Development of the FDDI Physical Layer

By Jerry D. Hutchison, Christopher Baldwin, and Bruce W. Thompson

Abstract

The engineering development of the FDDI physical layer resulted in the delivery of components, specifications, and protocols. The development presented new design problems related to the technology and to the operation of token rings. The choice of the most appropriate technologies for the chip set was based on technology issues, risk control, and costs. The chip set that emerged after the physical layer functions were partitioned uses both ECL and CMOS technology. Further, three design problems of general interest arose during development: the elasticity buffer and

circuitry related to the distributed clocks in an FDDI LAN, the multimode fiber-optic link using light emitting diodes, and the media error processes as related to correctness and fault isolation.

Introduction

second. The physical layer of FDDI-the topic of this paper - connects many stations, each of which may transmit information to any other station in the network. As in other LANs, packets of user data are encoded according to the physical layer protocol and are transmitted as a serial data stream over a physical media to other stations of the LAN. FDDI, however, is unique in its use of hundreds of individual, point-to-point, fiber-optic connections that form a ring network in the physical layer. The resulting LAN offers both a high data rate and a total physical extent of up to 100 kilometers (km).

The development of physical layer hardware, used in all FDDI products, included the physical protocol (encoding/decoding) device, a receive clock recovery device, a local clock generator, and optical transmitters and receivers. This paper

The fiber distributed data interface (FDDI) is a multiaccess, packet-switching local area network (LAN) that operates at 100 megabits (Mb) per

focuses on development of the physical layer hardware and describes some aspects of the design in detail. We first review the operation of the physical layer and

the functional partitioning of the implementation. We then present detailed discussions of the distributed clock scheme, the design of an optical link, and the methods to control the effects of bit errors in the physical layer. Some of the results of the development to improve the performance, correctness, and reliability of FDDI described here have been incorporated in the American National Standards Institute (ANSI) FDDI standards.

Operation of the Physical Layer

The FDDI physical layer is a collection of point-to-point links forming a "ring." The operation of the layer is described in terms of physical links, physical connections, and the functions of individual stations. The many station types allowed by the ANSI FDDI standards are constructed with a simple physical layer functional block called the PHY port.

A physical link contains a transmitter, a receiver, and a segment of physical medium which conducts the bits of a packet from one station to a second station. The topology of FDDI is arranged so that topologies for FDDI. Each bit of information received from one physical link is transmitted onto another physical link until the information travels around the loop and returns to where it started. The FDDI protocols provide for a single originator of data packets; other stations repeat the data so that each station on the ring receives the packet of information. The collection of many point-to-point links forms the ring, which is viewed as a multiaccess medium by the users.

The basic element in the topology of an FDDI LAN is the physical connection. A physical connection contains two physical links, also shown in Figure 1. Within the station, the circuitry that implements physical layer functionality needed for one physical connection is called the PHY port. The physical connection is a full duplex connection between exactly two PHY ports. Neighbors in the ring directly exchange the control information for each connection, allowing control protocols in FDDI station management (SMT) to establish the shared states for a connection: in-use, starting, and disconnected. The status "in-use" indicates that a

the collection of physical links forms a closed path, or ring, as shown in Figure 1. This simple topology illustrates the basic concepts common to even the most complicated

connection is part of the ring; other states indicate it is not. The control information exchanged over the physical connection is used to autoinitialize and autoconfigure the

connections in the LAN, a method of operation currently unique to FDDI rings.

There are several types of FDDI stations, and different types can support different numbers of physical connections.[1] A single attachment station (SAS) (as seen in Figure 1) can establish one physical connection with a single neighbor. The dual attachment station (not shown) has two PHY ports and may establish physical connections with two neighbors. A concentrator (CON) is a type of station that can establish connections with many neighbors, thereby providing attachment points for other stations. The CON shown in Figure 1 interconnects its PHY ports internally to configure a single ring.

Figure 2 shows the functions of and flow of data through a PHY port which implements the FDDI physical layer protocol (PHY) standard.[2] Data packets to be transmitted over the LAN are passed as a stream of bytes from the data link to the physical layer. Each byte contains two PHY symbols, and each symbol represents 4 bits of user data. The FDDI

from the coded signal) and bounds the low-frequency components of the signal spectrum. The code bits are then converted to a serial stream and transmitted as optical pulses on the fiber-optic media.

The station is coupled to the media with a media interface connector (MIC). The MIC provides the concrete interface necessary for interoperability between equipment from multiple vendors. The FDDI Physical Layer Media Dependent (PMD) standard specifies mechanical and optical properties of the MIC.[3] The MIC includes both a transmit and a receive interface.

