Cisco Systems

Cisco LAN Switching Fundamentals

The essential guide for understanding Ethernet switched networks

David Barnes, CCIE® No. 6563
Basir Sakandar, CCIE No. 6040

ciscopress.com

Cisco LAN Switching Fundamentals

David Barnes, CCIE No. 6563
Basir Sakandar, CCIE No. 6040

Cisco Press

Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

Cisco LAN Switching Fundamentals

David Barnes
Basir Sakandar

Published by:
Cisco Press
800 East 96th Street
Indianapolis, IN 46240 USA

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

First Printing July 2004

Library of Congress Cataloging-in-Publication Number: 2002100685

ISBN: 1-58705-089-7

Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Cisco Press or Cisco Systems, Inc., 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 Cisco LAN switching. 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 author, Cisco Press, and Cisco Systems, Inc., 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.

The opinions expressed in this book belong to the author and are not necessarily those of Cisco Systems, Inc.

Corporate and Government Sales

Cisco Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact:

U.S. Corporate and Government Sales 1-800-382-3419 corpsales@pearsontechgroup.com

For sales outside of the U.S. please contact: international@pearsoned.com

Feedback Information

At Cisco Press, 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 through e-mail at feedback@ciscopress.com. Please make sure to include the book title and ISBN in your message.

We greatly appreciate your assistance.

Publisher	John Wait
Editor-in-Chief	John Kane
Executive Editor	Brett Bartow
Cisco Representative	Anthony Wolfenden
Cisco Press Program Manager	Nannette M. Noble
Production Manager	Patrick Kanouse
Development Editor	Dayna Isley
Project Editor	San Dee Phillips
Copy Editor	Katherin Bidwell
Contributing Author	John Tiso
Technical Editors	Richard Froom and Geoff Tagg
Team Coordinator	Tammi Barnett
Cover Designer	Louisa Adair
Composition	Octal Publishing, Inc.
Indexer	Brad Herriman

CISCO SYSTEMS

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the **Cisco.com Web site at www.cisco.com/go/offices.**

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia • Czech Republic
Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland • Israel • Italy
Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal
Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

Printed in the USA

About the Authors

David Barnes is a senior manager for Cisco Advanced Services based in Richardson, Texas. He is CCIE No. 6563, MCSE+I, Master CNE, and a Certified Technical Trainer. David manages the Advanced Services team, which is responsible for providing technical expertise on all Cisco routing and switching products to many of the largest Cisco customers. He began his career at Cisco Systems, Inc., as a network consulting engineer specializing in LAN Switching in 1999. In the 10 years before he joined Cisco, David designed, implemented, and managed networks for numerous Fortune 500 companies.

Basir Sakandar, CCIE No. 6040, has dual certifications in Routing and Switching and in Security. He received his undergraduate degree from University of Southern California (USC) in 1997. His current role at Cisco Systems is an advanced services engineer out of Richardson, Texas, office. He has helped with design and implementation of various technologies for Fortune 100 companies.

Contributing Author

John Tiso, CCIE No. 5162, is a senior consultant for Networked Information Systems (NIS), a Cisco Systems Gold Partner in Woburn, Massachusetts. John has a bachelor of science degree from Adelphi University and a variety of industry certifications. John has been published in several industry trade journals and has been a technical editor for Cisco Press and McGraw-Hill for many years. John can be reached at johnt@jtiso.com.

About the Technical Reviewers

Richard Froom, CCIE No. 5102, is a technical leader for the Storage Area Network (SAN) team of the Internet Switching Business Unit—Financial Test Lab at Cisco Systems. Richard has been with Cisco for six years, previously serving as a support engineer troubleshooting customers' networks and a technical leader dealing with Cisco Catalyst products. Richard, being involved with Catalyst product field trials, has been crucial in driving troubleshooting capabilities of Catalyst products and software. He has also contributed substantially to the Cisco.com LAN Technologies Technical Tips and has written white papers dealing with 802.3 auto-negotiation and Hot Standby Router Protocol (HSRP). Richard is currently testing Cisco SAN solutions. Richard is also the coauthor of *CCNP Self-Study: Building Cisco Multilayer Switched Networks (BCMSN)*, Second Edition (ISBN 1-58705-150-8), and *Cisco Catalyst QoS: Quality of Service in Campus Networks* (ISBN 1-58705-120-6) from Cisco Press. He attended Clemson University where he completed his bachelor of science in computer engineering.

Geoff Tagg runs a networking consultancy in the UK, where he has more than 20 years experience of working with companies ranging from small local businesses to large multinationals. Prior to that, he had 15 years experience of systems programming on a variety of mainframe and minicomputers. Geoff's specialty is Internet Protocol (IP) networking over a range of LAN and WAN technologies, including Ethernet, Frame Relay, ATM, and ISDN. Geoff lives in Oxford, England, with his wife Christine and family and is a visiting professor at nearby Oxford Brookes University.

