes it necessary for the the bus to have a break in the communications path, where signals can leave the system.

The protocols for directional buses can be implemented using regenerators rather than passive taps when it is advantageous. Passive taps remove energy from the signal path, and the signal must be restored to its full strength after passing several stations. In addition, by removing information from the transmission medium after it is received, the medium can be reused and to support more communications.

The DQDB standard provides for erasure nodes (10,11), which remove information that has already been received. Erasure nodes are regenerators that read slots and, depending on the location of the destination, regenerate the slot or leave the bus empty.

Architecture

The dual bus in a DQDB network is configured as a bidirectional loop, as shown in Fig. 4. The signal on the outer bus propagates clockwise around the loop, and the signal on the inner bus propagates counterclockwise. The signal does not circulate around the entire loop, but starts at a head-end on each bus and is dropped off the loop before reaching the head-end.

To communicate, a station must know the location of the destination and the head-ends and transmit on the proper bus. For instance, station A transmits on the outer bus to communicate with station B, and station B transmits on the inner bus to communicate with station A.

The dual bus is configured as a loop so that the head-end can be repositioned to form a contiguous bus after a failure occurs. The head-end for each bus is moved so that the signal is inserted immediately after the failure and drops off at the failure. This system continues to operate after any single failure.

The ability to heal failures increases the complexity of stations on the DQDB network. To heal failures, the station that assumes the responsibility of the head-end must be able to generate clock and framing signals. In addition, after a failure each station must determine the new direction of every other station. For instance, after the failure in Fig. 4 is repaired, station A must use the inner bus, rather than the outer bus, to transmit to station B.

The Access Protocol

In a DQDB network transmission time is divided into fixed-size slots that stations acquire to transmit data. A station at the beginning of the bus periodically transmits a sync signal that each station uses to determine slot boundaries. The first bit in the slot is a “busy” bit. It is one when the slot is being

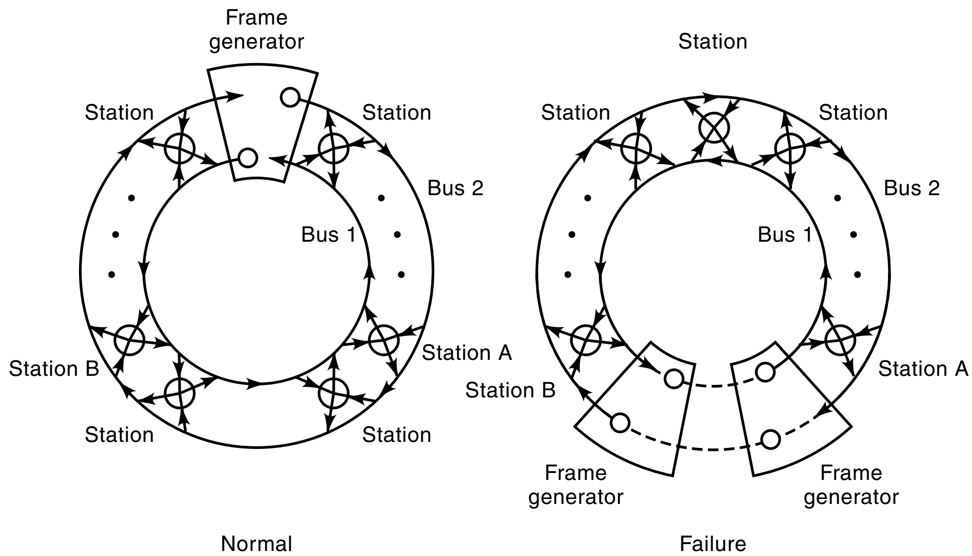

Figure 4. The topology of the DQDB network. This shows dual bus configured as a dual loop in order to survive single failures.

used, and zero when it is empty. When a station has data to send, it writes a one into the busy bit. A read tap precedes the write tap at each station. When a station writes a one into the busy bit, it also reads the bit from upstream to determine if it was already set. If the busy bit was zero, the station transmits data. If the busy bit was one, then the slot is full, but there is no harm in over writing the busy bit.

The problem with this type of a protocol is that stations that are closer to the head-end see more empty slots than stations that are farther away. To prevent unfair access to the medium, stations send reservations to the stations that can acquire slots before they have a chance. Reservation requests are transmitted to the upstream stations on the opposite bus as the data. There are two separate reservation systems, one for transmitting data on each bus. In each system, the bus that is used to transmit data is referred to as the data bus, and the other bus is the reservation bus.

Reservations are used to prevent upstream stations from acquiring all of the empty slots when downstream stations are waiting. A station places reservations in a queue with its own requests and services the entrys in this queue whenever an empty slot passes. If the entry is a reservation, the busy bit in the slot is left unchanged so that the slot is available for a downstream station. If entry is from the local source, the busy bit is set and data are inserted in the slot.

A station with a very large message may still dominate the bus by requesting a large number of slots. To prevent this, a station is only allowed to have one outstanding request at a time. When a message requires several slots, one request is transmitted on the reservation bus and one data slot is placed in the local queue. When the data slot is removed from the queue, a second request is transmitted on the upstream bus and the second data slot is placed in the queue. The process continues until the entire message is transmitted. When there is more than one active user, the active users are given one slot at a time and serviced in a round-robin order.

The queue in each reservation system is implemented with two counters that track the number of reservations. Counter C_B counts the reservations that precede this station's data slot, and counter C_A counts the reservations that follow this station's slot. When a station is not actively transmitting, it increments C_B whenever a reservation is received and decrements it whenever an empty slot passes. When a station has a data slot to transmit, it increments C_A whenever a reservation is received. When C_B is zero and an empty slot appears on the data bus, the station transmits its data slot and then transfers the count from C_A to C_B. If the station has another data slot, transferring the count from C_A to C_B gives downstream stations, which placed reservations while this data slot was queued, a chance to acquire the data bus before this station transmits a second data slot.

In a DQDB network with multiple priority levels for data, there is one reservation bit and two counters for each priority level. When empty slots are received, the counters for the higher priority levels are emptied first.

Protocol Example

Figure 5 shows the operation of five stations in the middle of a DQDB bus as data arrive and are then transmitted. The data bus propagates from left to right and the reservation bus from right to left. The figure shows the value of the busy bit on the data bus and the reservation bit on the reservation bus for each time slot, as the bus passes each station. The values in counters C_A and C_B, and whether or not a station is waiting to transmit data, are shown in between the data and request bus at each station. The operation is simplified and ignores questions of relative timing and propagation delay.