Signals received from a connection are decoded by the PHY port for processing by the station. The optical input signal is translated to an electrical signal. The remote bit clock is extracted from the signal and used to recover logic levels corresponding to the individual bits. A framer then establishes the original code group boundaries and converts the serial code bit stream into parallel form. Also, the elasticity buffer synchronizes the received data to the local clock

coding scheme, called 4B/5B encoding, translates each symbol into a code group containing 5 code bits. This encoding limits the maximum time between transitions on the media (allowing clock information to be derived reference and accounts for the frequency difference between the local and remote clock references. Finally, code groups of 5 code bits are decoded into symbols, and symbols are correctly paired to form the data bytes which

represent the received data. These data bytes are passed either to another PHY port or to the data link layer. A later section, Operation of the Distributed Clock Scheme, expands on the elasticity buffer design.

We have so far described the FDDI physical layer in terms of PHY ports and the physical connections between them. These basic elements form the physical layer for all types of FDDI stations. Different types of FDDI stations have one or many PHY ports, but the operation of an individual PHY port and physical connection is independent of station type and topology. In the next section, we discuss the functional partitioning of the PHY port and the reasons behind the partitioning chosen. Subsequent sections describe the distributed clock scheme, the design of the physical link, and the impact of physical link errors on the LAN.

Functional Partitioning

In this section, we describe the partitioning of the functions of the PHY port into the following components:

- PHY (physical protocol chip, also referred to
- FOT (fiber-optic transmitter)
- FOR (fiber-optic receiver)

Our choices for the appropriate partitioning and technology were founded on our decision to develop a highly integrated and low-cost chip set. After examining several alternatives, we chose a partitioning that enabled us to use mostly CMOS technology (complementary metal oxide semiconductor), a minimal amount of custom ECL (emitter coupled logic), and no ECL gate array technology. Although a 125-megahertz (MHz) serial channel requires ECL circuitry in the system, we wanted to minimize the amount of custom ECL technology. ECL consumes a substantial amount of current and is relatively expensive as compared to CMOS technology. We also considered ECL gate array technology, but decided against it because it was not a mature technology, lacked requisite analog functions for clock extraction, and was available from only a few vendors.

We determined that the CDCR and CDCT were the only functions that had to be implemented in ECL technology. This

as the ELM chip)

- CDCT (clock and data conversion transmitter)
- CDCR (clock and data conversion receiver)

determination was based on the need for the high transmission rates and for quick conversion to and from serial and parallel data streams. The CDCR receives a 125-megabaud

ECL serial data stream from the FOR. Using a phase lock loop, CDCR extracts a receive clock to recover the data bits and then converts the serial data to a 5-bit parallel bus. The CDCT receives a 25-MHz, 5-bit-wide parallel bus; then, by using a phase lock loop, it generates an internal 125-MHz transmit clock in phase with the local 25-MHz clock. CDCT then converts the 5-bit-wide parallel bus to a 125-megabaud ECL serial bit stream that is transmitted by the FOT.

We selected the 5-bit width for the parallel bus to obtain a 25-MHz bus rate. This rate is a convenient divisor of the 125-MHz serial rate and is within the operating range of the CMOS gate array technology used in the connected chips. The 5-bit bus also offered the advantage of enabling us to maintain a low pin count on the devices to which the bus is interfaced, thus further containing costs.

Another complication relative to the clock component was how to distribute a 125-MHz clock signal. As noted earlier, some FDDI products have many PHY ports, and those PHY ports must have a clock, and convert the 5-bit parallel bus to the serial stream. With this method, the highest clock rate distributed on our boards was a 25-MHz clock.

As a consequence of selecting the transmit phase lock loop, we chose to specify and build separate transmit (CDCT) and receive (CDCR) devices in custom ECL technology. We were very concerned that the combination of two asynchronous phase lock loops on a single chip would induce cross talk. Cross talk could cause false locking of the phase lock loops to one another, resulting in lost data. Therefore making a single chip was considered too risky for the initial implementation. Our solution was to specify the transmit and receive devices, thus eliminating the possibility of cross talk.

The balance of the logic for the physical layer protocols could now be designed in CMOS gate array technology. The use of CMOS gate arrays was important in meeting schedule since it allowed us to quickly implement changes. Changes were inevitable and therefore had to be accommodated

common clock line for the transmission of the serial data stream. We decided to add another phase lock loop in the transmit component that would lock onto the 25-MHz local clock, generate the 125-MHz serial

because the ANSI standard was not finished and stable during our design cycle. All of the physical layer functions such as the encoder, decoder, elasticity buffer, framer, and smoother

were implemented in CMOS gate arrays. We had simulated these functions in software; however, we were now able to build them using CMOS gate arrays and actually analyze their behavior in real networks. With the hardware, we quickly verified the protocols defined in the FDDI standards. Proper PHY operation is best confirmed by testing actual implementations.

The fiber-optic transmitter (FOT) converts a 125-megabaud electrical signal to light pulses to be transmitted to a receiving station. The fiber-optic receiver (FOR) receives the pulses from a transmitting station and converts them to an electrical data stream. We decided not to develop the FOT and the FOR components ourselves. Instead we chose to influence the specification of the system's functional requirements in the ANSI FDDI Committee and then depend on external vendors to develop the components.

