UC 405

LBL-29107
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Information and Computing Sciences Division

Submitted to Computer Communication Review

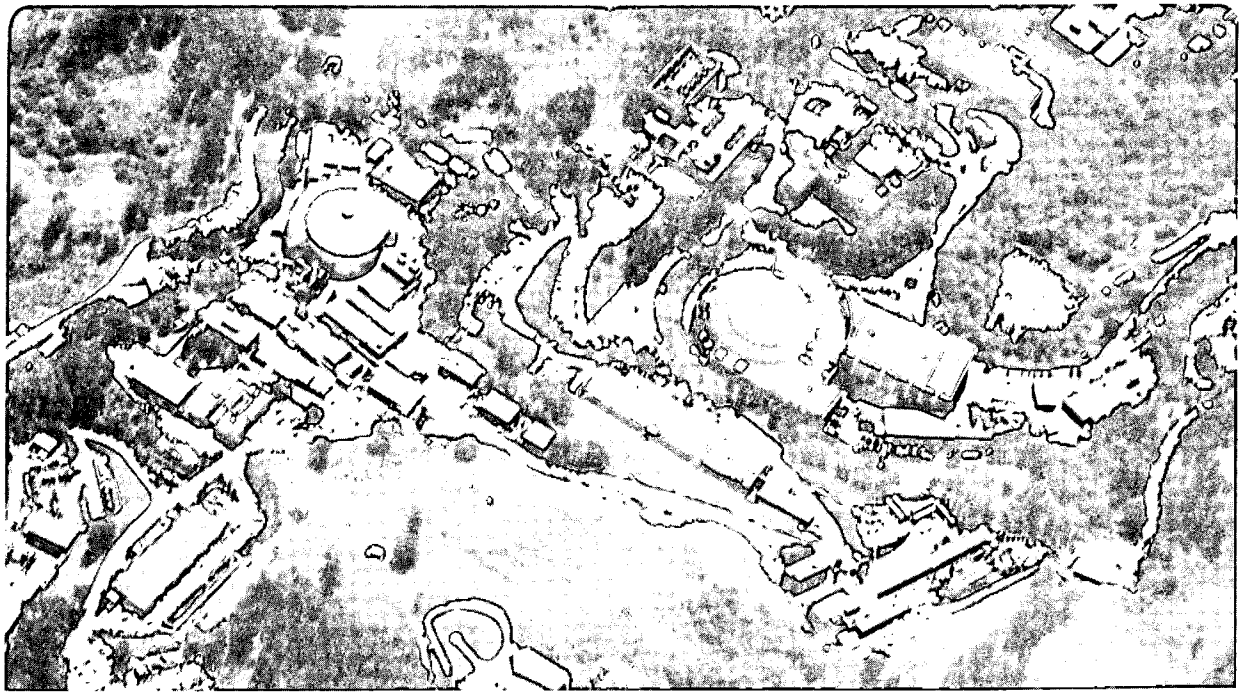
FDDI — A LAN among MANs

F.E. Ross, J.R. Hamstra, and R.L. Fink

May 1990

For Reference

Not to be taken from this room



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

LBL-29107
Copy 1
Bldg. 50 Library.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

FDDI - A LAN Among MANs

Floyd E. Ross
Unisys

James R. Hamstra
StanaTek

Robert L. Fink*
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Abstract - FDDI (Fiber Distributed Data Interface) is a 100 Mbit/sec local area network (LAN) based on a token ring protocol operating over an optical fiber medium. Though not a metropolitan area network (MAN), FDDI can operate over distances normally considered as the province of MANs. With the addition of FDDI-II capabilities, and the newly defined mapping for FDDI over SONET, FDDI now offers MAN functionality for large organizations needing to provide integrated services over a widely based high performance private network. This paper updates earlier efforts on this subject, giving an up-to-date overview of FDDI, FDDI-II and SONET mapping. At the end of the paper a brief look at the possible future for FDDI technology is given.

1. AN OVERVIEW OF FDDI

The Fiber Distributed Data Interface (FDDI) is a 100 megabit-per-second (Mbps) Local Area Network (LAN) that uses optical fiber as the medium. The FDDI protocol is based on a token ring access method. FDDI is being developed as an American National Standard (ANS) in Accredited Standards Committee (ASC) X3T9. FDDI is also being developed as an International Standard (IS) in ISO/IEC JTC1/SC 25.

The basic FDDI uses optical fiber with Light Emitting Diodes (LEDs) transmitting at a nominal wavelength of 1325 nanometers over multi-mode fiber links up to 2 Km. Connections between stations are made with a dual fiber cable employing a polarized duplex connector. A single-mode fiber enhancement to FDDI uses laser diode transmitters to allow individual links to be extended up to at least 60 Km on private fiber. For environments where a private LAN must be extended across public links, an FDDI-to-SONET mapping will provide transport for FDDI over SONET common carrier facilities.

FDDI data packets (frames) use the 48-bit address structure defined in the IEEE P802 LAN protocols. The assignment of unique addresses to stations is administered by the IEEE Standards Office, with IEEE 802 stations and FDDI stations sharing a common address space. Locally administered addresses are also permitted. The nature of the clocking, which requires an adjustment for the accumulated jitter between frames, limits frames to 4500 octets maximum. Multiple frames may, however, be transmitted on the same access opportunity.

The FDDI instantaneous data transmission rate is 100 Mbps. The effective sustained data rate at the data link layer can be well over 95 percent of this peak rate. The four-to-five group code used on the optical fiber medium requires a 125 megabaud transmission rate.

A maximum ring configuration of 500 stations and 100 kilometers of duplex cable has been used as the basis for calculation of the default values of the recovery timers. The choice of larger values for the recovery timers will allow larger networks to be configured; smaller values will reduce recovery time for smaller networks.

1.1 Organization Of FDDI Standards

FDDI is most easily described in terms of its component standards. There are seven standards in all. These are being developed in conformance with the OSI reference model and the OSI management

* This work was supported in part by the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

framework and layer management guidelines. The basic FDDI, when completed, will consist of the following set of four standards:

PMD	A Physical Layer Medium Dependent standard which specifies the optical fiber link and related optical components. (Reference 3.)
PHY	A Physical Layer Protocol standard which specifies the encode/decode, clocking, and data framing. (Reference 2.)
MAC	A Media Access Control standard which specifies access to the medium, addressing, data checking, and frame generation/reception. (Reference 1.)
SMT	A Station Management standard which specifies the FDDI station configurations, ring configurations, and the control required for proper operation of stations in an FDDI ring. (Reference 6.)

The following standard entities are being developed as extensions to the basic FDDI:

SMF-PMD	A Single-Mode Fiber (SMF) version of the PMD standard which provides an alternate to the basic PMD, increasing the permissible fiber links from 2 to 60 kilometers in length. (Reference 6.)
SPM	An FDDI-to-SONET Physical Layer Mapping Function standard which provides an alternate to the basic PMD, providing a transport for FDDI over SONET common carrier facilities. (Reference 7.)
HRC	A Hybrid Ring Control standard which specifies an upward-compatible version of FDDI, commonly known as FDDI-II, which adds the capability for circuit switched services to the packet services of the basic FDDI, thus creating an integrated services LAN. (Reference 4.)

