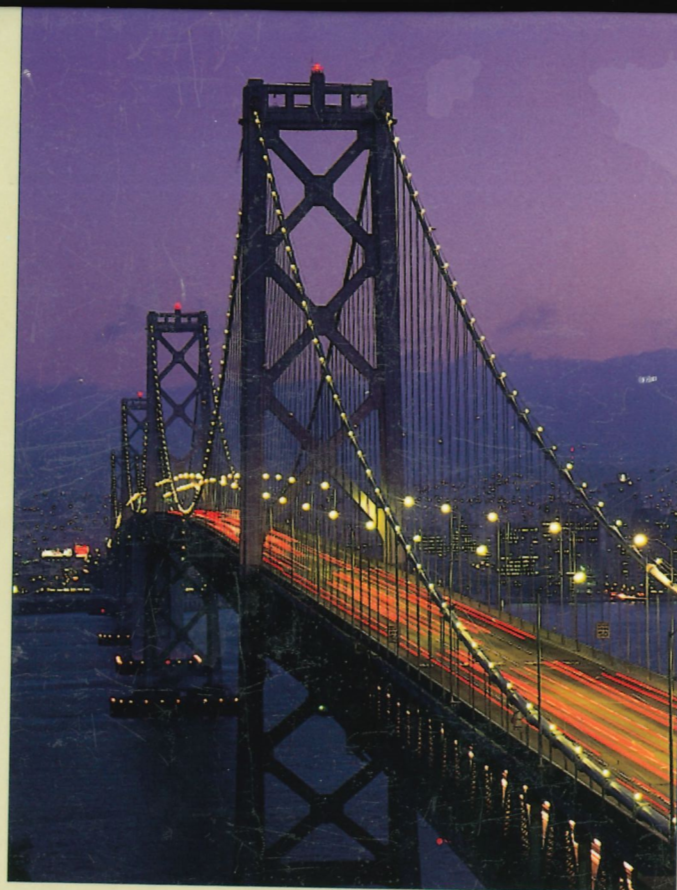


MACMILLAN NETWORK
ARCHITECTURE AND
DEVELOPMENT SERIES



SWITCHED, FAST, AND GIGABIT ETHERNET

THIRD EDITION

Understanding, Building, and Managing

High-Performance Ethernet Networks

Robert Breyer & Sean Riley

MACMILLAN NETWORK ARCHITECTURE
AND DEVELOPMENT SERIES

Switched, Fast, and

Gigabit Ethernet,

Third Edition

Robert Breyer
Sean Riley



Copyright © 1999 by Macmillan Technical Publishing

THIRD EDITION

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without written permission from the publisher, except for the inclusion of brief quotations in a review.

International Standard Book Number: 1-57870-073-6

Library of Congress Catalog Card Number: 98-85496

2002 01 00 99 4 3 2 1

Interpretation of the printing code: The rightmost double-digit number is the year of the book's printing; the rightmost single-digit, the number of the book's printing. For example, the printing code 99-1 shows that the first printing of the book occurred in 1999.

Composed in Palatino and MCPdigital by Macmillan Computer Publishing

Printed in the United States of America

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Macmillan Technical Publishing cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Warning and Disclaimer

This book is designed to provide information about Switched, Fast, and Gigabit Ethernet. Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied.

The information is provided on an as-is basis. The authors and Macmillan Technical Publishing shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the discs or programs that may accompany it.

Feedback Information

At Macmillan Technical Publishing, our goal is to create in-depth technical books of the highest quality and value. Each book is crafted with care and precision, undergoing rigorous development that involves the unique expertise of members from the professional technical community.

Readers' feedback is a natural continuation of this process. If you have any comments regarding how we could improve the quality of this book, or otherwise alter it to better suit your needs, you can contact us at ciscopress@mcp.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher

Jim LeValley

Executive Editor

John Kane

Managing Editor

Caroline Roop

Acquisitions Editor

Brett Bartow

Development Editor

Kitty Wilson Jarrett

Project Editor

Theresa Mathias

Copy Editors

Hugh Vandivier

Keith Cline

Indexer

Tina Trettin

Acquisitions Coordinator

Amy Lewis

Manufacturing Coordinator

Brook Farling

Book Designer

Anne Jones

Cover Designer

Karen Ruggles

Proofreader

Megan Wade

Production

Steve Balle-Gifford

Darin L. Crone

TABLE 3.5 SUMMARY OF DIFFERENT 10Mbps ETHERNET/802.3 PHY STANDARDS

Parameter	10BASE5	10BASE2	10BASE-F	10BASE-T
Maximum segment length	500m	185m	Varies from 400m to 2000m	100m
Topology	Bus	Bus	Star	Star
Medium	50- Ω thick coax	50- Ω thin coax	Multimode fiber (a few vendors provide for single-mode equipment, too)	100- Ω UTP
Connector	NICDB15	BNC	ST ¹	RJ-45
Medium attachment	MAU bolted to coax	External or on NIC	External or on NIC	External or on NIC
Stations/cable segment	100	30	N/A	2 (NIC and repeater)
Maximum segments	5	5	5	5