It was important to encourage the optical vendors to standardize their components so costs would decrease, and so that more than one source of optical components would be available to us. Accordingly we did

Operation of the Distributed Clock Scheme

In the FDDI distributed clocking scheme, each station uses an independent, local clock reference when transmitting or repeating data packets. The station must synchronize the receive data with its own reference clock prior to further processing. Although this distributed clock reference scheme simplifies many problems, it also can give rise to data integrity problems and packet loss rate issues that must be solved for the scheme to work effectively. Data must be synchronized to the local clock reference in a way that prevents detected and undetected errors caused by metastability problems. Further, interpacket gap shrinkage that can result in an unacceptable packet loss rate must be controlled. In the sections Elasticity Buffer and Smoother below, we describe how these problems are addressed in the physical layer protocol.

Elasticity Buffer

Each PHY port of a station must accept data packets

not combine the optical transmitter and receiver with any other physical layer functions. The optical link design is the subject of a later section in this paper.

from another station with a slightly different clock frequency and bit transmission rate. It is the function of the elasticity buffer within the PHY port to synchronize the incoming

data to the local clock reference. The buffer is also designed to control synchronizer metastability, a source of undetected data corruption. As a result of the elasticity buffer operation, the size of the gap between two data packets varies as the packets are repeated around the ring.

The elasticity buffer is a collection of storage registers that are written to and read from at different rates. (See Figure 3.) The buffer forms a circular queue due to the movement of two independent pointers: the input pointer selects the register to be written to and moves at the recovered clock rate; the output pointer selects the register to read from and moves at the local clock rate. The location of the pointers is based on the gray code counters. The input pointer is controlled by the input state machine and the output pointer by the output state machine. These state machines position the pointers at a controlled distance from one another. Therefore pointer control prevents data from being written while it is being read from the same register, even though the pointers are moving at different rates.

during the idle time between data packets, known as the interpacket gap. When a minimum interpacket gap time is detected by the input state machine, a reset control signal is sent to the output state machine. The reset signal is synchronized by the output state machine to avoid metastability. When an input signal to a register is changing at the time the register is clocked, its output may become indeterminant and assume multiple values over a time called a period of metastability.[4] The reset signal could be changing when it is sampled by the output state machine, so this signal is synchronized by waiting an interval after each sample for the sample value to settle before the sampled value is utilized, or considered valid. The reset signal is delayed by this process. The circuitry guarantees that the present address of the input pointer has been in the holding register on the reset condition for a sufficient duration and thus its stability is ensured. The output state machine then loads the address that the input state machine stored in

During normal operation, the input and output pointers approach each other and must be periodically repositioned. The repositioning occurs the holding register. The output pointer moves to that location plus an offset in order to keep a minimum distance from the input pointer.

This approach to repositioning pointers ensures that all data or signals are stable in their respective registers before being sampled by the local clock. As a consequence, we were able to specify in the design that only one control line be synchronized in the elasticity buffer. The signal that crosses the clock boundaries is the reset signal, which is generated by the remote clock and sampled by the local clock. The reset signal triggers all events required for the elasticity buffer to correctly receive data. Since there is only one control line within the elasticity buffer that needs synchronization, the implementation is very robust.

We also had to anticipate occurrences outside normal operation. Therefore we designed circuitry that detects when the pointers point to the same register for more than a minimum amount of time.[5] This circuitry prevents the buffer from reading the register while its contents are changing; if the buffer were read, data corruption and consequently undetected errors could result during abnormal operation. The input and output pointers buffer has overflowed or underrun.

The input and output pointer counters are compared using the local clock; that is, the input pointer counter bits can change with respect to the local clock, as shown in Figure 4. Gray code counters limit to one the number of bits that can change in the pointer counters at any sample interval. The comparator circuit is sampled twice using the local clock and the local clock shifted by 90 degrees. If one sample of the D flipflop notes a change, the changing bit settles down before the other sample happens; thus metastability problems are controlled. A logical AND of the sampled outputs signifies that the two pointers have had the same address for at least one quarter of the local clock interval. When the addresses are the same for at least one quarter of the local clock time, the error flag is raised. When the error flag is raised, an overflow or underrun is imminent. The cause is somewhere in the network. For instance, the clocks are out of specification, or a downstream station

are 3-bit, asynchronous gray code counters. The pointers are compared to one another to determine whether they coincide, indicating that data in the is sending larger data packets than are permitted. In any case, the condition is detected and prevented from increasing the risk of undetected data corruption.

Smoother

Another function in the physical layer, called the smoother, prevents the loss of data packets that can result from shrinkage of the interpacket gap. Data packets can be lost, or discarded, at two places. The decision to discard can be made at the physical layer as the result of an elasticity buffer overflow or underflow. Also, packets can be discarded at the media access control (MAC) layer. The MAC layer is not required to copy a packet that has less than 6 idle bytes of interpacket gap preceding the packet.