Figure 1 provides an overview of each of these as they relate to the OSI reference model. FDDI MAC provides a superset of the services required by the Logical Link Control (LLC) protocol developed by IEEE P802.2. Figure 1 depicts the use of the IEEE 802.2 LLC as the upper sublayer of the data link layer. Any other appropriate protocol may be used.

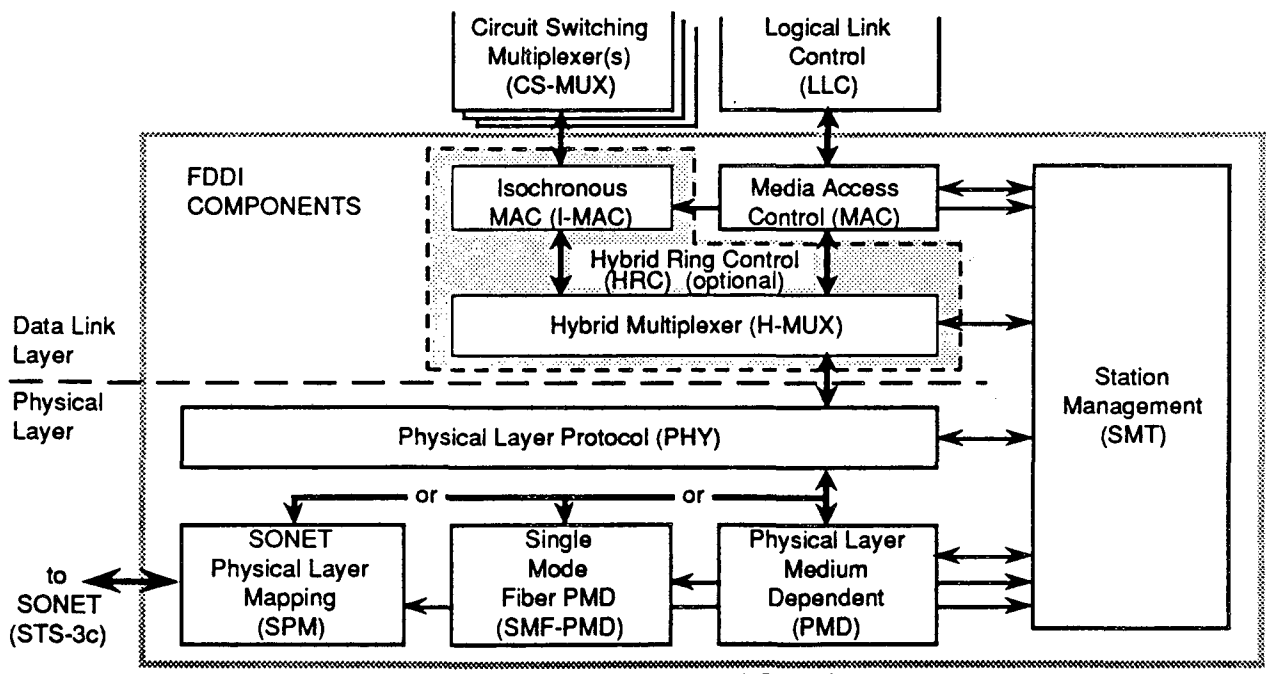


Figure 1 - FDDI Architectural Overview

2. FDDI NETWORK CONFIGURATIONS

2.1 Station Types

FDDI defines two classes of stations, based on whether or not they are intended to attach directly to the dual (counter-rotating) trunk ring. A dual attachment station (DAS) has two ports (PHY/PMD pairs), designated as peer A and B ports. It can attach directly into the trunk ring (the trunk ring is two logical rings). A dual attachment station has one or more MAC entities. Figures 2 and 3 are examples of dual attachment stations having a single MAC (SM) and dual MACs (DM).

Another method of dual ring attachment is provided by a concentrator. A concentrator has additional ports beyond those required for its own attachment to the FDDI ring. These additional ports, designated as master (M) ports, provide for the attachment of additional slave stations. These slave stations, while logically part of the ring, are physically isolated from the trunk rings by the concentrator. Varying levels of functionality, including multiple MACs, are permitted in concentrators. A dual attachment concentrator (DAC) can attach directly to the trunk ring, and could provide the capability to insert its slave stations into either, or both, of the logical rings provided by the trunk ring. A concentrator would configure its MAC(s) downstream of any inserted slave stations in either ring.

Figure 5 depicts a single attachment station (SAS). It has one port, designated as a slave (S) port, and therefore would not normally attach directly into the trunk rings. Instead, it would be attached to the ring via the M port of a concentrator.

Additional station types beyond these examples are permitted. As an example, a single attachment concentrator (SAC) is a concentrator that would attach to the trunk rings via another concentrator.

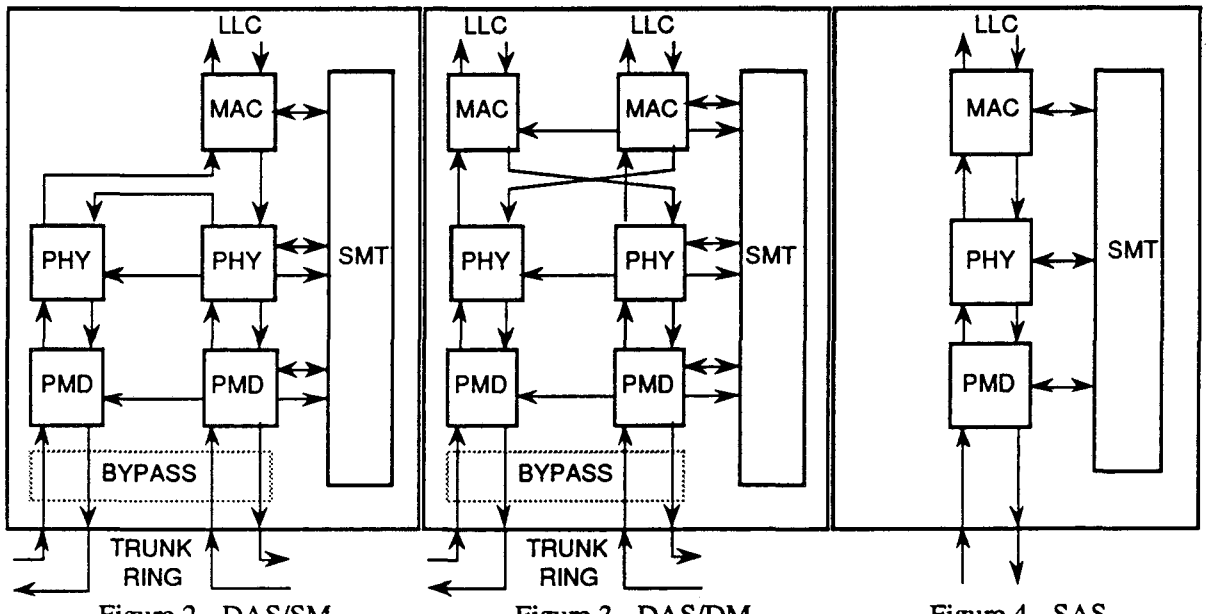


Figure 2 - DAS/SM Figure 3 - DAS/DM Figure 4 - SAS

2.2 Network Topologies

The baseline FDDI network topology is a dual counter-rotating ring of trees, *i.e.* one pair of counter-rotating logical rings (the trunk ring) with trees of cascaded concentrators and slave stations (see Figure 5). Only one dual trunk ring is allowed whereas there may be multiple trees of varying depths attached to it. All subsets of this topology are supported.