Starting at slot time 1, a single data slot arrives at station 5, followed by two data slots at station 2, and another single data slot at station 3. When the data slots arrive, the network is busy transmitting other data slots, so the data is queued rather than being transmitted immediately. The service order for the messages is station 5, station 2 slot 1, station 3, and station 2 slot 2. The protocol operates as a first in first out (FIFO) queue with a round-robin strategy for multiple slot messages.

In the first three slots, while the reservations arrive, the data bus is busy carrying slots from stations that are closer to the head-end, which were previously queued. Since the stations are upstream, the reservations are not in the queues at these five stations. In the first slot, station 5 inserts a one on the reservation bus that is entered in the queues at stations 1 to 4. In the second slot station 2 places a reservation that is only entered in the queue at station 1. In the third slot station 3 places a reservation that is entered in the queues at stations 1 and 2. After three time slots, station 1 has 3 reservations and station 2 has two reservations, one that will be serviced before its own request and one after. At station 3 there is only one reservation, since station 3 did not receive the reservation from station 2. The fact that station 3 cannot contribute to seeing that station 2 is serviced in its fair turn will not matter, since station 2 has earlier access to the data bus. Stations 4 and 5 have 1 and 0 reservations, respectively.

The reservations are serviced starting in slot 4. Since C_B is greater than zero in stations 1 to 4, these stations let the empty slot pass by, and each removes one reservation by decrementing C_B. Station 5 transmits in slot 4. In slot 5 both stations 2 and 3 are poised to transmit, with C_B at zero and a data slot waiting. Station 2 has the first access at the slot and acquires it, demonstrating that it is unimportant that station 3 did not have an entry for station 2. Station 2 has a second slot that it must transmit, but C_A indicates that one other downstream station has not been serviced. Station 2 moves the count from C_A to C_B before entering its next request in the queue. Slot 6 is acquired by station 3, since C_B is greater than zero at station 2, and slot 7 is acquired by the second slot from station 2.

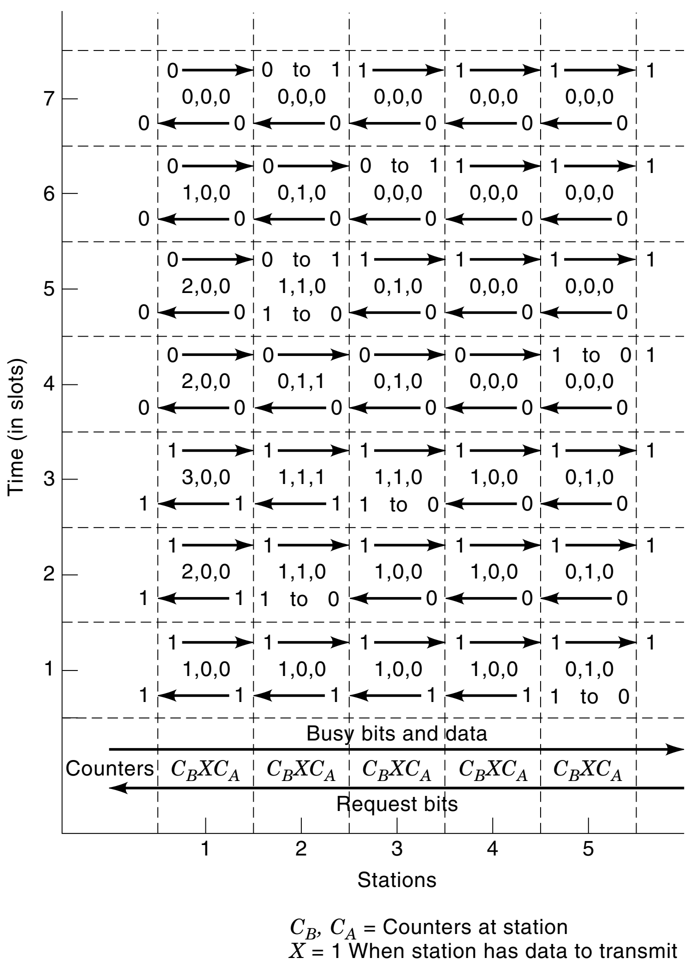

Figure 5. The simplified operation of the DQDB protocol. This shows the operation of the counters at five stations on the bus as data packets arrive at the stations and are transmitted on the bus.

Isochronous Traffic

There is no mechanism in the distributed queueing protocol to provide service guarantees on delay or bit rate. This problem has been sidestepped in the standard by creating a separate protocol that shares the bus with the DQDB protocol, as in FDDI. The slots leaving the head-end are grouped into 125 μs frames. In some slots the busy bit is zero and the slots are available for the data transfer protocol. In other slots the busy bit is one so that stations practicing the data transfer protocol do not try to access these slots.

The busy slots that are generated by the head-end occur at regular intervals and contain a unique header so that they are recognized by stations that require guaranteed rates. These slots are partioned into octets (bytes) that can be reserved. A station that reserves a single octet in a frame acquires a guaranteed 64 kbits/s channel with at most a 125 μs delay. This is the same guarantee provided by the telephone system. As in FDDI, the channels are circuit switched and referred to as isochronous channels.

Protocol Unfairness

Soon after the IEEE 802.6 standard was passed, it was noted that because of the distance-bandwidth product of the network, there was a potential for gross unfairness (12,13). In this work we will explain the source of the problem, and a particularly simple solution, called bandwidth balancing (BWB) (14), which eliminated most of the problem and was added to the standard.

The IEEE 802.6 standard is designed to operate at 155 megabits/s, with 53 byte slots, and is compatible with ATM. At these rates, a cell is only about 0.4 miles long. The standard spans up to 30 miles. Therefore, there may be 75 cells simultaneously on the bus. Assume that a station near the head-end of the bus has a long file transfer in progress when a station 50 cells away requests a slot. In the time it takes the request to propagate to the upstream station, that station transmits 50 slots. When the request arrives, the upstream station lets an empty cell pass and then resumes transmission. An additional 50 slots are transmitted before the empty cell arrives at the downstream station. When the empty slot arrives, the downstream station transmits one slot and submits a request for another. The round trip for this request to get to the upstream station and return an empty slot is another 100 slots. As a result, the upstream station obtains 100 times the throughput of the downstream station.