The decision to implement the smoother came after simulations of the elasticity buffer revealed the then current draft ANSI PHY protocol would result in an unacceptable packet loss rate. In a series of nodes with a random distribution of clocks, some stations add to and others delete from the interpacket gap. If nodes add or delete without regard to the size of the interpacket gap, we found that the interpacket gap could be deleted entirely or reduced to a minimum size. At this size, the MAC is not required to copy

- One hundred one elasticity buffers
- Maximum size data packet length
- Pseudorandom clock distribution

The solution is to monitor the interpacket gap.[6] If it falls below the 7-byte minimum size, the smoother adds to the interpacket gap. The addition of interpacket gap to the output stream causes the elasticity buffer to use a buffer to delay the output data and then send an interpacket gap byte. The amount of buffering within the elasticity buffer is finite so that the delay within a station is not long. The smoother also reclaims storage elements by deleting bytes of interpacket gap one byte at a time from long preambles. This reclamation occurs any time 8 bytes or more of interpacket gap appear at a station.

The smoother is a distributed algorithm that cannot be adequately proven in simulation. We built a gate array that contained the elasticity buffer with the added smoother function to prove in hardware that our new algorithm would operate properly. We built a 200-

a frame, and data packets are lost. Our simulation showed that an unacceptable 10 percent packet loss would occur due to 6 or less bytes of interpacket gap under the following conditions:

node ring of elasticity buffers that could handle at least the test case (101 elasticity buffers) used in the simulation. Each elasticity buffer had a variable oscillator to allow control of the

distribution of clock

frequencies that we felt would induce interpacket gap shrinkage. We also included special test features in the chip to monitor the interpacket gap at every node in the tester.

During four weeks of ring operation, we observed the interpacket gap of 6.72 billion maximum size packets (4500 bytes). The minimum interpacket gap observed was 7 bytes, which resulted in no packet loss. The experiment indicated a packet loss rate of less than 2E-10. Since no packet loss occurred, the actual loss rate is unknown; but this result gives us confidence that the loss rate predictions made by analysis are correct.

The 200-node hardware test bed demonstrated that our algorithm worked effectively. The smoother protocol was adopted as a mandatory part of the final ANSI FDDI PHY standard.[2] The standard allows a variety of different designs, some having as little as one byte of smoothing, depending on the number of preamble bytes required by the MAC implementation in the same station. All allowed designs have worst-case loss rates below 1E-14 (by analysis) in a homogeneous ring. The PHY protocol, the FDDI distributed clocking scheme does not significantly contribute to packet loss.

FDDI Optical Link Design

As noted in the earlier section, Operation of the Physical Layer, the physical connection is the basic element in the FDDI LAN topology. The physical connections between FDDI stations are full duplex, fiber-optic links that deliver a serial code bit stream from one station to another with a bit error rate (BER) less than 2.5E-10. Each station has a separate transmit and receive link, and both links are cabled together to the same destination. The optical link requirements are defined and measured at the MIC. Any set of conforming FDDI stations connected together with a compliant cable plant in a legal topology are guaranteed to provide the required transmission service. Conformance to the optical requirements can be measured independently of both the interconnecting media and the attached station. Measurements can be taken from either end of a physical connection.

The technology choices we confronted and the design methods we used in the

worst-case packet loss rate in a heterogeneous ring, one with multiple types of smoother designs, has a packet loss rate below 1E-10 (by analysis). Given the addition of the smoother to

development of the optical link are summarized in the following sections. These methods can be applied to any transmission system design problem with

similar requirements. The Physical Layer Medium Dependent (PMD) Working Group of the FDDI committee adopted these methods, and Digital played a leading role in the design of the PMD Standard. The Working Group developed the design in a manner that combined theoretical analysis with empirical modifications in an iterative process to arrive at the specifications for the system. The full detail of the models has been documented previously in the literature.[7,8]

Technology Choices

The FDDI distance and bit rate requirements clearly mandate the use of a fiber-optic transmission system. However, the choices are not equally obvious between laser- or LED-based transmitters, between 850-nanometer (nm) and 1300-nm operation, and between single-mode and multimode fiber operation. FDDI development initially focused on the transmission distance requirements of LANs which serve as local office networks as well as network backbones. Accordingly the technology chosen for the optical link should be cost-effective and capable of spanning approximately 2-km distances. For multimode fiber prevailed over lasers operating with single-mode fiber because at the design time the former was more reliable and had a better chance of achieving the cost goals required by short-distance office interconnection spans. The selection of appropriate technology was especially difficult because the technology was rapidly evolving. The Working Group made the basic technology choices in the 1984-1986 time frame; the chosen technology represented the best compromise between available technology and reasonable anticipated improvements. The FDDI committee later addressed the long-distance requirements (greater than 2 km) of a campus LAN with a single-mode fiber and laser transmitter PMD (SMF-PMD). That development effort is not addressed in this paper.