In configuring the dual trunk ring, each optical fiber connection is intended to have an A port at one end and a B port at the other. In configuring a concentrator subtree, each optical fiber connection is intended to have an M port at one end and a A, B or S port at the other end, depending on the slave station type.

Stations with multiple MACs may configure them in both of the counter rotating rings, or all MACs may be in the same ring. Direct communication between MACs in different rings is not possible unless the ring has been reconfigured to the wrap state due to some fault. Thus both MACs may have independent access to a full 100 Mbps data capability in a normal configuration.

A dual attachment station performs as a pair of single attachment stations insofar as its attachment to the ring via a concentrator is concerned. Operation of both ports of a dual attachment station in slave mode is allowed; however no interconnection is permitted between the two ports within the station, to enforce the logical tree topology.

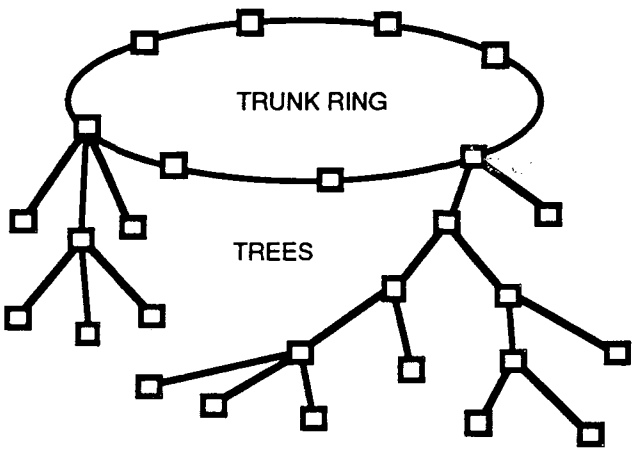


Figure 5 - Ring of Trees

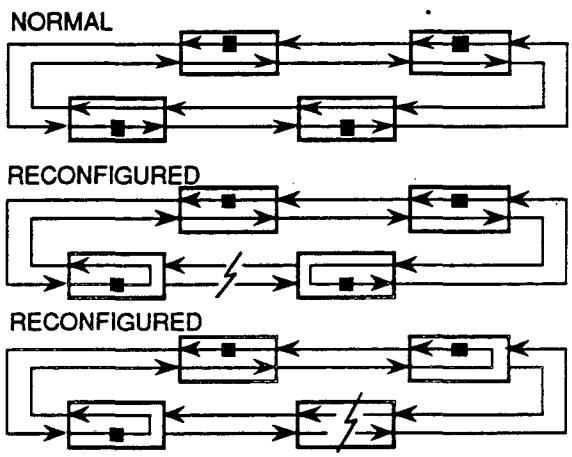


Figure 6 - Reconfiguration

2.3 Reliability Considerations

Ring topologies allow for the isolation of failing attachments through several mechanisms. Counter-rotating rings are basic to the FDDI structure. The counter-rotating trunk ring concept uses two signal paths between stations – one clockwise and the other counter-clockwise. When a failure in a link occurs, the stations on either side reconfigure internally (see Figure 6). The functional stations adjacent to the break make use of their links in the reverse direction to close the ring, thus eliminating the bad link. Should a station itself fail, the stations on either side reconfigure to isolate the failing station and both of the links to it. This results in joining the surviving links into a single logical ring, designated as *wrap* mode.

As shown in Figures 2 and 3, DAS stations may offer a bypass capability. An optical switch is used to bypass a station's peer (AB) ports so that the signal from the previous station is passed directly to the next station. Bypassing may be activated by a station itself, by a human operator, automatically at the removal of power, or by some overall network-controlling function. This maintains dual logical rings between the remaining stations, designated as *thru* mode. (*Thru* and *wrap* are properties of a station.)

Concentrators provide an additional level of isolation between the trunk ring and the attached slave stations. A concentrator can monitor all of its slave stations and deinsert (remove) any faulty station detected. This isolation also provides for graceful insertion and removal of slave stations from the ring, without losing the token or data frames.

The use of these techniques allows FDDI networks to be configured to tolerate a variety of station or link failures and physical network reconfigurations without catastrophic consequences. When failures occur, the network automatically reconfigures, isolating any failing element and maintaining logical ring operation. Continuous monitoring of the failed link or station allows the network to automatically reconfigure and restore normal operation when repair is effected. These reconfiguration actions may result in the loss of individual frames, which then need to be retransmitted.

3. FDDI PROTOCOL DATA UNITS

The basic FDDI encodes all information into frames for transmission on the medium. LLC supplies MAC a Service Data Unit (SDU), which is encapsulated into a frame for transmission. The element of data across the PHY-to-MAC interface is a symbol. At this interface there are two kinds of symbols: data symbols which each represent 4 bits of binary data, and control symbols which may not occur within the body of a frame. Each symbol requires five code bits on the medium for its transfer. The signals on the medium use an NRZI code, where each polarity change (off-on or on-off) denotes a code bit one, and absence of a polarity change denotes a code bit zero.

3.1 Symbols

As shown in Table 1, in the 32-member symbol set, seven symbols of the set are not to be transmitted since they violate code run length and DC balance requirements. Note that the Quiet symbol, however, is a necessary member of the line-state symbol set since it is used to indicate the absence of any functional signal.

DECIMAL	CODE GROUP	SYMBOL	NAME	ASSIGNMENT	
00	00000	Q	QUIET	LINE STATE SYMBOL	
31	11111	I	IDLE	"	
04	00100	H	HALT	"	
24	11000	J		STARTING DELIMITER	
17	10001	K		"	
05	00101	L		"	
13	01101	T		ENDING DELIMITER	
07	00111	R	RESET	CONTROL INDICATOR	
25	11001	S	SET	"	
30	11110	0		DATA SYMBOL	0000
09	01001	1		"	0001
20	10100	2		"	0010
21	10101	3		"	0011
10	01010	4		"	0100
11	01011	5		"	0101
14	01110	6		"	0110
15	01111	7		"	0111
18	10010	8		"	1000
19	10011	9		"	1001
22	10110	A		"	1010
23	10111	B		"	1011
26	11010	C		"	1100
27	11011	D		"	1101
28	11100	E		"	1110
29	11101	F		"	1111
01	00001	V	VIOLATION	NOT TRANSMITTED	
02	00010	V	VIOLATION	"	
03	00011	V	VIOLATION	"	
06	00110	V	VIOLATION	"	
08	01000	V	VIOLATION	"	
12	01100	V	VIOLATION	"	
16	10000	V	VIOLATION	"	

Table 1 - Symbol Coding

The Quiet, Halt and Idle symbols are used for line state signaling, which is provided by the physical layer hardware. The Idle symbol is also used as a timing fill between Data Link Layer PDUs.

The remaining symbols are used by the Data Link Layer. The J, K and L symbols are used for starting delimiters. The JK symbol pair is the physical layer starting delimiter sequence, and is uniquely recognizable independent of previously established symbol boundaries. The L symbol is not used in the basic FDDI, but is used in FDDI-II to embed the MAC starting delimiter within the cycle structure. Sixteen

symbols are used for data quartets, each representing four bits of ordered binary data. The T symbol is used for ending delimiters, and the R and S symbols are used as control indicators.