¹ ST is the official connector, but SC and MIC are actually more popular.

Half- and Full-Duplex Ethernet

Regular Ethernet is a shared-media access method. All networks that utilize a MAC are based on a shared-media operation (hence the term *media access*, which implies that the transmission media are shared and access is negotiated and not always given).

By definition, a shared-media LAN transmission method also implies half-duplex operation. *Half-duplex* means that a station is either transmitting or receiving, but not both, at the same time. That's because in CSMA/CD, a station has to listen to see whether the channel is available, and only if it is can a station start transmitting. When one station is transmitting data, all others have to listen. Therefore, it is an either-or situation for all stations on the LAN. This method of operation was very efficient for the coaxial cable on which Ethernet grew up. Running a single coaxial cable throughout an entire office and providing everyone an access opportunity to the cable every few milliseconds was very efficient. Sharing the transmission media also brought with it collisions, something that has given Ethernet a bit of a bad name. The truth is that collisions are a very effective and efficient method of preventing overload.

Full-duplex, on the other hand, means that a station is simultaneously transmitting *and* receiving. In the early 1990s, many events occurred that have made Full-Duplex Ethernet a reality:

- The introduction of 10BASE-T wiring offered the capability for separate transmit and receive data paths. Before the arrival of 10BASE-T, coaxial cable didn't offer this capability; only one electrical (coax) wire made simultaneous transmission and reception impossible.

Note

Technically speaking, a single cable can carry separate transmit and receive data streams in two directions. This is known as dual-duplex and requires a specially designed physical layer. The electronic circuitry required to do this is very complex, and typically requires a sophisticated DSP. Until recently, the cost and complexity associated with DSP technology was considered prohibitive. 100BASE-T, however, will utilize simultaneous transmit and receive on four pairs of UTP. With 10BASE-T, the existing physical layer could be kept for full-duplex transmission.

- The emergence of multiport Ethernet bridges or switches meant that the physical media were no longer being shared by multiple users but were increasingly being used to connect two switches or a switch and a NIC together in a point-to-point manner.

In 1992, Kalpana seized on this opportunity so that it could effectively double the speed of Ethernet through full-duplex transmission. Kalpana started working with several other industry vendors to establish a de facto industry standard for full-duplex Ethernet over UTP wire. Kalpana proposed the following scheme:

- One pair of UTP wire (or one fiber strand) would be used exclusively for transmission and one for reception of data.
- A new kind of multiport bridge or switch would be required at both ends of the wire.

The Ethernet MAC algorithm would be modified as follows:

- No carrier-sense (CRS) would be required, because a cable pair would now be dedicated for both transmission and reception.
- Similarly, no collision-detection (CDT) jam or exponential backoff are required because collisions are no longer present: They happen in a multiple-user segment only.

Essentially, *full-duplex* Ethernet means running Ethernet without the CSMA/CD MAC in operation (refer to Figure 3.2, without steps 1, 2, 3, 5, and 6).

Ethernet frames are transmitted and received simultaneously on two pairs of UTP or fiber at any time. The only possible problem is one of data overflow. We

can deal with this issue through flow control. Kalpana proposed that artificial collisions be generated by the receiving station if an overload condition exists.

Figure 3.7 illustrates the evolution of Ethernet from a shared-media coaxial-cable LAN with collisions to a dedicated media full-duplex technology.