Optical Link Overview

The optical link is composed of three basic elements: a transmitter, a cable plant, and a receiver. The transmitter is provided with a serial 125-megabaud code bit stream and creates an amplitude modulated 1300-nm optical version of the bit stream. The code bit

these applications, the superior bandwidth and loss characteristics of 1300-nm LED systems prevail over 850-nm LED technology; 125-megabaud transmission over 2 km is not possible with 850-nm LEDs. LEDs and

stream has previously been encoded with a 4-bit into 5-bit (4B/5B) non return to zero invert (NRZI) coding scheme that ensures that the serial sequence has sufficient transitions to allow recovery of the

transmit station's timing clock at the distal end of the link. The cable plant uses a glass, graded-index, multimode optical wave guide to ferry the signal to the receiver; the cable has an arbitrary number of junctions (e.g., connectors). The cable plant is described by its optical loss and bandwidth. The receiver in turn converts the optical signal back into a logic-level code bit stream. When a station is not sending data, the transmitter is provided with special code bit sequences which ensure that there is always an optical signal on the medium between packet transmissions. Thus, whenever the link is a part of a ring, the optical system stays in its equilibrium operating conditions, and the clock recovery circuit is always synchronized with the incoming stream.

Design Methods

The design of any digital transmission system must provide sufficient end-to-end bandwidth and signal power. Further, the design must demonstrate bounded jitter characteristics in order to provide the required data transfer at the desired BER. The bandwidth allocation, jitter budget, and loss

Bandwidth Allocation and Models Nyquist communications theory requires a system bandwidth of at least one half the baud rate to prevent error rate degradation due to intersymbol interference. Practical systems require a somewhat greater bandwidth. We determined the LED-fiber bandwidth for FDDI by measuring the sensitivity of commercial 125-megabaud optical receivers as a function of increasing input rise time (decreasing bandwidth) and by observing when the channel bandwidth started to cause a penalty in the measured receiver BER performance. The 0.5 decibel (dB) optical power penalty point was found at 95 MHz. That point is the bandwidth requirement for the LED and fiber combination in a worst-case maximum length link; lower bandwidth causes increasingly higher penalties in BER performance and must be prevented.

The bandwidth of an LED and multimode fiber optical system is modeled with three components which add in a root mean square (RMS) fashion as shown in equation 1 in Figure 5. The design problem confronted is as follows: how are the three different bandwidth components rationally allocated to

budget for the FDDI optical system are described next.

meet the 95-MHz LED-fiber requirement, and what is the maximum distance that can be achieved and still meet this requirement? Although the electrical and

modal bandwidth limitations are well known (equations 2 and 3 in Figure 5), the chromatic bandwidth limitation caused by the LED and fiber combination was not well understood.

Chromatic bandwidth limitation is caused by the interaction of the LED spectral width with the wavelength dispersion of the glass fiber. Thirteen hundred-nm LEDs are not monochromatic; their emission spectrum is typically 170-nm wide at the half optical power point. The propagation velocity of light in glass is a function of the wavelength of the light; light of different wavelengths experiences differential delay or dispersion. Accordingly a signal of appreciable optical spectral width experiences dispersion that causes an increase in the signal transition times and limits the bandwidth. The amount of dispersion experienced by a pulse is a function of the length of the fiber, of the optical spectral width, and of the separation of the pulse central wavelength from the zero dispersion wavelength of the fiber. The wide spectral width of 1300-nm LEDs is sufficient to cause systems based on their bandwidth; the equation for the model is 4 in Figure 5. Equations 1 through 4 are the complete model for the bandwidth of the FDDI optical system. The inputs to the model are

the transmitter spectral center wavelength, spectral width, transmitter rise and fall times, the fiber length, the fiber modal bandwidth, the fiber's zero dispersion wavelength ((0)), and the zero dispersion slope (S(0)). These parameters completely define the constituents of the bandwidth of a multimode fiber-optic transmission system. In an iterative sequence of calculations with the model, we evaluated a trade-off of fiber length, fiber modal bandwidth, LED chromatic attributes, and LED rise and fall times to arrive at a 2-km maximum fiber length with transmitter chromatic and temporal requirements that could reasonably be met by vendors. The transmitter requirements were described by a series of curves that balanced transmitter rise and fall times and chromatic attributes. These requirements guarantee the 95-MHz LED-fiber bandwidth requirement for a 2-km fiber. Thus transmitters are allowed

use to be distance limited by chromatic dispersion, even though the system is operating at 1300 nm, which is the nominal zero dispersion window of fiber. A model was developed and verified for the chromatic slow rise times if they have narrow spectral widths or central wavelengths that match the minimum dispersion wavelength of the fiber. Transmitters with wider spectral widths and central wavelengths

displaced from the zero dispersion wavelength have fast rise time requirements. The curves in ANSI FDDI PMD Figure 9 show the final allowed transmitter spectral and temporal trade-offs.[3] They were generated with a slight modification to the basic model described above that used explicit fast Fourier transform (FFT) descriptions of the LED electrical bandwidth component. The transmitter requirements depend on the fiber meeting modal bandwidth and chromatic dispersion specifications. We empirically established the 500 MHz·km minimum modal bandwidth distance product requirement and the allowed range of dispersion parameters shown in the ANSI PMD Figure 14.[3]