The encoding of control symbols has been designed to provide maximum code space between them, to ensure at least single bit transmission error detection.

3.2 Line States

Line states are used to indicate the condition of the physical link, and for signaling for link control functions.

Four line states are used in the low-level signaling sequences of SMT. They are generated by continuously transmitting the same line state symbol or symbol pair. Signaling line states include Quiet line state, Halt line state, Master line state and Idle line state. Quiet line state indicates the absence of signaling on the medium. The other three line states generate continuous square wave patterns at different frequencies, which can be reliably detected in the presence of considerable noise.

Four additional conditions are defined, indicating normal or erroneous operation. Active line state is entered whenever the physical (JK) starting delimiter is recognized. Cycle line state is entered when the starting delimiter is followed by a cycle header in FDDI-II. Noise line state is entered upon detection of a persistent invalid link condition. Line state unknown is defined as the condition when the criteria for none of the other line states are satisfied. Noise line state and Line state unknown are never intentionally transmitted.

3.3 Frame and Token Formats

MAC data is transmitted on the FDDI ring in frames which are variable in length. Tokens are special short fixed-length 'frames' that are used to signify the right to transmit data. Figure 7 shows the frame and token formats.

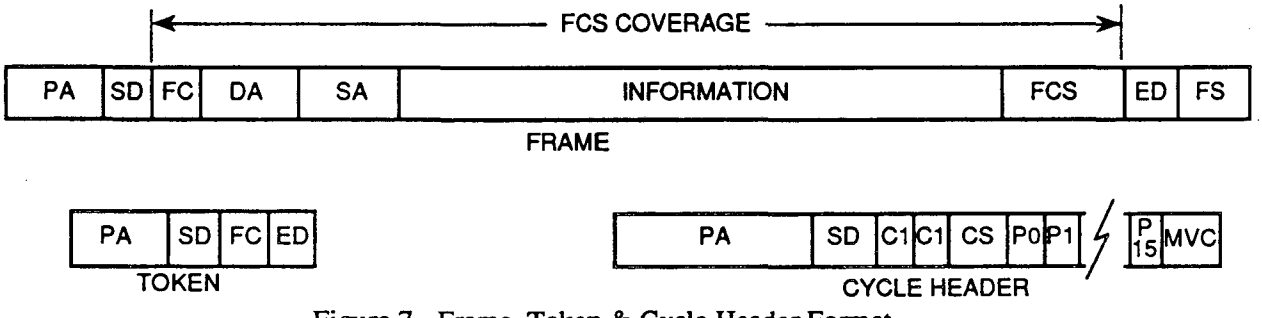


Figure 7 - Frame, Token & Cycle Header Format

The Preamble (PA) field (nominally 16 Idle symbols) is used for establishing and maintaining clock synchronization between frames. The Starting Delimiter (SD) field consists of a JK symbol pair in the basic FDDI. In FDDI-II, frames use an IL symbol pair embedded within the packet data channel as a starting delimiter.

The Frame Control (FC) is an 8 bit field that identifies the type of MAC PDU (frame or token). For frames, the FC defines the service class (synchronous or asynchronous), the length of the address fields (16 or 48 bits), the MAC Service Access Point (SMT, LLC or implementer) and other type-dependent information. For tokens, the FC field defines the token class (restricted or nonrestricted). Two Ending Delimiter (ED) symbols (TT) complete a token.

For frames, the Destination Address (DA) and Source Address (SA) fields may be either 16 or 48 bits long depending on the FC value. DA may be either an individual or a group address, the latter of which has the potential to be recognized by more than one station. The Info field is a variable number of 8 bit octets, subject to the minimum and maximum total PDU size constraints (4500 octets for the basic FDDI and 8600 octets for FDDI-II).

The 32-bit Frame Check Sequence (FCS) field is a cyclic redundancy check using the standard polynomial used in the IEEE P802 LAN protocols. It covers the fields from the FC through the FCS, which may consist only of the 16 data symbols. Data symbols are not used in fields not covered by the FCS.

The Ending Delimiter (ED) field of a frame is one delimiter symbol (T). It is followed by the Frame Status (FS) field that has a minimum of three control indicator symbols that are subject to modification by each station as it repeats the frame. These indicate, when set, that an error has been detected in the frame by the station (E), that the addressed station has recognized its address (A), and that the frame has been copied by the station (C).

3.4 Cycle Format

FDDI-II employs a cycle structure to control the multiplexing of packet (MAC) data and circuit switched (isochronous) data in the same ring. This structure repeats on the ring every 125 microseconds. Figure 8 shows the cycle format. The order of transmission is left to right by row starting with the top row.

The cycle header format is shown in Figure 7. The Preamble (PA) field (nominally 5 Idle symbols) is used for establishing and maintaining clock synchronization between cycles. The Starting Delimiter (SD) field consists of a JK symbol pair. The Cycle Control (CC) field contains two control indicators that control cycle synchronization and sequencing. This field also serves to distinguish FDDI-II cycles from basic FDDI frames. The Cycle Sequence (CS) field provides a modulo 192 cycle sequence count.

The 16 symbols of programming information (P0 - P15) determine whether the corresponding wideband channel (WBC0 - WBC15) is allocated to packet or isochronous traffic. Thus, each Pn controls the multiplexing of one of the columns (WBCn) shown in Figure 8. The Isochronous Maintenance Channel (IMC) byte provides a 64 Kbps voice channel for maintenance purposes.

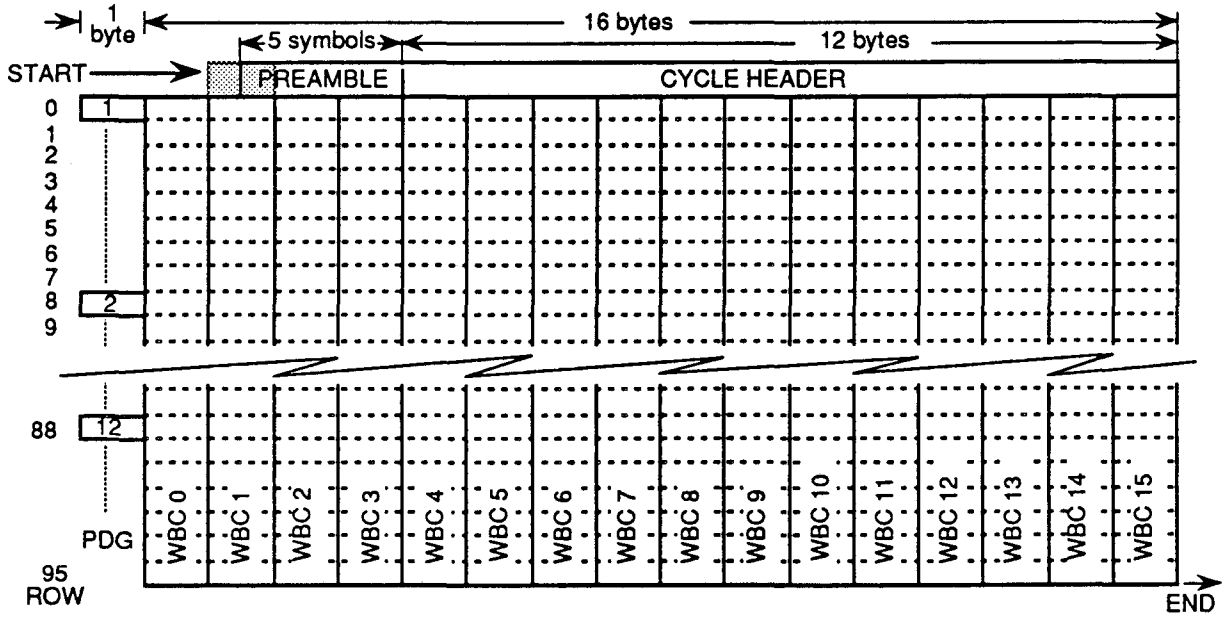


Figure 8 - FDDI-II Cycle Format

The remainder of the cycle structure consists of the interleaved wideband channels and packet data group (PDG). Each column in Figure 8 represents one of these data streams.