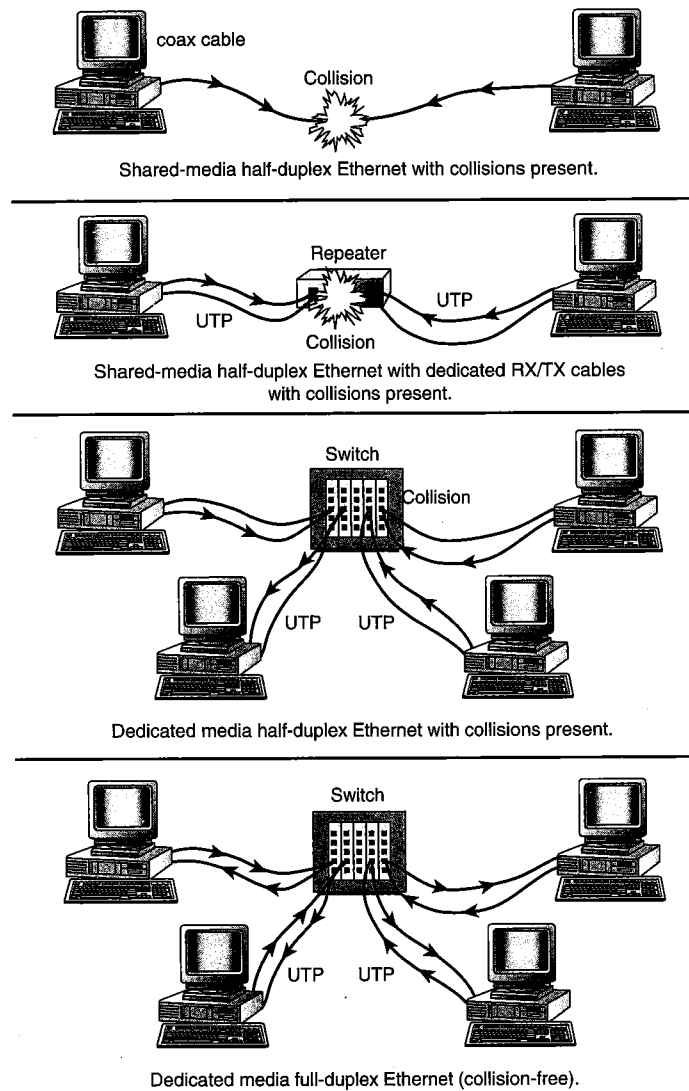


FIGURE 3.7 The evolution of Ethernet from shared-media, half-duplex to dedicated media full-duplex.

Note

A two-user switched connection is a prerequisite to Full-Duplex Ethernet because full-duplex requires a point-to-point connection with only two stations present. Note that Switched Ethernet does not automatically imply full-duplex operation.

Full-duplex has played an increasingly important role with the introduction of Fast and Gigabit Ethernet. In 1997, the 802.3X full-duplex/flow control standard was approved. (Chapter 4 covers this in more detail.) Some people in the IEEE even advocated making Gigabit Ethernet a full-duplex-only technology, which means dropping the all-too-familiar CSMA/CD MAC concept altogether. In the end, the half-duplex followers prevailed, and CSMA/CD operation was included in the standard. It is doubtful whether we will ever see any half-duplex shared Gigabit Ethernet equipment in the marketplace, however.

A trend is currently in place to move from repeated half-duplex to Switched full-duplex Ethernet. With the advent of Gigabit Ethernet in particular, it looks like CSMA/CD and collisions are heading for extinction. What will live on is the language of Ethernet, namely the 802.3 frame format. Our well-known Ethernet frames will be zooming along at 1, 10, 100, 1000, and someday 10000Mbps.

100BASE-T/Fast Ethernet

100BASE-T is the 100Mbps version of the classic Ethernet standard. The IEEE officially adopted the new IEEE 802.3u Fast Ethernet/100BASE-T specification in May 1995. The salient features of 100BASE-T are as follows:

- The 100BASE-T MAC uses the original Ethernet MAC operating at 10 times the speed.
- The 100BASE-T standard is designed to include multiple physical layers. Three different 100BASE-T physical layers are part of the 802.3u standard: two for UTP and one for multimode fiber. An additional UTP PHY was added later as the 802.3y specification.
- Like 10BASE-T and 10BASE-F, 100BASE-T requires a star-wired configuration with a central hub.
- 100BASE-T also includes a specification for an MII, a 100Mbps version of today's AUI. The MII layer is a digital interface connecting MAC and PHY and allows for external transceivers.

The differences between 10BASE-T and 100BASE-T are in the PHY standards and network design areas. That's because the new IEEE 802.3u 100BASE-T specification contains many new rules for repeaters and network topology. Figure 3.8 provides an overview of the new IEEE 802.3u standard.