Jitter Budget In most high-speed serial digital communications systems, the clock used to recover the received data must be extracted from the bit stream. The recovered clock is used to sample the data, and the sampling transition is nominally in the middle of the bit interval. If the sampling clock location overlaps with the signal transition between bits, errors occur. Jitter is time dither of the bit stream signal transitions; the measured a probability equal to the BER requirement.

A jitter budget tracks the accumulation of jitter in the bit stream edge position and allocates it to different components. The budget ensures there is a jitter-free opening, or window, for the placement of the sampling clock. Jitter consists of three basic types:

- Duty cycle distortion-DCD
- Data dependent jitter-DDJ
- Random jitter-RJ

DCD is static and is caused by switching threshold variation and mismatched rise and fall times in driver circuits. DDJ is caused by bandwidth limitations in transmission components and is also a function of the transmitted code bit stream. We developed a worst-case test pattern that evinces high-frequency DDJ components caused by local run length variations in the transmitted bit stream and low-frequency DDJ components caused by variations in the average power of the unbalanced 4B/5B code bit stream. RJ is caused primarily by thermal noise corrupting the signal in receivers and is apparent at low optical

value is a function of the probability of its occurrence. Because jitter is the predominant source of communications system error, it is measured at powers. RJ adds in a root-mean-square fashion with other RJ components; DDJ and DCD add linearly to RJ.

The FDDI jitter budget tracks these three components of jitter through the optical link. The budget ensures a sufficient allocation for the clock recovery implementation to place the clock correctly in the jitter-free window to retime the data. The specific values of jitter allotted to each link element were determined largely by empirical methods. The sum of all jitter allocations must not exceed the code bit width (8 ns). Table 1 summarizes the jitter budget, showing the totals for each jitter component as it adds through the link.

Only the jitter components visible at the PMD MIC (PMD out and PMD in) are enforceable parts of the standard. Note the sum of the jitter components at PHY in (the exit of the receiver function) is 5.87 ns, leaving a 2.13-ns jitter-free window remaining in the 8-ns bit cell. This window is allocated to the static alignment error and RJ of the clock recovery implementation. Digital jitter-free window is the measured receiver component requirement.

Table 1 FDDI Jitter Budget Example (ns Peak to Peak)

Measurement Point	DCD	DDJ	RJ
PHY out	0.4	0.0	0.32
PMD out	1.0	0.6	0.76
PMD in	1.0	1.2	0.76
PHY in	1.4	2.2	2.27

Optical Loss Budget The optical loss budget for FDDI is the difference between the minimum optical power launched into the fiber and the minimum optical power required at the receiver. Decreased optical power in a receiver causes a reduction in the signal-to-noise ratio and is evinced on the serial data stream as an increase in its RJ. The optical power requirements are defined in terms of the performance measured with 62.5-micron-core multimode fiber. The fiber core size is specified because the launch power of a particular transmitter is a function of the fiber type

developed specialized test equipment to generate and receive the DDJ test pattern and to signal received bit errors; the error rate at the worst-case optical conditions (minimum power, maximum jitter) was measured as a function of clock sampling position to measure the jitter-free window at the receiver exit. The 2.13-ns

used. This fiber type is the prevalent standard for fiber-optic LANs and with it an 11-dB loss budget is provided for the optical link.

The 11-dB loss budget is apportioned by a user between bulk fiber losses (1.5 dB/km typical, 2.5 dB /km worst case), connector losses (0.6 dB typical, 1.0 dB worst case) and splice losses (0.2 dB typical, 0.5 dB worst case). With this loss budget, users can construct cable plants of up to 2 km in length with any number of connectors and splices, provided the total loss is less than 11 dB. There is no minimum loss required because the maximum launch power is equal to the maximum input power; stations may be operated back to back without saturating the receiver function.

In summary, the design methods we used guaranteed the bit error rate of the serial data stream transmission between stations. The optical bandwidth was allocated and guaranteed by design to prevent BER degradation due to intersymbol interference; the jitter accumulation from different link elements was budgeted to prevent BER degradation due to received data sampling errors, and an optical power budget was defined to control BER degradation due to inadequate receiver signal-to-noise ratio.

Physical Link Error Process

An important part of the physical layer development was the analysis of the media bit error processes. In the previous section, we presented the design of the optical link to control the bit error rate. This section considers the effect of the error process and the isolation of certain types of faults that cause errors.