4. PHYSICAL LAYER (PHY & PMD) OPERATION

PHY provides the protocols and PMD the optical data links between FDDI stations. Each port (PHY/PMD pair) simultaneously receives and transmits. The transmitter accepts symbols from the Data Link Layer, converts these to five-bit NRZI code groups and transmits the encoded serial data stream on the medium. The inverse process occurs in the receiver.

4.1 Link Clocking

PHY provides the bit clocks for each station. The transmit bit clock uses a local fixed-frequency reference. The receive bit clock is recovered from the NRZI data stream, and is locked to the frequency of the upstream transmit bit clock.

An elasticity buffer (EB) is always inserted in the repeat path between the receiver and the transmitter to compensate for the difference in frequency between the local clock and that of the upstream station by adjusting the bit delay through the station. The elasticity buffer in each station is reinitialized to its center position during the preamble (PA) between frames or cycles. This has the effect of increasing or decreasing the length of the PA as it proceeds around the ring. The transmit clock has been specified with 0.005 percent stability. With a minimum elasticity buffer of 10 bits, frames of up to 4500 octets in length can be transmitted without exceeding the limits of the elasticity buffer.

4.2 Smoothing Function

The operation of multiple cascaded elasticity buffers tends to redistribute the PA lengths in long rings. Longer PAs created no problem by themselves but came at the expense of shorter, or even negative, length PAs, resulting in the loss of the entire frame.

This problem was solved by means of a smoothing buffer function incorporated into PHY. The smoother examines the PA length between frames and either inserts or deletes, as the case may be, preamble symbols (or bytes) in order to maintain the PA near the nominal 16 symbol length. Simulation work presented to the technical committee showed that the algorithm chosen, even under worst case conditions, reduced the probability of frame loss for the maximum length frames to less that 10^{-12}. Extended testing of a large physical ring configuration has reaffirmed these results.

In FDDI-II, preamble length variations translate into jitter in the 8 KHz cycle synchronization function. The smoothing function for FDDI-II limits variation in PA length within a narrow band around the nominal 5 symbol length. The smoothing algorithm is designed so that each smoother stage provides more jitter attenuation than the amount of jitter gain introduced by the action of the corresponding EB stage. Thus the effect of cascaded stations around the ring is that of cascaded jitter filters, so that the resultant cycle synchronization jitter decreases as the length of the ring increases. This has been confirmed by simulation.

4.3 Signal Quality

FDDI stations are required to receive an optical signal stream with an average power of -31 dBm and repeat it with a bit error rate (BER) not to exceed 2.5×10^{-10}. An increase of 2 dB in the power of the input signal requires that the BER decrease to 10^{-12}. The data integrity goals for a full FDDI ring have been represented as a BER of 10^{-9} with a packet loss rate of 10^{-9}. It is probable that many implementations will exceed the requirements by several orders of magnitude.

Input signals with an average power as high as -14 dBm are permitted. Average output power levels are specified at -14 to -20 dBm. The difference between -20 and -31 dBm yields a maximum permitted cable plant loss of 11 dB.

Each physical link of an FDDI ring is monitored by a Link Error Monitor (LEM). Whenever the estimated link error rate exceeds a threshold specified by SMT then that link is excluded. The default value of the threshold is 10^{-7} with 10^{-4} being the maximum permitted value.

The line state signaling mechanism functions reliably when the physical link BER is 10^{-2} or better. The MAC layer recovery mechanisms, which rely on the transfer of short frames, will function reliably when the network error rates are 10^{-3} to 10^{-4} or better. These conclusions were the basis for the choice of 10^{-4} as the maximum LEM threshold.

4.4 Cable Plant

PMD specifies the station characteristics and the cable plant on the basis of 62.5/125 micron multi-mode fiber and 8/125 micron single-mode fiber. This is important in that it provides a single point of reference for conformance verification. Other multi-mode fiber sizes may be used, and PMD provides data in an appendix on the loss factors that must be taken into consideration for the other common fiber sizes. This data includes 100/140, 82.5/125, and several variations of 50/125 micron fiber.

PMD specifies a duplex optical connector with a fixed shroud design for both multi-mode and single-mode fiber.

4.5 Single Mode Fiber PMD

Some FDDI networks are required to support station-to-station distances longer than the 2 km limit of the multi-mode fiber PMD. The SMF-PMD, a single-mode fiber version of the multi-mode fiber PMD, provides an alternative to multi-mode fiber use, increasing the permissable fiber length to the 40-60 km range. Two categories of Active Output and Input Interfaces are specified, and all four combinations of the two categories are supported, thus accomodating cable plant losses from 0 to 32 dB.

The Category I and II Active Output Interfaces provide average power in the range of -20.0 to -14.0 dBm, and -4.0 to 0.0 dBm respectively. The Category I and II Active Input Interfaces operate with average power inputs of -31.0 to -14.0 dBm, and -37.0 to -15.0 dBm respectively. Nominal mode field diameter of the single-mode fiber is 9-10 μm±10%. Zero Dispersion Wavelength of the fiber is 1295-1322 nm.

4.6 FDDI to SONET Physical Layer Mapping

Some FDDI network applications have a need to use SONET common carrier facilities as an alternate FDDI physical medium for certain links. This will allow FDDI network users to easily and reliably extend links of an FDDI network used for private network applications across private or common carrier SONET facilities.

SONET (Synchronous Optical NETwork) is an ANSI standard (T1.105-1988) for a fiber optic transmission service at various rates in a Synchronous Digital Hierarchy. This hierarchy specifies a Synchronous Transport Signal (STS) in multiples of N = 3, 9, 12, 18, 24, 36, & 48, of 51.840 Mbps. The level 3 (STS-3) service is 155.520 Mbps, with a payload of 149.760 Mbps, which is well suited to carry the FDDI signaling rate of 125 Mbps. Thus the concatenated Super-Rate Mapping STS-3c, where multiple STS-1 services are concatenated together, is used for FDDI.

A SONET mapping (SPM) for FDDI signal streams at 125 Mbps has being proposed and is under joint review for approval by the X3T9 and T1X1 communities. Unused bandwidth in the STS-3c service will be available for other services, such as multiple T1 services at 1.544 Mbps.

Initially SPM will be provided in external 'repeater' devices, such that an existing FDDI PMD can interface to SONET without basic change to the specific FDDI PMD implementation. Later, SPM-compliant ports will be provided for FDDI concentrators. This will be more cost-effective, as well as providing more flexible reconfiguration functions for these relatively critical and costly links. (This is also the most likely SMF-PMD evolutionary model.)