Although the reason is less obvious, a similar imbalance can occur in favor of the downstream station when that station starts transmitting first. While the downstream station is the only source, it transmits in every cell, while placing a reservation in every slot. When the upstream station begins transmitting, there are no reservations in its counter, but there are 50 reservations on the bus. While the upstream source transmits a slot, a reservation is received. Therefore, the upstream station must allow one slot to pass before transmitting its second slot. During the time it takes to service the reservation and the upstream station's next slot, two more reservations arrive. Therefore, the upstream station lets two empty slots pass before transmitting its third slot. The reservation queue at the upstream station continues to build up each time it transmits a slot, and the upstream station takes fewer of the available slots. The imbalance between the upstream and downstream station is sustained after the 50 reservations pass the upstream station because, at the other end of the bus, the downstream station places a reservation on the bus for each of the empty slot that the upstream station releases. The imbalance is not as pronounced as when the upstream station starts first, but it is considerable. The exact imbalance depends on the distance between two stations and the time that they start transmitting relative to one another (14).

Bandwidth Balancing

The bandwidth balancing (BWB) mechanism, which was added to the standard to overcome the unfairness, is based on two observations:

1. Each station can calculate the exact number of slots that are used, whether or not the data physically pass the station.
2. It is possible to exchange information between stations by controlling the fraction of the slots that are not used.

A station sees a busy bit for every slot transmitted by an upstream station and a reservation for every slot transmitted by a downstream station. By summing the busy bits and reservation bits that apply to a data bus and adding the number of slots that the station transmits, the station calculates the total number of slots transmitted on the bus.

Table 1 shows an example of how stations can communicate and achieve a fair utilization by using the average number of unused slots. Two stations, station A and station B, each try to acquire 90% of the unused bandwidth on a channel. Station A starts first and acquires 90% of the total slots. When station B arrives, only 10% of the slots are available, but station B does not know if the slots are being used by a single station taking its allowed maximum share, or many stations transmitting sporadically. Station B uses 90% of the available slots, or 9% of the slots in the system. Station A now has 91% of the slots available. When station A adjusts its rate to 90% of 91% of the slots, it uses 82% of the slots, making 18% of the slots available to station B. Station B adjusts its rate up to 90% of 18%, which causes station A to adjust its rate down, and so on until both stations arrive at a rate of 47.4%. Note that this mode of communications can only be used when stations try to acquire less than 100% of the slots.

Table 1. Convergence of Rates When Two Stations Use 90% of the Slots Available to Them

Station A		—	Station B	
Measure Bsy + Rqst	Take		Measure Bsy + Rqst	Take
0	$0.9 * 1 = 0.9$		—	—
0	$0.9 * 1 = 0.9$		0.9	$0.9 * 0.1 = 0.09$
0.19	$0.9 * 0.91 = 0.82$		0.82	$0.9 * 0.18 = 0.16$
0.16	$0.9 * 0.84 = 0.76$		0.76	$0.9 * 0.24 = 0.22$
0.22	$0.9 * 0.78 = 0.7$		0.7	$0.9 * 0.3 = 0.27$
0.27	$0.9 * 0.73 = 0.66$		0.66	$0.9 * 0.34 = 0.31$
0.31	$0.9 * 0.69 = 0.62$		0.62	$0.9 * 0.38 = 0.34$
. . .	. . .		. . .	. . .
0.474	$0.9 * 0.526 = 0.474$		0.474	$0.9 * 0.526 = 0.474$

The implementation of BWB in the standard is particularly simple. A station acquires a fraction of the slots available by counting the slots it transmits and placing extraneous reservations in the local reservation queue when the count reaches certain values. In this way, a station lets slots pass that it would have acquired. For instance, if stations agree to take 90% of the slots that are available, they count the slots that they transmit and insert an extra reservation in C_A after every ninth slot that they transmit. As a result, every tenth slot that the station would have taken passes unused.

With BWB, the fraction of the throughput that station i acquires, T_i, is a fraction α of the throughput left behind by the other stations:

$$T_i = \alpha \left\{ 1 - \sum_{j \neq i} T_j \right\}$$

When N stations contend for the channel, they each acquire a throughput:

$$T = \frac{\alpha}{1 + \alpha(N - 1)}$$

The total throughput of the system increases as α approaches one or the number of users sharing the facility becomes large. The disadvantage with letting α approach one is that it takes the network longer to stabilize. We can see from the example in Table 1 that the network converges exponentially toward the stable state. However, as $\alpha \rightarrow 1$ the time for convergence goes to infinity. When $\alpha = 1$, BWB is removed from the network and the original DQDB protocol is implemented.

THE MANHATTAN STREET NETWORK

The Manhattan Street Network (MSN) (15) is a network of two-by-two switches. A source and destination may be attached to each switching node. The logical topology of the network resembles the grid of one-way streets and avenues in Manhattan. Fixed size cells are switched between the two inputs and outputs using a strategy called deflection routing.

The MSN resembles a distributed ATM switch. The two-by-two switching elements may be in a large number of wiring centers rather than a central location. However, the same interconnection structure is used for switching elements in different wiring centers as for elements in the same location. Routing is simpler in the structured MSN than in a general network of small switches.

In deflection routing, packets can be forced to take an available path rather than waiting for a specific path. Each packet between a source and destination is routed individually and may take different paths. The packets may arrive at the destination out of order and have to be resequenced.

Deflection routing has several advantages over virtual circuit routing. The overhead associated with establishing and maintaining circuits is eliminated, and the capacity is shared between bursty sources without large buffers and without losing packets because of buffer overflows. Deflection routing is also being used for routing inside some ATM switches (16,17).