To evaluate the error process, we had to know the source of the errors and study their effect on the protocols for the FDDI. The error process must be considered in light of two metrics:

- Correctness of the protocol. Error events may lead to undetected corruption of user data. Detected errors reduce performance, but an undetected error may have nearly unbounded bad effects for the user. Therefore the undetected error rate must be very low.
- Isolation of error source to a component of the network when the error rate is too high. Error rates may exceed acceptable levels as the result of a misconfigured network or a fault. Isolation

of the problem is the first step in a repair process.

A discussion of protocol correctness and fault isolation must consider more sources of errors than the normal bit error process discussed in the previous section. To provide correctness and fault isolation, the design must account for misconfigured links and common faults. A misconfigured LAN may provide poor performance, but it is always unacceptable for a data packet to be delivered with undetected errors.

Good examples of misconfigured links include those with cables that are too long and that use too many connectors in the cable plant. Common faults in the system include transmitters that are too dim, dirty or partially plugged connectors, and cables that are kinked (for example, by a misplaced chair leg). One can write an endless list of possible faults and can posit a fault with an arbitrarily complex symptom. The faults listed above are important because they are likely to occur during normal use of the components. Many of these faults can be traced to a careless or uninformed user. A design must ensure that these external causes of abnormal error rate do

As discussed earlier for the design of an optical link, bit errors are caused by transition jitter resulting from bandwidth and power budget limits. The important faults and misconfigurations reduce the channel bandwidth or increase the optical loss beyond the design limits. The error rate may exceed the design limit but the physics of the error process remains the same.

We analyzed the impact of this error process given the FDDI encoding/decoding and error-detecting protocols. An example of an error event is shown in Figure 6 to illustrate the effect of media noise on the FDDI encoding schemes. The code bits on the media are encoded as NRZI, where a signal transition represents a code bit 1 and a lack of transition (for a bit time) represents a code bit 0. With this encoding, a single noise event results in two code bit errors where the resulting pair of bits are the complement of the original bits. In Figure 6, the pair of code bits 0,1 are changed to 1,0. There are four possible pairs of code bits-00, 10, 01, 11 - that change to 11, 01, 10, 00, respectively, by an error

no lasting damage and that they can be detected and isolated.

The error process resulting from important faults is similar to the error process of a correctly operating optical link. event. The FDDI PHY uses a block code in which 5 code bits represent a symbol, and a symbol contains 4 data bits. In the example, the single error event changes 2 code bits, which results in a decoded symbol

with 4 incorrect data bits. The number of data bits affected by an error event is multiplied by the decoding process.

Error detection is provided by redundancy in the data packet. Errors are detected by the MAC protocol based on a frame check sequence (FCS). The probability of an undetected error is related to the number of error events in the packet and to the specific symbols created. Our analysis, based on a draft of the FDDI MAC protocol, indicated that undetected data corruption could occur with high probability.[9] In the important case, a new frame was created when a noise event changed a data symbol into an ending delimiter and created a smaller frame. This truncation process resulted in an undetected packet error rate of 3E-14 for large rings (500 stations).[10] Our design requirements include the much more strict limit of 1E-21 on this rate. For this reason, an enhancement to strengthen the ending delimiter was proposed and accepted by ANSI X3T9.5 for the MAC protocol.[11] In accord with this enhancement, a frame is valid only if its ending delimiter is followed by a symbol that This enhancement results in an undetected error rate of 5E-24 for the protocols, allowing significant margin for actual implementations.[10]

To isolate a faulty physical link, we need to know which of many links exceeds the design-specified error rate. Each error event must be detected and counted at one point in the topology. Using a traditional method, we would isolate faults based on the information provided by the MAC FCS error counters. Although this method works reasonably well for a bus topology, it is more difficult to use with FDDI topologies. The quantity of physical links may greatly outnumber the MACs in the topology. The errors from more than a single physical link may be counted by one MAC, thus masking which links exceed the error rate. For example, in a wrapped dual ring of single-MAC, dual attachment stations, data errors occurring in only half the physical links in the network would be counted by a single event counter. A similar situation occurs in an FDDI tree topology. The MAC error counters are not associated with a particular physical link.

Fault isolation must be

cannot be created by the noise event that created the ending delimiter. Thereby, undetected corruption was greatly reduced in the final, standard MAC protocol.[12]

based on facilities present for each physical link. For this purpose, we developed a protocol called link error monitor (LEM). LEM takes advantage of the requirement in the PHY

standard that a set of code bit groups representing violation symbols and certain sequences of control symbols not be transmitted (repeated) onto a physical link. Our study of the error process indicated that roughly 30 percent of the error events could be detected by the physical layer decoder.[10] This accuracy is acceptable as BER variations of many orders of magnitude are often the most important. LEM counts the decode violations that are received only at one point in the LAN immediately after the error event occurs. Errors not counted by LEM are those in which the created symbol may be repeated by a PHY port, such as when a data symbol is changed to another data symbol. An instance of LEM protocol may observe each PHY port and detect events associated with a particular physical link.