5. MEDIUM ACCESS CONTROL OPERATION

5.1 Data Transmission

A token is a special sequence that indicates that the medium is available for use. A properly functioning FDDI ring contains at most one token. Only after having captured a token is MAC allowed to transmit a frame or frames. When finished, MAC issues a new token to signify that the medium is available for use by another station. Priority requirements, necessary to assure the proper handling of frames, are implemented in the rules of token capture. Under these rules, if a given station is not allowed to capture the token, it must repeat it to the next station in the ring.

The basic concept of a ring is that each station examines and (usually) repeats the frames that it receives from its upstream neighbor to its downstream neighbor. If the destination address (DA) of the frame matches the station's MAC address and there is no error indicated, then the frame is copied into a local buffer with MAC notifying LLC (or SMT) of the frame's arrival. MAC modifies the indicator symbols in the FS field of the frame as it repeats it to indicate the detection of an error in the frame, the recognition of its own address, and the copying of the frame.

Each frame propagates around the ring to the station that originally placed it on the ring. The transmitting station may examine the indicator symbols in the FS field to determine the success of the transmission.

The MAC of a transmitting station is responsible for removing from the ring all of the frames that it has placed on the ring (a process termed stripping). Each MAC recognizes its returning frames by matching their SA fields with its own address. When the SA field matches, repeating ceases, and Idle symbols are placed on the medium to strip the remainder of the frame. Remnants of stripped frames occur because a station's decision to strip a frame is based upon recognition of its own address in the SA field, which cannot occur until after the initial part of the frame has already been repeated. These remnants cause no ill effects because various criteria, including recognition of an ED, must be met before a frame is accepted as valid.

Stripping is an important aspect of FDDI ring operation. When a station, having captured a token, initiates transmission, the ring could contain a number of frames in transit that were placed on the ring earlier by other stations. These frames should be removed from the ring by their originating stations, before they arrive (for what would be the second time) at the currently transmitting station's receiver. Thus, the transmitting station should receive Idle symbols and the remnants of stripped frames, until its own frames return. While the station is transmitting, it will be removing from the ring anything that arrives at its receiver. After it issues a token, the station generates Idle symbols until the SD of a frame is received. At this time it resumes the examination of all arriving frames, stripping those that it has originated and repeating all others.

5.2 Data Integrity

MAC has the responsibility for the recognition of frames that are addressed to it. MAC must ensure that frames accepted are indeed intended for it. Furthermore, MAC must ensure they are valid frames and not the result of some data corruption process. The required reliability of frame reception is enforced by the valid frame criteria.

A valid frame is required to be correctly formed, including proper starting and ending delimiters, with its body composed entirely of the sixteen data symbols. The FC field must be one of the ones allowed for frames. In addition, all fields within the frame must conform to the proper length restrictions, e.g., the information field must contain an even number of symbols. The FCS field must be correct and the E indicator is required to be reset, the latter indicating that no other station has detected any errors as the frame progressed around the ring.

With these criteria in place, the probability of MAC improperly accepting a frame as valid, i.e., an undetected error, has been shown to be of the order of 4×10^{-25} under pessimistic assumptions, *e.g.* with 1000 physical links (a maximum default size ring in the wrap condition), all with a worst case BER of 2.5×10^{-10} and with all frames having the maximum length of 4500 bytes.

5.3 Ring Scheduling

The FDDI MAC uses a Timed Token Rotation (TTR) protocol to control access to the medium. Under this protocol, the MAC in each station measures the time that has elapsed since the token was last received. The initialization procedures establish a Target Token Rotation Time (TTRT) equal to the lowest value that is bid by any of the stations. This TTRT value is retained by each MAC and becomes the value of the Token Rotation Timer (TRT).

The use of the TTR protocol imparts some useful operational characteristics. It allows stations to request and establish guaranteed bandwidth and response time for synchronous frames via SMT procedures. It

establishes a maximum response time for the ring because, in the worst case, the time between the arrival of two successive tokens will never exceed twice the value of TRT. It also establishes a maximum ring utilization equal to (TRT-RL)/TRT, where RL is the physical ring latency, *i.e.* the time for a token to propagate around an idle ring.

Low values of TTRT (*e.g.* four milliseconds) may be used to establish an average token rotation time of four milliseconds and a guaranteed response time not exceeding eight milliseconds. This would be useful in a time-critical application (*e.g.* packetized voice). Larger values of TTRT may be used to permit very high ring utilizations under heavy loads. For instance, using a TTRT value of 50 milliseconds and a ring latency (RL) of 0.25 milliseconds (reasonable for a ring consisting of 75 stations and 30 kilometers of fiber) a utilization of 99.5 percent can achieved.

Two classes of service are defined. Synchronous service is used where predictable units of data are to be delivered at regular intervals. Synchronous service uses any received token. Ring access is guaranteed with a delay not exceeding twice TTRT. Asynchronous service uses only early tokens, *i.e.* when the time since the last received token is less than the negotiated TRT.

The bandwidth required for synchronous service is allocated (via SMT protocol) from the total available FDDI bandwidth. However, it is dynamically assigned as synchronous traffic is enqueued. Unused bandwidth is dynamically assigned to asynchronous packet traffic.

Asynchronous service is available in two modes. Highest priority is given to restricted token mode. Cooperating stations may enter a restricted dialogue, which allows them to issue and use restricted tokens, after negotiating an agreement using nonrestricted tokens. Restricted token mode operation allows stations to contend for available asynchronous bandwidth on a dialogue basis.

Lowest priority is given to asynchronous traffic that may be transmitted only by capturing a nonrestricted token. This mode of operation allows stations to contend for the available asynchronous bandwidth on a single frame basis. Multiple levels of priority for nonrestricted mode asynchronous frames may be provided within a station by specifying additional (more restrictive) time thresholds for token rotation.

5.4 Ring Initialization and Recovery

MAC is responsible for ring initialization to ensure that one, and only one, token is generated. This is accomplished through the claim process which is started whenever MAC is first activated. This process is also used to negotiate the same value for TRT in all of the stations on the ring.

During this process, MACs in different stations may generate a series of claim frames, each containing that MAC's bid for the TTRT value. The lowest TTRT bid is given precedence. In the event of a tie on the TTRT values, the highest MAC address is given precedence. This process continues until eventually only one MAC (the one with the lowest TTRT bid that has the highest address) is generating claim frames, which are being repeated by all of the other MACs. When a MAC receives its own claim frame back it knows that it has won the claim process and it then issues one token.

The beacon process is invoked when the claim process fails to complete and under other circumstances, such as when the operation of the ring is suspect. The beacon process is entered when the MACs in one or more stations transmit a beacon frame. During the beacon process the MAC in each station yields to an incoming beacon frame and repeats it.

Two conclusions to this process are possible. In one case, the ring is apparently healed. Either all beacon frames have been discarded and a timer expires, or at least one MAC recognizes the SA in a received beacon frame as its own. The MAC(s) then enter(s) the claim process and, when that process completes, one token will be generated and the ring again made operational.

In the other case, the fault persists. Eventually, a quasi-stable state is reached with the MAC immediately downstream of the fault domain generating beacon frames which are being repeated by all stations up to the fault domain. Since the beacon frames contain the SA of the beaconing MAC, the location of the fault domain is apparent to a manager located at any of the stations. Alternatively, the SMT of the beaconing station may take the necessary corrective action to heal the ring if the beaconing persists for too long.