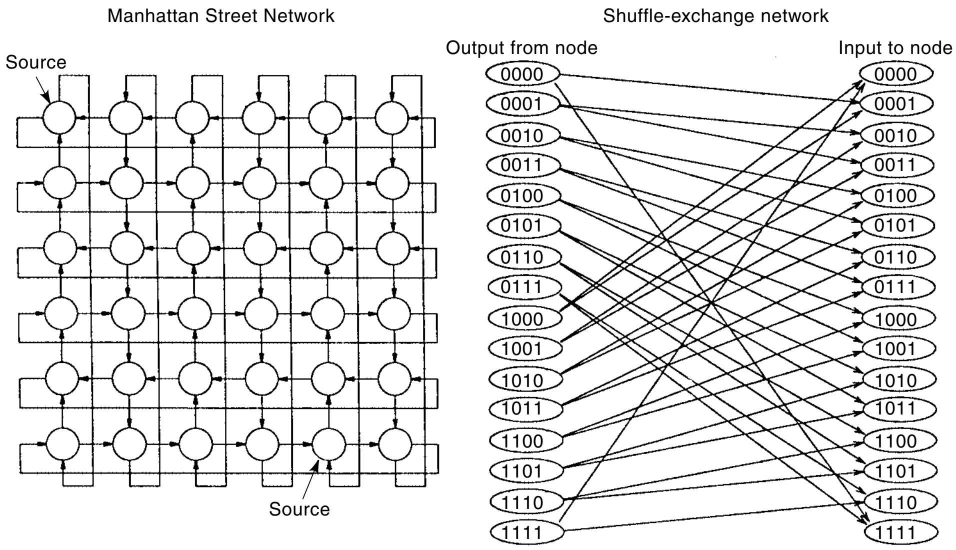

Figure 6. Regular Mesh Topologies. This figure shows the connectivity of the nodes in the Manhattan Street Network and the Shuffle-Exchange Network. The same nodes appear in the left and right columns of the Shuffle–Exchange Network in order to make the regular structure easier to see.

Topology

The MSN is a two-connected topology. In two-connected networks, other than linear structures like the dual bus or dual loop, a path must be chosen at every intermediate node. Two two-connected topologies that have simple routing rules are the MSN and the shuffle-exchange network (15), shown in Fig. 6.

The MSN is a grid of one-way streets and avenues. The directions of the streets alternate. By numbering the streets and avenues properly, it is possible to get to any destination without asking directions or having a complete map.

The grid is logically constructed on the surface of a torus instead of a flat plane. The wraparound links on the torus decrease the distance between the nodes on the edges of the flat plane and eliminate congestion in the corners.

In the shuffle exchange network node i is connected to nodes $2{*}i$ and $2{*}i + 1$, modulo the number of nodes in the network. In the figure for the shuffle exchange network each node appears in both the left- and right-hand column in order to make it easier to draw the connections. The two links leaving each 2×2 switching element are shown in the left-hand column and the two links arriving at each switching element are shown in the right-hand column. The shuffle exchange network has a simple routing rule, based on the address of the destination. The simple routing rule is one of the reasons why this structure is used in most ATM switches.

Deflection Routing

Deflection routing is a rule for selecting paths for fixed-size cells at network nodes with the same number of inputs and outputs, as shown in Fig. 7. The cells are aligned at a switching point. If both cells select the same output, and the output buffer is full, one cell is selected at random and forced to take the other link. The cell that takes the alternate path is deflected. Deflection routing can operate without any buffers in the nodes.

Deflection routing gives priority to cells passing through the node. Cells are only accepted from the local source when there are empty cells arriving at the switch. Cells are never dropped due to insufficient buffering because the number of cells arriving at the node never exceeds the number of cells that can be stored or transmitted. There is an implicit assumption that the source can be controlled. This assumption is common in data networks with variable throughputs, such as the Ethernet.

The MSN is well suited for deflection routing for three reasons:

1. At any node many of the destinations are equidistant on both output links. Cells headed for these destinations have no preference for an output link and do not force other cells to be deflected.
2. When a cell is deflected, only four links are added to the path length. The worse that happens is that the cell must travel around the block.

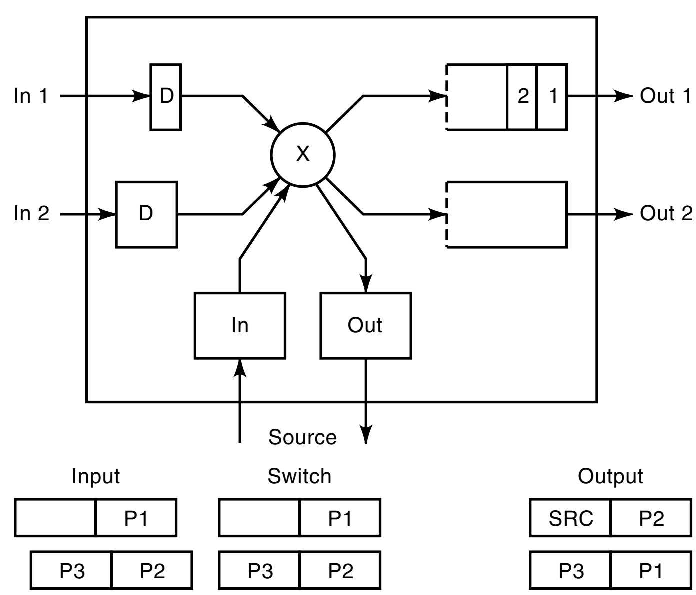

Figure 7. A block diagram of a node in a deflection routing network. The arriving packets are delayed so that they aligned. The packets can be exchanged and placed in their respective output buffers.

3. Deflection routing can guarantee that cells are never lost, but cannot guarantee that they will not be deflected indefinitely and never reach their destination. It has been proven that this type of livelock will never occur in the MSN (18).

Deflection routing is similar to the earlier hot potato routing (19), which operated with variable-size packets, on a general topology, with no buffers. Fixed-size cells, the MSN topology, and two or three cells of buffering converted the earlier routing strategy, which had very low throughputs, to a strategy that can operate at levels exceeding 90% of the throughput that is achieved with infinite buffering.

Reliability

The MSN topology has several paths between each source and destination. The alternate paths can be used to communicate after nodes or links have failed.

There are two simple mechanisms to survive failures in the MSN. Both mechanisms, shown in Fig. 8, are adopted from loop networks. Node failures are bypassed by two normally closed relays that connect the rows and columns through. The missing node in the grid in Fig. 8 has failed. Link failures are detected by a loss of signal, as in loop networks. Nodes respond to the loss of signal by not transmitting on the link at right angles to the link that has stopped. When one link fails, three other links are removed from service and the node at the input to the failed link stops transmitting on it. The dotted link in Fig. 8 has failed and nodes stop transmitting on the dashed links. This link removal procedure works with any number of link failures.

Since the number of input and output links are equal at all of the operating nodes, deflection routing continues to operate without losing cells. In addition, it has been found that the simple routing rules that are designed for complete MSNs continue to work on networks with failures.

COMPARISON OF FDDI, DQDB, AND THE MSN

The FDDI, DQDB, and MSN networks are all two-connected topologies. There are two links entering and leaving each node. Some units in the FDDI network may not be two-connected, but the main part of the network is a dual loop. Two-connectivity distinguishes these networks from the earlier loop and bus networks, used in LANs, which were one-connected. The increased connectivity makes these networks better suited for the increased number of users and longer distances spanned in a MAN.