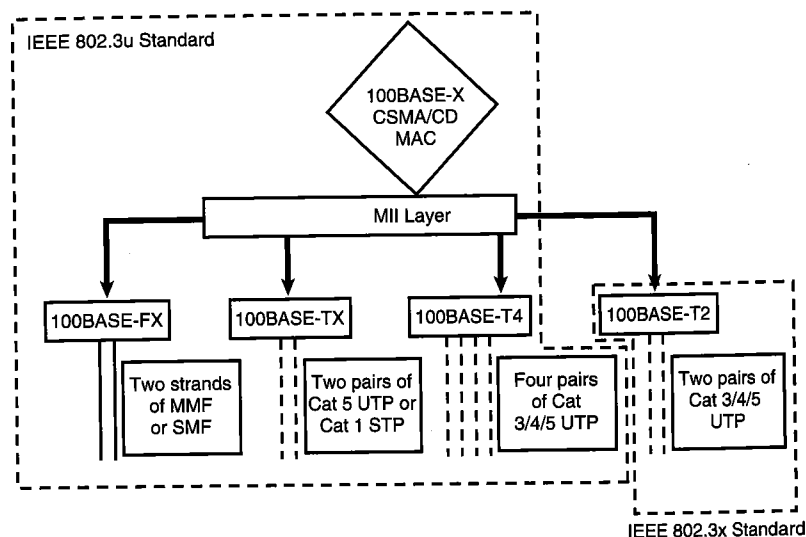


FIGURE 3.8 An overview of the 100BASE-T 802.3u standard showing MAC, MII, and the three original PHYs. (100BASE-T2 was added as 802.3y a year later.)

The Fast Ethernet CSMA/CD MAC

The 100BASE-T MAC is almost identical to the 10Mbps “classic” Ethernet MAC. As mentioned earlier, the 802.3 CSMA/CD MAC is inherently scalable, which means that it can run at different speeds and be interfaced to different physical layers. StarLAN/1BASE5 took advantage of this scalability to run Ethernet at 1Mbps. Table 3.6 compares 10Mbps and the new 100Mbps Ethernet MAC standards. Note that the 100Mbps Ethernet MAC retains all the 10Mbps Ethernet MAC parameters except for InterFrameGap, which has been decreased to one-tenth its original value, from 9.6μs to 0.96μs. A Fast Ethernet frame has the same framing format as a 10Mbps Ethernet frame, except that it is transmitted across the wire at 10 times the speed.

TABLE 3.6 10Mbps COMPARED TO 100Mbps ETHERNET MAC PARAMETERS (MOST VALUES ARE IDENTICAL WHEN MEASURED IN BIT TIMES, BUT ONE-TENTH THE VALUE WHEN MEASURED IN REAL TIME)

Parameter	Ethernet/802.3	Fast Ethernet/802.3u
slotTime	512 bit times	Same ¹
Minimum InterFrameGap	96 bit times (9.6μs)	Same (=0.96μs)
attemptLimit	16 (tries)	Same
backoffLimit	10 (exponential number)	Same
jamSize	32 bits	Same

Parameter	Ethernet/802.3	Fast Ethernet/802.3u
maxFrameSize	1518 bytes	Same
MinFrameSize	64 bytes (512 bits)	Same
AddressSize	48 bits	Same

¹ Same means same value in bit times as 10 Ethernet (for example, the value in the column to the left).

You can calculate the 100BASE-T bit time as follows:

$$1 \text{ bit-time} = \frac{1 \text{ bit}}{100\text{Mhz}} = 10\text{ns}$$

Fast Ethernet PHYs

As with 10BASE-T, 100BASE-T combines the CSMA/CD MAC with different physical layer specifications. The IEEE 802.3u specification contains three new physical layers for 100Mbps Ethernet:

- 100BASE-TX, which requires two pairs of Category 5 UTP or Type 1 STP cabling
- 100BASE-FX, which uses two strands of multimode fiber
- 100BASE-T4, which requires four pairs of Category 3 or better cabling

Then there is the IEEE 802.3y 100BASE-T2 PHY standard, which was completed about a year after the original Fast Ethernet standard. 100BASE-T2 utilizes only two pairs of Category 3 cable, but is virtually unheard of. 100BASE-TX and 100BASE-FX are the most popular 100BASE-T physical layers. They were based on the FDDI/CDDI physical layer, utilizing the 4B/5B encoding method. (Chapter 6, "Cabling and More Physical Layer Details," gives more information on the 4B/5B encoding method.) 100BASE-TX is the de facto standard for high-speed horizontal connections; 100BASE-FX is becoming increasingly popular for vertical/backbone connections. 100BASE-T4 is rarely used, whereas we know of no 100BASE-T2 products that are shipping.