Many of the potential faults that would otherwise cause the beacon state to be entered are detected by the LEM or other physical layer monitoring functions. In these cases SMT will reconfigure around the faulty link so that the stations at both ends of the fault domain will be reconfigured to the wrap state, thereby healing the ring.

6. STATION MANAGEMENT OPERATION

SMT provides the FDDI-specific subnetwork management functions, including the control required for proper operation of an FDDI station in an FDDI ring. SMT monitors ring activity and exercises overall control over station activity. Each FDDI station has one SMT entity.

SMT operation is divided into three broad categories: Connection Management (CMT), Ring Management (RMT), and Operational Management.

6.1 Connection Management

CMT is primarily concerned with establishing and maintaining the physical and logical topology of the FDDI network. The physical topology consists of full-duplex physical connections between ports in adjacent nodes. From this physical topology CMT builds a logical topology consisting of counter-rotating logical rings. To this end, primitive signaling sequences use the four signaling line states to establish each physical connection and to exchange control information between the ports.

CMT manages the physical layer resources of an FDDI node. Other functions performed by CMT include the Link Error Monitoring function, and the Trace function, which is used to force nodes in a suspected fault domain to leave the ring and enter self-test in an effort to localize a fault.

6.2 Ring Management

RMT is primarily concerned with establishing and maintaining correct logical ring operation, *i.e.* with a usable token in circulation.

RMT manages the MAC layer resources of a station. Other functions performed by RMT include the detection of MAC error conditions, such as the MAC of a station stuck in the beaconing state or in beacon/claim oscillation (*e.g.* due to duplicate MAC addresses). When appropriate, RMT initiates the CMT trace function to attempt to localize the fault. RMT may provide notification to the nearest upstream neighbor of detected error conditions by transmitting a special 'directed beacon' frame that is addressed to it.

6.3 Operational Management

The remaining functions of SMT deal with the management of the FDDI network in the operational state. These services extend across multiple stations and include a well-defined set of SMT frames to achieve an interoperable implementation of the layer management protocol.

Although it is the intention that SMT provide all of the facilities required for management of FDDI in an OSI environment, it is also necessary that SMT provide frame based management tools for other management protocols. This includes the capability to obtain and modify management attributes of any of the local and remote FDDI entities, (MAC, PHY, PMD, *etc.*). Manageable attributes includes such items as policy and parameter settings, current status, counter values, timer values, error information, addresses, and connection information. These attributes are supplied to the management agent process(es). There is also a protocol for reporting significant event information, which includes such things as changes in availability of the managed entity and certain error events.

7. FDDI-II

7.1 FDDI-II Concepts

FDDI-II is an upward-compatible enhancement of the basic FDDI that adds a circuit switched (or isochronous) service to the existing packet service. The circuit switched data service can be described as a constant stream of data. The data rate is matched to the application, with, for example, 64 Kbps being used for a digital voice data stream. Other data stream rates, even up to many Mbps in the case of video, are used for other applications.

The contrasting nature of packet and circuit data services is of interest. Some packet data traffic occurs in random quantities and at random times. This is referred to as asynchronous traffic. Other packet traffic, more regular in nature and occurring in relatively predictable quantities on a regular time basis, is referred to as synchronous (packet) traffic. In contrast, isochronous data occurs in precise amounts on a precise time basis. It typically represents a sequence of digital samples from a sensor (*e.g.* voice or video). More importantly, isochronous data must be synchronized with clock information to ensure the accurate regeneration of the sampling clock (as distinct from the bit clock) to minimize distortion in data reconstruction. Isochronous data is more easily transferred in a circuit switched network.

The FDDI-II isochronous service provides continuous, synchronized channels at fixed data rates between two or more end points. These channels are established via a management protocol. Instead of using addresses attached to the data, a channel is identified by the location of its time slot, or slots, that occur regularly relative to a synchronization signal. The public networks use a standard 125 microsecond (8 KHz) synchronization signal. (In North America this is known as the Basic System Reference Frequency (BSRF)). FDDI-II can use an external (*e.g.* public network) synchronization signal when it is available; otherwise, an internal reference is used.

Networks that carry isochronous data must maintain precise synchronism. An FDDI-II ring is organized as a synchronized periodic structure known as a cycle, repeating every 125 microseconds. One station (called the cycle master) must insert a delay for all isochronous data so that the isochronous channel delay appears to be an exact multiple of the 125 microsecond cycle period.

FDDI-II allocates bandwidth to isochronous service in units called Wideband Channels (WBC). Each WBC is 6.144 Mbps (96 bytes per cycle), which is four times the North American, and three times the European, basic access rate to the public digital network. WBCs are full duplex and are independently allocatable. Each WBC is one of the columns in Figure 8.

At 100 Mbps there are 16 WBCs, providing a maximum of 98.304 Mbps isochronous payload. The remaining bandwidth is used for the overhead of the cycle structure, plus a minimum of 768 Kbps dedicated to the packet service. This residual bandwidth (12 bytes per cycle), designated the Packet Data Group (PDG), is interleaved with the 16 WBCs as shown in Figure 8. H-MUX augments the PDG with the bandwidth of any WBCs that are not assigned to isochronous service. This is an efficient system that allows the bandwidth of all unallocated WBCs to be used by the packet channel. For example, with eight of the WBCs assigned to isochronous service, 49.920 Mbps of bandwidth would be available for packet service. The total packet plus isochronous payload is 99.072 Mbps, or more than 99% efficiency.

The allocation of transmission channels within the isochronous WBCs is allowed with a variable granularity. Supported channel sizes include 8 , 16, 32 and 64 Kbps, plus any multiple of 64 Kbps up to 6.144 Mbps. This includes the commonly used ISDN data rates. Mixtures of these data rates in the same WBC are allowed. In addition, the aggregate of any or all of the isochronous WBCs may be used as one (super-rate) circuit service, satisfying the needs of such applications as high resolution video. Thus a multiplicity of channels may be provided within the same FDDI-II ring.

7.2 FDDI-II Operation

FDDI-II introduces an additional standard called Hybrid Ring Control (HRC). HRC defines the additional Data Link Layer entities to support isochronous data. These entities are a Hybrid Multiplexer (H-MUX) and an Isochronous MAC (I-MAC).

H-MUX functions as an additional sublayer between the PHY and MAC sublayers. H-MUX establishes and maintains the cycle structure, including the programmable Wideband Channels. The WBCs provide a dynamic bandwidth division mechanism, with a granularity of 6.144 Mbps, for multiplexing data between the (packet) MAC and the Isochronous MAC. This requires that the (packet) MAC be able to transmit and accept data on a non-continuous basis because packet data is interleaved with isochronous data.

An FDDI-II ring can be initialized in basic (frame) mode and switched to hybrid (cycle) mode after a station has negotiated for and won the right to be cycle master. Alternatively, one or more potential cycle masters (monitor stations) can directly initialize the ring in hybrid mode. If more than one monitor attempts to become cycle master a monitor contention process elects the highest ranking monitor to be cycle master (usually within one ring circulation delay).

The cycle master generates cycles at an 8 Khz rate (every 125 microseconds) and inserts the latency required to maintain an integral number of synchronous cycles in the ring. This delay is only incorporated for isochronous WBCs, so there is no unnecessary delay in the packet traffic. This is essential in providing an integrated services network with acceptable packet service.