Like the earlier LANs, DQDB and FDDI are linear topologies. Logically, the nodes are arranged in a one-dimensional line. In linear topologies the throughput per user and reliability are not as high as they are in other two-connected topologies, which have shorter distances and more paths between nodes. There were early proposals (20) to connect the second paths in loop networks to achieve higher throughputs and reliability than the bidirectional loop, but these more complicated topologies lost some of the more desirable characteristics of loops and buses. The MSN and shuffle-exchange topology achieve both goals while approaching the ease of routing and growth that is associated with linear topologies.

Throughput

A disadvantage with linear topologies is that the average throughput that each user can obtain decreases linearly with the number of users. In linear topologies, the average distance between nodes increases linearly with the number of nodes in the network. With uniform traffic, the number of users sharing a link increases linearly with the number of users in the network and the average rate per user decreases accordingly. The protocols used in FDDI and some DQDB networks do not reuse slots. The throughput in these networks is reduced for all traffic distributions.

By contrast, in the MSN the distance between nodes increases as the square root of the number of nodes in the network, and in the shuffle exchange the distance increases as the log of the number of nodes. As a result, the reduction in the throughput per user, which occurs as networks become large, is much less in the MSN and the shuffle-exchange networks than in the FDDI or DQDB network.

In the DQDB network, the penalty for large networks can be reduced by breaking the network into segments and erasing data that have already been received when they reach the end of a segment. This strategy works particularly well when the traffic requirements are nonuniform. When there are communities of users that communicate frequently, if those users are placed on the same segment of the bus, the traffic between them does not propagate outside the segment and interfere with users in other segments. A similar strategy for reusing capacity does not exist for FDDI. In addition, the concept of communities is not as meaningful in FDDI. FDDI operates as a single loop. When user A has a short path to user B, user B

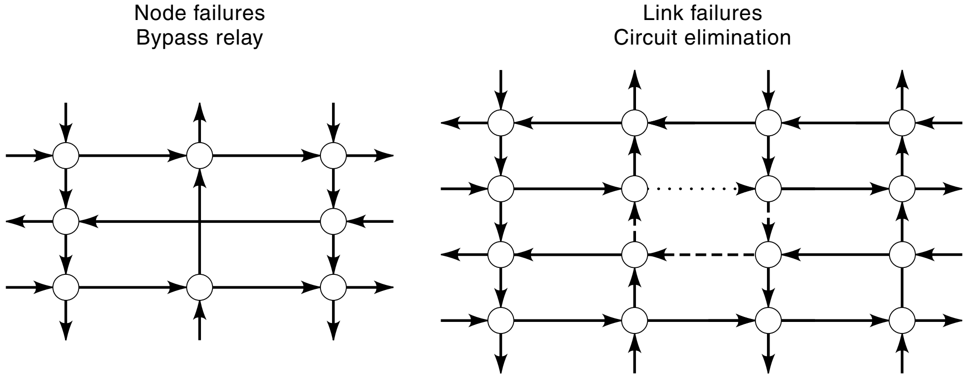

Figure 8. Failure recovery mechanisms in the MSN. Node failures are survived with bypass relays at the failed node. Link failures are survived by eliminating a circuit that includes the failed link, thus preserving the criterion necessary to perform deflection routing at every node.

must traverse the rest of the loop to get to user A, and therefore has a long path.

In the MSN and shuffle-exchange network special erasure nodes are not needed because the protocol removes cells that reach their destination. In the MSN, communities of users are supported in a very natural way. If nodes that communicate frequently are located within a few blocks, they only traverse the paths in those few blocks and do not affect the rest of the network. The concept of communities does not exist in shuffle-exchange networks.

There is less interference between neighborhoods in the MSN than in DQDB networks. In the MSN, when there is heavy traffic within a neighborhood, communications between other neighborhoods can continue without passing through that neighborhood. In deflection routing, because of the random selection process when there are equal length paths or oversubscribed nodes, cells naturally avoid passing through congested neighborhoods. By contrast, when a community in the middle of the bus becomes congested in a DQDB network, communications between nodes at opposite edges of the bus must still pass through that community.

Reliability

Both DQDB and FDDI have a structure that enables them to survive single failures without losing service. FDDI has a redundant loop that is pressed into service to bypass single failures, and DQDB repositions the head-end of the bus. In both of these networks if two or more failures occur in adjacent components, the mechanism bypasses those components without affecting the operating components. However, in the more likely event that the multiple failures occur in components that are separated from one another, the network is partitioned into islands of nodes that cannot communicate with one another.

Nodes in the MSN are not cut off from one another until at least four failures occur. When four failures have occurred, the likelihood of nodes being disconnected, and the number that are actually disconnected, is small. A quantitative comparison of the MSN, DQDB, and FDDI networks is presented in Ref. 21.

Routing Complexity

An advantage of linear topologies is that routing is relatively simple. All of the data that enter an FDDI system are transmitted on a single path, and there is only one path to select at any intermediate node. In a DQDB network the source must decide which of the two paths leads to the destination, but once the data are in the network there are no choices to make. The MSN and shuffle-exchange networks have simple rules to select a path, but a choice must be made at each node.

In a linear topology everything that is destined for an output link either originates at the source at the node or is received from a single input link. When the data that arrive on the input link have precedence over the data from the source, there is no need to delay or buffer the data from the input link. Deflection routing has the same result in any network where the in-degree equals the out-degree at each node.

Growth

An important consideration in any large network is how easily it can be modified to add or delete users. In the early days of LANs the correspondence between the topology of the network and the physical distribution of users was considered important. A loop network was supposed to be a daisy chain between adjacent offices and a bus network was supposed to pass down a hallway. With experience we realized that it is more difficult to change the wiring between offices as users move than to have all offices connected to a wiring cabinet and change the interconnection in that cabinet. As a result, most LANs are physically a star network, or a cluster of stars, of connections between users and a wiring cabinet. The users are connected together inside the wiring cabinet to form a logical loop or bus or mesh network. The number of wires that must be changed in the wiring cabinet determines how difficult it is to add or delete users from a network.

In a bidirectional loop or bus network, adding or deleting a user is a relatively simple operation. To add a user, the connection between two users is broken and the new user is inserted between them. In the wiring cabinet, two wires are deleted and four are added. In the FDDI loop, there is only one path to the new user. In the DQDB network, every user must determine which bus to use to transmit to the new user.