Like 10BASE-T and 10BASE-F, all 100BASE-T PHY specifications require a star configuration with a central hub. All 100BASE-T UTP PHYs use the RJ-45 connector and specify a maximum 100-meter cable distance.

This section provides an overview of the different 100BASE-T physical layers.

100BASE-FX: Fast Ethernet for Fiber-Optic Cabling

100BASE-FX (like 100BASE-TX) borrows its physical layer from the ANSI X3T9.5 FDDI physical layer media-dependent (fiber PMD) standard. 100BASE-FX utilizes two strands of the multimode (62.5/125µm) fiber cabling made popular by FDDI. The maximum segment length for fiber-optic connections varies.

For two switches or a switch-adapter connection using multimode fiber, 412 meters is allowed. You can increase this number to 2000 meters if the link is full duplex. (The IEEE did not specify SMF as part of the 802.3u standard, but numerous vendors are selling SMF-based equipment with distances ranging from 10km to 20km.) 100BASE-FX repeater diameters can be up to 320 meters.

100BASE-TX: Fast Ethernet for Category 5 UTP

100BASE-TX is based on the ANSI-developed copper FDDI twisted-pair physical layer-dependent sublayer technology (TP-PMD), also known as CDDI. 100BASE-TX has numerous similarities to 10BASE-T. It utilizes the same two pairs of unshielded twisted-pair cable, the same RJ-45 connector, and the same maximum segment length of 100 meters. As in 10BASE-T, one pair of wiring transmits data, and the second receives data.

100BASE-T4: Fast Ethernet for Four-Pair Category 3 UTP

100BASE-T4 was the only new PHY developed for 100BASE-T. It was designed to cater to the huge installed base of Category 3 voice-grade wiring. 100BASE-T4 utilizes four pairs of voice or data grade unshielded twisted-pair Category 3 or better cable. Unlike 10BASE-T and 100BASE-TX, no separate dedicated transmit and receive pairs are present, so full-duplex operation is not possible. Note that all four pairs need to be terminated for 100BASE-T4 operation.

Why Did T4 Flop?

Both 100BASE-TX and 100BASE-FX were based on the FDDI/CDDI physical layer technology. Originally, the 100Mbps MAC and these two original PHYs were collectively known as 100BASE-X. (The X was chosen because Fast Ethernet was based on the 100Mbps Ethernet MAC coupled to the FDDI PHY technology.) 100BASE-X is also synonymous with prestandard products. When the IEEE standard was approved, however, everyone quickly adopted the official designation terminology.

Borrowing from the existing ANSI FDDI technology meant that the IEEE 802.3u standard could be finished relatively quickly. More importantly, however, 100BASE-X products were able to get

to market very quickly. FX and TX products could be built and sold *before* the actual 802.3u standard was officially ratified because existing FDDI/CDDI semiconductor and magnetics components were available to build TX or FX NICs, hubs, and switches.

T4 was also part of the original 802.3u specification and was supposed to cater to the huge installed base of Category 3 voice-grade wiring. Unfortunately, T4 has been a market flop because T4 products only started shipping a year after the standard was approved. This two-year delta gave TX products a head start. As a result, 100BASE-TX is by far the most widely used physical layer specification for

continues

continued

100BASE-T today, accounting for more than 95% of all Fast Ethernet shipments.

T4 required four pairs, whereas both 10BASE-T and 100BASE-TX only utilized two pairs. This issue caused a lot of confusion. The original EIA/TIA 568 cabling rules specify four pairs to be installed, which meant using two pairs for 10BASE-T and two spare pairs. Yet some people chose to use the two spare pairs for voice or only installed two pairs in the first place,

making the cabling plant unsuitable for T4.

In some ways, T4 was also a defensive marketing effort in the controversy with 100VG-AnyLAN, covered in Chapters 1 and 2. At the time, the 100VG advocates made a big deal of the fact that 100BASE-TX required Category 5 cable, whereas 100VG ran on four pairs of Category 3. The adoption of the IEEE T4 standard effectively neutralized the 100VG argument. Later on, the 100VG people adopted a two-pair Category 5 version as well, but 100VG had already lost the battle by that time.