I-MAC provides access to the individual channels within isochronous WBCs. I-MAC accesses a channel as L bits beginning at byte M after the cycle synchronization signal in WBC number N. The last attribute is necessary because FDDI-II has 16 WBCs that may be independently assigned to either packet or isochronous data. This definition allows channel data rates from 8 Kbps (N=1) up to the 6.144 Mbps data rate of a WBC. Multiple WBCs may be joined to accommodate higher circuit data rates (*e.g.* for video).

Data steering logic in I-MAC steers the received data streams to the I-MAC data service interface, and the transmitted data streams from the service interface, for channels that are open at the station.

FDDI-II defines but does not standardize a Circuit Switching Multiplexer (CS-MUX) function above the I-MAC data service interface. CS-MUX adapts the FDDI-II transmission channel service to provide any necessary virtual circuit signaling and synchronization functions. At the endpoints of the virtual circuit, CS-MUX provides any required adaptation to the virtual circuit user. Between subnetworks (*e.g.* FDDI-to-FDDI or FDDI-to-ISDN) CS-MUX provides any required inter-subnetwork adaptation functions, such as frame alignment and time slot interchange.

7.3 FDDI-II Applications

FDDI-II expands the range of applications that can be supported by FDDI rings. An FDDI-II ring can connect high performance workstations, processors and mass storage systems, plus bridges, routers and gateways to other LANs, MANs and WANs. The same ring could have some of its bandwidth allocated to isochronous services. This isochronous bandwidth may in turn be suballocated into a variety of virtual circuit services such as video, voice and possibly control or sensor data streams. The division of bandwidth between these services may be adjusted based on the time of day or other requirements.

In practice, no single FDDI-II ring is likely to connect to all these types of equipment, and certainly not to all the equipment of a large site. Instead, multiple interconnected FDDI-II rings would be used, with high traffic units, such as processors, attaching to several rings.

The evolution of FDDI-II applications will parallel that of the basic FDDI. Initially FDDI was used as a backbone for existing lower-speed local networks such as Ethernet. However, as its cost decreases and it is designed into newer equipment (rather than being added on), the dominant use will be for high speed data distribution. Similarly, FDDI-II could initially be used as a backbone for interconnecting PABXs and other ISDN-compatible equipment. However, its dominant use will be for real time video and other image distribution. There will be a substantial demand for both interactive and non-interactive video and image services. Interactive applications include video conferencing and real time visualization. Non-interactive applications include real time surveillance (*e.g.* video, radar and sonar) and medical and industrial testing and diagnostics (*e.g.* optical, X-ray and magnetic resonance imaging).

It is anticipated that FDDI-II implementations will follow those of the basic FDDI by one to two years. ASC X3T9.5 has specified the rules of coexistence of these two FDDI implementations to ensure

interconnectability and interoperability of FDDI-II implementations operating in basic mode with the basic FDDI implementations. This also allows the use of FDDI-II chips for all applications once they become available.

This migration strategy means that many installations will contain a mixture of basic FDDI and FDDI-II equipment. In these installations basic mode and hybrid mode rings will coexist. Concentrators and bridges will be used to interconnect these subnetworks. Whenever any station is plugged into a concentrator, the concentrator will automatically determine the capability of the station and insert it into an appropriate ring. Bridges will relay both packet and circuit data between rings as necessary.

8. BEYOND FDDI

FDDI has become the high-speed LAN of choice. Its success is the result of (sometimes painful) cooperation between the various manufacturer and user communities. The manufacturers include system, semiconductor, fiber-optic component, connector and cable producers. This is resulting in very wide utilization of FDDI. If one can view Ethernets and IBM Token Rings as the RS-232 service of the 1980s, it is equally likely that FDDI will become the equivalent service of the 1990s.

This wide acceptance of FDDI, and the lengthy process of converting to it, argues persuasively for the future development of LANs beyond FDDI-II in performance and function. As FDDI use evolves from a backbone for lower speed local networks into distribution of higher speed services, demand is emerging for an even higher speed data and video backbone LAN.

The development and implementation cycles for networks such as Ethernet and FDDI have been very long. For example, Ethernet was originally developed in 1973, saw limited use in the late 1970s, was redesigned in the early 1980s, started to come into wide use in the mid-1980s and by 1990 was widely used. FDDI was originally developed in 1982, saw limited use in the late 1980s, and will be in wide use by the mid-1990s. Clearly it is desirable to build as much as possible on existing architectures to minimize the development time, and thus ensure the wide-spread acceptance of future LAN products.

At this time, the likely requirements for the next generation LAN can be identified:

- Ability to provide a backbone for multiple FDDI networks.
- Ability to provide efficient interconnections to wide area networks, *e.g.* B-ISDN.
- Support for a wide variety of 'integrated' services, *e.g.* video, voice, data, graphics.
- Data rate at least 600 Mbps, but less than 1.25 Gbps.
- Data rate matched to the Synchronous Digital Hierarchy.
- Duplex links for reliability.
- Physical ring and tree topologies supported so existing FDDI cable plant may be used.
- Ability to operate over leased public network (*e.g.* SONET) links on specific segments.
- Support for both single- and multi-mode fiber.
- Media accessing modes for ATM, Isochronous and FDDI-like packet services.

Additional issues that must be considered are:

- The extent of the network – network size tradeoff with access and recovery protocols.
- Physical encoding – FDDI 4/5 or more effective encodings.
- Media access method – FDDI token ring or other access methods.
- Logical topology – FDDI dual counter-rotating rings or other logical topologies.

To the extent that next generation LAN requirements can fit the FDDI architecture, the development times for standards and products can be greatly reduced. The best example of this is the tradeoff between an FDDI-like LAN supporting ~600 Mbps over existing FDDI cable facilities, versus a 1+ Gbps LAN in-

compatible with FDDI that requires new cable plants and significantly greater time-to-market. Clearly these are key issues for the future.

At this time it appears highly likely that an FDDI-like LAN operating at STS-12c data rates (~600 Mbps), utilizing the existing FDDI cable plant and evolutionary development of the existing FDDI architecture, will be feasible at a very competitive cost. If this is so, the evolution of FDDI will continue well into the 21st century.

REFERENCES

1. American National Standard, FDDI Token Ring Media Access Control (MAC), ANSI X3.139-1987.

2. American National Standard, FDDI Token Ring Physical Layer Protocol (PHY), ANSI X3.148-1988.

3. Draft Proposed American National Standard, FDDI Token Ring Physical Layer Medium Dependent (PMD), ASC X3T9.5 Rev. 9, March 1989, (ANSI X3.166-199x).

4. Draft Proposed American National Standard, FDDI Hybrid Ring Control (HRC), ASC X3T9.5 Rev 6, May 1990, (ANSI X3.186-199x).

5. Draft Proposed American National Standard, FDDI Token Ring Station Management (SMT), ASC X3T9.5, Rev. 6.2, May 1990.

6. Draft Proposed American National Standard, FDDI Token Ring Single-Mode Fiber Physical Layer Medium Dependent (SMF-PMD), ASC X3T9.5 Rev. 4.2, May 1990, (ANSI X3.184-199x).

7. Project Proposal, Fiber Distributed Data Interface to Synchronous Optical Network (SONET) Physical Layer Mapping Function (FDDI-SPM), February, 1990.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
INFORMATION RESOURCES DEPARTMENT
BERKELEY, CALIFORNIA 94720