In the shuffle-exchange network, adding or deleting nodes is very complex. Complete networks are only defined for certain numbers of users. The shuffle–exchange network shown in Fig. 6 is only defined when the number of users is 2^n. When a network is replaced by the next larger network, virtually every wire must be removed and the network reconfigured from scratch.

In a complete MSN, two complete rows or columns must be added to retain the grid structure. There are, however, partial MSNs in which rows or columns do not span the entire grid, and a technique is known for adding one node at a time to a partial MSN to construct eventually a complete MSN (22). With this technique the number of links that must be changed in the wiring cabinet is the same as in the loop or bus network. In addition, an addressing scheme is known that allows new partial rows or columns to be added without changing the addresses of existing nodes or the decisions made in the routing strategy.

Multimedia Traffic

Multimedia traffic is playing an increasingly more important role in networks. Non-real-time video or audio, as is currently delivered by the Web sites on the Internet, is adequately handled by the data modes on any of the MANs that have been described. However, the data mode of operation cannot provide the guarantees on throughput or delay that are required for real-time voice or video applications; nor can it guarantee the sustained throughput that is needed to view movies while they are being received. DQDB and FDDI have an isochronous mode of operation that provides dedicated circuits to support real-time traffic.

The isochronous mode is well integrated into the DQDB protocol. Nodes that only require the data mode do not have to change any protocols or hardware when isochronous traffic is added to the network. The only change that these nodes notice is that some slots seem to enter the network in a busy state. By contrast, when isochronous traffic is added to an FDDI network, every node must be able to perform context switches to move between the data and circuit modes. The MSN does not have an isochronous mode of operation. It is unlikely that dedicated circuits can be added to the mesh structure without constructing a separate circuit switch at every node in the network.

CATV

The current CATV network is an existing MAN that delivers TV programs to a large number of homes. The network is designed for unidirectional delivery of the same signal to a large number of receivers. The channels have a very wide bandwidth relative to telephone channels. In bidirectional CATV systems there are many more channels from the head-end to the home than in the opposite direction.

The growth of the use of the Web in homes has created the need for wide-band channels to homes. The traffic on the Web is predominantly from servers to clients, and most homes are clients rather than servers. Increasingly, the Internet is being used for multicast communications (23) of video or audio programming, in which the same signal is received by more than one receiver. CATV networks are naturally suited for the Web and multicast communications.

The simplest CATV systems are hybrid networks that use the CATV network to send addressed packets to the home, and the telephone network to receive data from the home. The CATV network provides a means of quickly getting a large amount of data to the home. The home terminals share a CATV channel. They receive all of the packets and filter out the packets that are intended others. If the data are considered sensitive, encryption is used to keep sensitive data from being intercepted. Shared CATV channels are particularly useful when high data rates are needed for short periods, as when receiving new pages from Web servers. The telephone channel is a relatively low bandwidth channel that adequately handles the traffic to Web servers. The IEEE 802.14 working group is currently considering standards for hybrid networks.

The application of the Internet is not stationary. Packet telephony or an increase in the number of servers in homes for publishing or small businesses can quickly change the current unidirectional traffic demands. Data MANs that are implemented on CATV networks should be flexible so that they can track changes in the applications. In the early MANs, experimental CATV systems were used for two-way voice and data (24). There was not sufficient demand for data to homes at that time, and work on these networks stopped. The renewed interest in data applications of CATV networks makes it reasonable to reconsider the earlier work. In this section, we describe one such network, the Homenet (25).

Homenet

The Homenet transmission strategy partitions the CATV network into smaller areas, called Homenets. Stations only contend with stations in their own Homenet to gain access to the network. Because of the size of the Homenets, the contending stations satisfy the distance constraints imposed on the CSMA/CD protocol used in Ethernet LANs. Any station on the CATV network can receive the signal on any Homenet. Two stations communicate by transmitting on their own Homenet and receiving on the other stations Homenet.

The Homenet strategy is readily tailored to load imbalances, such as the directional imbalance in Web traffic. There is no relationship between the bandwidth that is available on a station's transmit and receive channel. Bandwidth is assigned to transmitters by adjusting the number of stations that they contend with for a Homenet. For instance, an Internet service provider (ISP) can be given one or more Homenets, so that traffic from the servers can access the network without contention. Many clients, in homes, may share the same Homenet, since the traffic level from a client to the servers is much lower.

A convenient characteristic of the Homenet strategy, in comparison with hybrid networks, is that the network can be modified easily to match the changes in load imbalances. If the traffic distribution changes and more traffic originates in some homes, those homes can be placed on less heavily populated Homenets to increase the amount that they can transmit.

Access Strategy. The stations in a Homenet access the channel using Movable Slot Time Division Multiplexing (MSTDM) (26), which is a variation on the CSMA/CD (Carrier Sense Multiple Access/Collision Detection) protocol used in Ethernet. MSTDM is implemented by a very small change in the standard Ethernet access unit, and sets up telephone quality voice connections on an Ethernet.

Figure 9 shows the transmission strategy that is used within a Homenet. The CATV taps are directional so that a station only receives signals from downstream and can only transmit upstream. The stations in a Homenet transmit upstream in an assigned frequency band. At a reflection point the upstream frequency band is received and transmitted downstream in a different frequency band. Before transmitting, a station listens to the downstream channel to determine if it is busy; CSMA then transmits on the upstream channel. When a station receives the signal that it transmitted, it knows that it has not collided with any other station, CD, so that any station that receives this channel can receive its data.

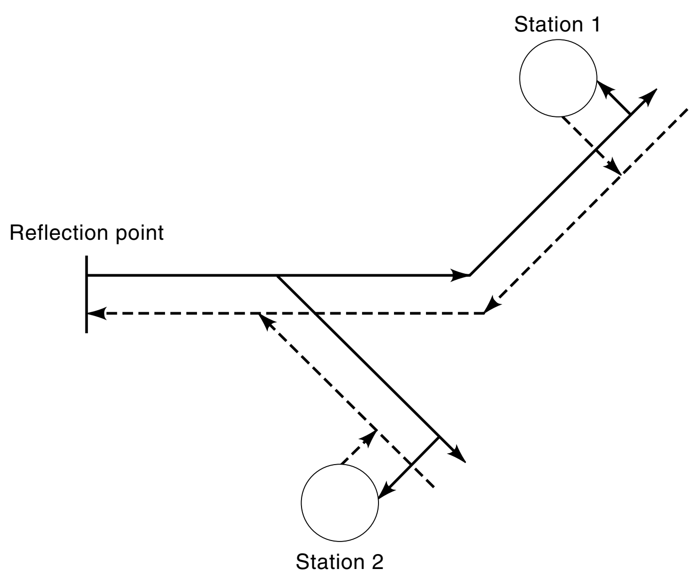

Figure 9. Transmission strategy within a Homenet. Stations transmit upstream on the CATV tree using one channel. At the root of the homenet the signal that is received on this channel is retransmitted downstream on a different channel.