The accuracy of a LEM BER estimate is comparable to other methods and has the advantage of providing better fault isolation. The accuracy of a LEM estimate is affected by the statistics given above for the error process and the length of packet transmitted on the ring. Generally we MAC FCS error counters as well. For instance, the FCS-based estimate of BER also depends on packet length and additionally on ring utilization. The FCS error counters count errors in valid packets only, so estimates of error rate are strongly affected by ring utilization. The LEM estimate includes error events in tokens, stripped frames, and those that occur during the idle period between packets.

The LEM protocol counts errors and provides a BER estimate for each link in the FDDI LAN. Network management applications may collect this data and identify marginal links within the LAN. In addition, LEM provides a mechanism to automatically eliminate faults from the network. This fault-recovery procedure preserves the integrity of the ring when physical links would otherwise prevent ring operation. The LEM protocol was proposed and included in the draft FDDI Station Management (SMT) proposed standard.[13]

The analysis of the physical layer error process resulted in two important changes that reduce the impact of errors. A change proposed

only assign significance to the order of magnitude of the estimate, i.e., the exponent of the BER written in scientific notation. This type of accuracy problem is shared by BER estimates based on

and adopted in the FDDI MAC protocol greatly reduced the rate of undetected corruption. The isolation of components contributing to a high error rate is facilitated by LEM, now a part of

the draft FDDI Station Management standard. These developments have improved the correctness and maintainability of the FDDI LAN.

Summary

Our development work on the FDDI physical layer provided physical layer components, specifications and new protocols. This paper has described the operation of the FDDI physical layer and the functional partitioning of the chip set. The functional partitioning resulted in greater integration and lower cost for the chip set. Much of the work on the physical layer centered on the need to control the error characteristics of both the constituent links and interplay of many asynchronous links as a system. Three important design problems were solved during the development effort. First, the elasticity buffer and smoother protocols, which were developed for the distributed clocking scheme, resolve data integrity and data loss problems. Second, the design of the fiber-optic link for FDDI required methods to allocate system bandwidth and power margins. The Finally, the analysis of the physical link error process resulted in increased correctness through a reduction of the undetected error rate and enhanced fault isolation provided by the link error monitor, LEM.

Acknowledgments

The authors have reported on the results and work of many individuals. Special acknowledgment is extended to Raj Jain, Don Knudson, Charlie Kaufman, and Herman Levenson for their original contributions to the development and for their help in writing this paper.

References

1. P. Hayden, R. Graham, and W. Hawe, "Fiber Distributed Data Interface Overview," Digital Technical Journal, vol. 3, no. 2 (Spring 1991, this issue): 10-18.
2. FDDI PHY, ANSI X3.148-1988, Token Ring Physical Layer Protocol (New York: American National Standards Institute, 1988).
3. FDDI PMD, ANSI X3.166-1990, Physical Layer Medium Dependent (New York: American National Standards Institute, 1990).

bandwidth, jitter, and loss budgets provided a means to allocate channel margin between individual components and can be applied to the design of many transmission systems.

4. T. Chaney and C. Molnar, "Anomalous Behaviour of Synchronizer and Arbiter Circuits," IEEE Transactions on Computers (April 1973): 421-422.

5. B. Thompson and J. Iannarone, Method and Apparatus for Detecting Impending Overflow and/or Underrun of Elasticity Buffer, U.S. patent no. 4,945,548.

6. C. Kaufman, M. Kempf, and J. Hutchison, Method and Apparatus for Nodes in Networks to Avoid Shrinkage of an Interframe Gap, U.S. patent no. 4,878,219.

7. J. Hutchison and D. Knudson, "Developing Standards for a Fiber Optic LAN - FDDI," SPIE Proceedings, vol. 715 (1986).

8. D. Hanson and J. Hutchison, "LED Source and Fiber Specification Issues for the FDDI Network," IEEE Computer Society Proceedings COMPCON (1987).

9. Draft of proposed standard for FDDI MAC, Accredited Standards Committee (ASC) X3T9.5 /83-16, Token Ring Media Access Control, Rev. 10 (1986).

10. R. Jain, "Error Characteristics of Fiber Distributed Data Interface," IEEE Transactions on Communications, vol. 38, no. 8 (1990): 1244-1252.

11. R. Jain, "FDDI Error Analysis," FDDI X3T9.5 Working Group on SMT, Committee Document SMT-80 (1987).

12. FDDI MAC, ANSI X3.136-1987, Token Ring Media Access Control (New York: American National Standards Institute, 1987).

13. H. Levenson, "Link Error Monitor (LEM)," FDDI X3T9.5 Working Group on SMT, Committee Document X3T9.5/88-251 (October 1989). See also "LEM Error Monitor," Document X3T9.5/88-198 (August 1989).