100BASE-T2: Fast Ethernet for Two-Pair Category 3 UTP

100BASE-T2 is virtually unknown. It is an IEEE standard, but no products ever made it to market. We wanted to cover it mainly for completeness and because 100BASE-T2 technology will reappear soon with the 1000BASE-T PHY. IEEE engineers came up with 100BASE-T2 because they thought 100BASE-T4 had two shortcomings they could improve. First, T4 requires four pairs of Category 3 wire, but some installations only have two pairs present or usable. The other issue with T4 is that it cannot do full-duplex. When 100BASE-T4 was designed, it was thought that four pairs would be required to transfer 100Mbps. Due to advances in digital signal processing (DSP) and integrated circuit chip technology, it has become possible to transfer 100Mbps over only two pairs of Category 3 UTP. The IEEE worked for about two years on the 100BASE-T2 standard. By the time the standard was complete, however, 100BASE-TX already dominated the market to such an extent that no T2 products were ever built. The following are the main features of 100BASE-T2:

- Utilizes two pairs of voice or data grade unshielded twisted-pair Category 3, 4, or 5 cable.
- Uses both pairs for simultaneous transmitting and receiving on both pairs, also known as dual-duplex.
- Uses a more complicated five-level encoding scheme called PAM5x5. (PAM stands for pulse amplitude modulation.)

TABLE 3.7 A SUMMARY OF THE DIFFERENT 100BASE-T PHYSICAL LAYERS

PHY Parameter	10BASE-T	100BASE-TX
Origin of Technology		CDDI TP PMD (ANSI X3T9.5)
IEEE standard	802.3i-1990	802.3u-1995
Encoding	Manchester	4B/5B
Cabling required	UTP Category 3/4/5	UTP Category 5 or STP Type
Signal frequency	20MHz	125MHz
Number of pairs required	2	2
Number of transmit pairs	1	1
Distance	100m	100m
Full-duplex capable?	Yes	Yes

¹ 150m for repeater-DTE, 412m for DTE-DTE, 2000m for full-duplex DTE-DTE, and 10km for single-mode full-duplex DTE-DTE.

Table 3.7 compares 10BASE-T with the four 100BASE-T physical layer specifications. We discuss the different physical layers and their cabling requirements in more detail in Chapter 6.

Fast Ethernet MII

MII is similar to AUI for 10Mbps Ethernet. The MII layer defines a standard electrical and mechanical interface between the 100BASE-T MAC and the various PHY layers. This standard interface works like AUI in the classic Ethernet world in that it allows manufacturers to build media- or wiring-independent products, with external MAUs being used to connect to the actual physical cabling.

The electrical signals differ between MII and AUI. AUI has a stronger signal, capable of driving 50-meter cable lengths; the MII signals are digital logic-type signals, capable of driving 0.5 meters of cable. MII uses a 40-pin connector, similar to the SCSI connector, although it is smaller. Figure 3.9 shows an external MII 100BASE-FX transceiver. We don't think you will see too many of these because most Fast Ethernet hubs and NICs already include the transceiver.

MII never became as popular as AUI. AUI became a de facto standard because Thick Ethernet hardware wasn't available with an onboard transceiver. That only came later, with 10BASE2 or 10BASE-T. Yet with 10BASE2 and 10BASE-T, the AUI connector provided the capability for backward compatibility through an external transceiver. With 100BASE-T, most hardware manufacturers decided to skip MII and offer different products that integrate the physical layer transceiver already. NIC manufacturers, for example, often sell the same basic design in a 100BASE-TX and in a 100BASE-T4 version. You have to choose up front which media type you need. Think of it as a choice between Thinnet-only or 10BASE-T-only products, as opposed to combo or AUI-only cards using

100BASE-T4	100BASE-FX	100BASE-T2
New; developed from 1995 to 1996	FDDI PMD (ANSI X3T9.5)	New; developed from 1994 to 1995
802.3u-1995	802.3u-1995	802.3y-1996
8B/6T	4B/5B	PAM5x5
UTP Category 3/4/5	Multimode or single-mode fiber	UTP Category 3/4/5
25MHz	125MHz	25MHz
4	2	2
3	1	2
100m	150/412/2000m ¹	100m
No	Yes	Yes

external transceivers. An example of a device with an MII connector is the Intel Pro/100 Server adapter, which uses an MII connector to provide you with the flexibility of either attaching an external fiber or a UTP PHY transceiver.



FIGURE 3.9 Fast Ethernet MII transceivers are similar to the old Ethernet AUI transceivers, but are very rare because most Fast Ethernet hardware includes an onboard transceiver. MII uses a 40-pin connector very similar to SCSI-2.