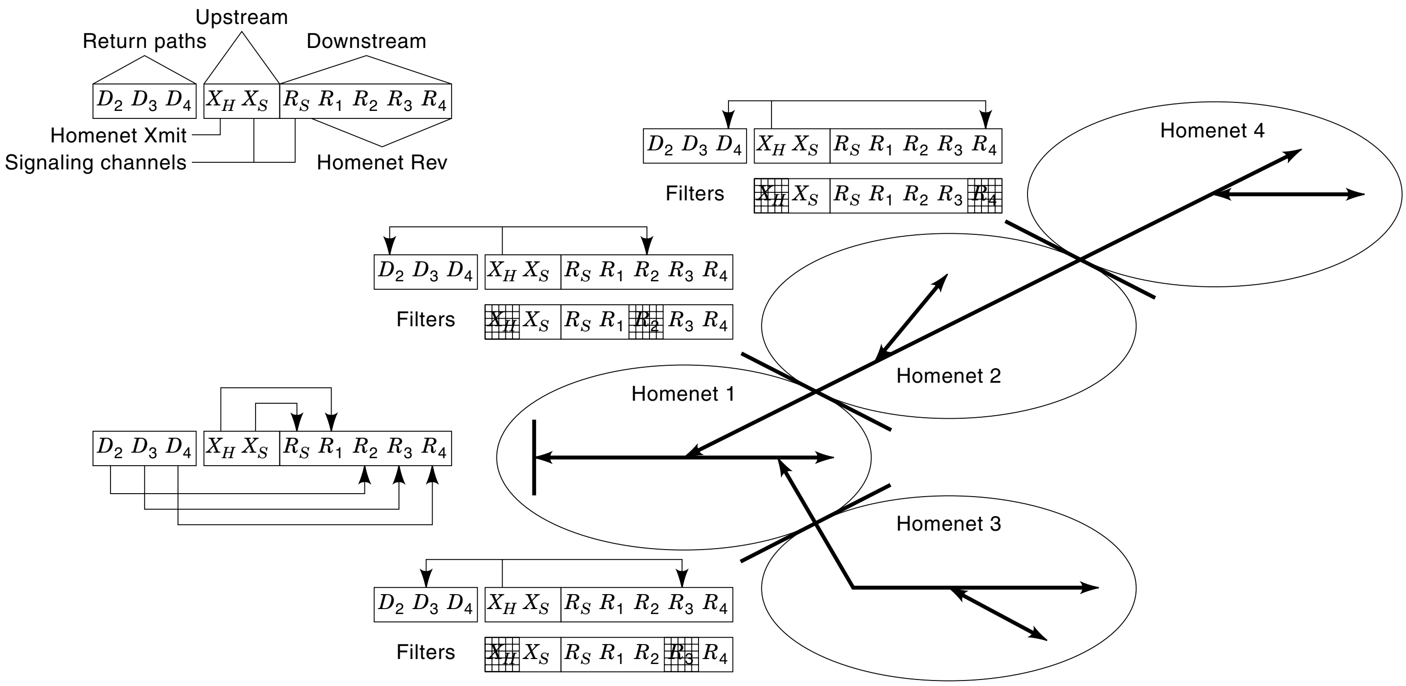

Figure 10. Transmission plan for a CATV network with 4 Homenets. The *X*'s are upstream, transmit channel and the *R*'s are downstream, receive channels. The cross-hatched channels are filtered out at the edge of the homenets.

Interconnection Strategy. Figure 10 shows the boundaries of Homenets, the assignment and filtering of frequencys, and the interconnections for a four-Homenet system. There are many fewer upstream channels in a CATV network than downstream channels. The same upstream data channel is reused in every Homenet by filtering that channel at the boundaries between Homenets. The stations in every Homenet contend for the same upstream channel, but the transmissions from stations in different Homenets do not interfere with one another.

Each Homenet is assigned a unique receive channel. At the Homenet's reflection point the upstream signal on the transmit channel is received and retransmitted on the Homenet's receive channel. The signal is also carried by a point-to-point link to the head-end of the CATV network, where it is again transmitted on the Homenet's receive channel. A Homenet's receive channel is filtered at the reflection point for the Homenet so that the signal from the head-end does not interfere with the signal that is inserted. The reflection point of each Homenet appears to be the root of the entire CATV tree for that Homenet's receive channel, and any station on the CATV network can receive the signal that a station transmits in its Homenet.

To see how the transmission strategy works, consider a station in Homenet 2. The station transmits on the common upstream channel. This transmission may collide with transmissions from other stations in Homenet 2. The transmission will not collide with transmissions from stations in Homenet 4 because the transmit channel is filtered at the boundary between Homenets 2 and 4. At the reflection point for Homenet 2, the signal on the transmit channel is placed in the downstream channel 2. A station that has transmitted detects collisions by listening to the downstream signal on channel 2. A successfully transmitted packet on Homenet 2 can be received by any station in Homenets 2 and 4. The signal on the transmit channel is also transferred to the head-end of the CATV network, where it is placed in channel 2, so that stations in Homenets 1 and 3 can also receive the signal from the station in Homenet 2. The signal from the head-end does not interfere with the signal inserted in channel 2 at the reflection point because the signal from the head-end is removed at the entry to Homenet 2.

In Fig. 10 there is a second upstream and downstream channel that is used by all of the stations on the CATV network. This is a signaling channel and is used for communications between stations during call setup. A station places a call by transmitting on the upstream signaling channel. At the head-end of the CATV network this signal is placed on the downstream signaling channel. The CSMA/CD contention rule cannot be used on this channel because of the larger distances involved, but the low utilization of the channel makes it reasonable to use a less efficient Aloha (27) contention rule. A station does not listen to the channel before transmitting, but does listen to the receive channel to determine if its signal collided with the signal from another station. If a station can receive its own signal, so can all other stations. When a station is not busy, it listens to the downstream signaling channel and responds to a connect request.

CONCLUSION

The principal reasons for using different technologies in different area networks are economic. As MANs have developed, the economics have changed, and continue to change. The initial application of MANs was interconnecting LANs in a city.

When the economics changed, WAN technology was used for this function.

At present there is a growing demand for wider bandwidth channels to individual homes in a MAN. The bandwidth is not available. In the short term we should expect the bandwidth to be made available using existing networks, either the CATV infrastructure or ADSL technology over the telephone local loop. As the demand continues to grow, additional capacity will be installed. Until the demand reaches a level where individual fibers to a home are justified, channel sharing techniques, such as FDDI, DQDB, and the MSN, are likely to be used to share the added capacity.

Just as the capabilities of WANs have affected MANs, the capabilities of MANs can affect LANs. An interesting question is whether or not the future growth of MANs will lead to the end of LANs. When inexpensive, wide-bandwidth channels are available from every desk to the telephone central office or an ISP, will it still be economical for companies to install and maintain private LANs?

BIBLIOGRAPHY

1. E. W. Zegura, Architectures for ATM switching systems, *IEEE Comm. Mag.,* **31** (2): 28–37, 1993.
2. F. E. Ross, An overview of FDDI: The fiber distributed data interface, *IEEE JSAC,* **7** (7): 1043–1051, 1989.
3. W. Stallings, *Local and Metropolitan Area Networks,* Upper Saddle River, NJ: Prentice Hall, 1997.
4. M. J. Johnson, Proof that timing requirements of the FDDI token ring protocol are satisfied, *IEEE Trans. Comm.,* **COM-35**: 620–625, 1987.
5. H. E. White and N. F. Maxemchuk, An experimental TDM data loop exchange, *Proc. ICC '74,* June 17–19, 1974, Minneapolis, MN, pp. 7A-1–7A-4.
6. R. M. Newman, Z. L. Budrikis, and J. L. Hullett, The QPSX man, *IEEE Comm. Mag.,* **26** (4): 20–28, 1988.
7. R. M. Newman and J. L. Hullett, Distributed queueing: A fast and efficient packet access protocol for QPSX, *Proc. 8th Int. Conf. on Comp. Comm.,* Munich, F.R.G., Sept. 15–19, 1986, published by North-Holland, pp. 294–299.
8. L. Fratta, F. Borgonovo, and F. A. Tobagi, The Express-Net: A local area communication network integrating voice and data, *Proc. Int. Conf. Perf. Data Commun. Syst.,* Paris, Sept. 1981, pp. 77–88.
9. J. O. Limb and C. Flores, Description of Fasnet—A unidirectional local area communications network, *BSTJ,* **61** (7): 1413–1440, 1982.
10. M. Zukerman and P. G. Potter, A protocol for Eraser Node Implementation within the DQDB framework, *Proc. IEEE GLOBECOM '90,* San Diego, CA, Dec. 1990, pp. 1400–1404.
11. M. W. Garrett and S.-Q. Li, A study of slot reuse in dual bus multiple access networks, *IEEE JSAC,* **9** (2): 248–256, 1991.
12. J. W. Wong, Throughput of DQDB networks under heavy load, *EFOC/LAN-89,* Amsterdam, The Netherlands, June 14–16, 1989, pp. 146–151.
13. J. Filipiak, Access protection for fairness in a distributed queue dual bus metropolitan area network, *ICC '89,* Boston, June 1989, pp. 635–639.
14. E. L. Hahne, A. K. Choudhury, and N. F. Maxemchuk, Improving the fairness of distributed-queue dual-bus networks, *INFOCOM '90,* San Francisco, June 5–7, 1990, pp. 175–184.
15. N. F. Maxemchuk, Regular mesh topologies in local and metropolitan area networks, *AT&T Tech. J.,* **64** (7): 1659–1686, 1985.
16. S. Bassi et al., Multistage shuffle networks with shortest path and deflection routing for high performance ATM switching: The open loop shuffleout, *IEEE Trans. Commun.,* **42**: 2881–2889, 1994.
17. A. Krishna and B. Hajek, Performance of shuffle-like switching networks with deflection, *Proceedings INFOCOM '90,* June 1990, pp. 473–480.
18. N. F. Maxemchuk, Problems arising from deflection routing: Live-lock, lockout, congestion and message reassembly, In G. Pujolle (ed.), *High Capacity Local and Metropolitan Area Networks,* Springer-Verlag, 1991, pp. 209–233.
19. P. Baran, On distributed communications networks, *IEEE Trans. Comm. Sys.,* **cs-12**, 1–9, 1964.
20. C. S. Raghavendra and M. Gerla, Optimal loop topologies for distributed systems, *Proc. Data Commun. Symp.,* 1981, pp. 218–223.
21. J. T. Brassil, A. K. Choudhury, and N. F. Maxemchuk, The Manhattan Street Network: A high performance, highly reliable metropolitan area network, *Computer Networks and ISDN Systems,* 1994.
22. N. F. Maxemchuk, Routing in the Manhattan Street Network, *IEEE Trans. Commun.,* **COM-35**: 503–512, 1987.
23. S. Deering, Multicast routing in internetworks and extended LANs, *Proceedings of ACM SIGCOMM '88,* Aug. 1988, Stanford, CA, pp. 55–64.
24. A. I. Karchmer and J. N. Thomas, Computer networking on CATV plants, *IEEE Network Mag.,* pp. 32–40, 1992.
25. N. F. Maxemchuk and A. N. Netravali, Voice and data on a CATV network, *IEEE J. Sel. Areas Commun.,* **SAC-3** (2): 300–311, 1985.
26. N. F. Maxemchuk, A variation on CSMA/CD that yields movable TDM slots in integrated voice/data local networks, *BSTJ,* **61** (7): 1527–1550, 1982.
27. N. Abramson, The Aloha system—Another alternative for computer communications, *Fall Joint Computer Conference, AFIPS Conference Proceedings,* **37**, pp. 281–285, 1970.

N. F. MAXEMCHUK
AT&T Labs–Research

MHDCT. See HADAMARD TRANSFORMS.
MHD POWER PLANTS. See MAGNETOHYDRODYNAMIC POWER PLANTS.
MICROBALANCES. See BALANCES.
MICROCOMPUTER. See MICROPROCESSORS.