Dedications

David Barnes: For Papa

Basir Sakandar: To my mom and to my wonderful nephews and niece

Acknowledgments

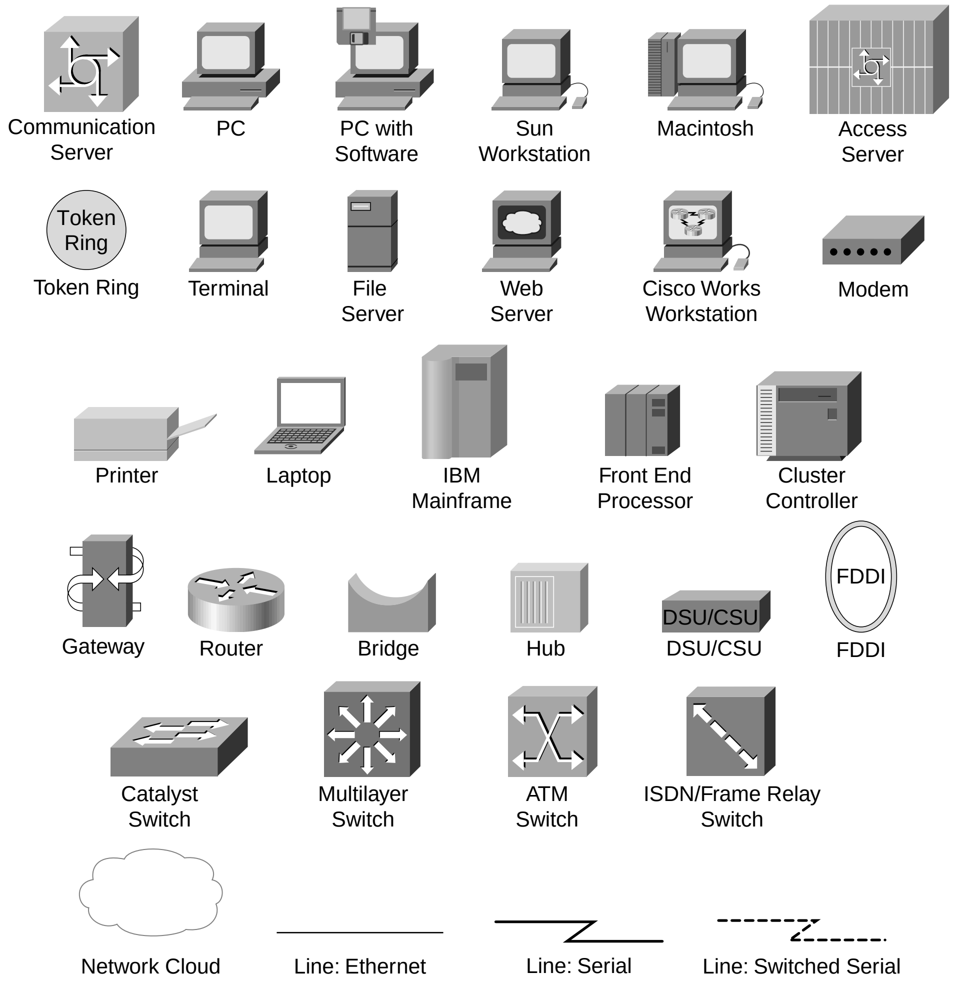
David Barnes: I would like to thank my father for bringing home a computer in 1985 instead of a drum set. While I sometimes still dream of being a rock star, I am glad I am in what I believe is an exciting and ever-changing industry, and I am fortunate to work for the best company in the world—Cisco Systems. Thanks goes to John Tiso for his ideas and for getting this started, Richard and Geoff for their efforts to make this a better book, and Brett Bartow and Dayna Isley at Cisco Press for all of their assistance on this project.

Basir Sakandar: I would like to thank everyone who participated in developing this book.

Contents at a Glance

Table of Contents

Icons Used in This Book



Command Syntax Conventions

The conventions used to present command syntax in this book are the same conventions used in the *Cisco IOS Command Reference*, as follows:

- **Boldface** indicates commands and keywords that are entered literally as shown. In examples (not syntax), boldface indicates user input (for example, a **show** command).
- *Italics* indicates arguments for which you supply values.
- Square brackets ([and]) indicate optional elements.
- Braces ({ and }) contain a choice of required keywords.
- Vertical bars (|) separate alternative, mutually exclusive elements.
- Braces and vertical bars within square brackets—for example, [x {y | z}]—indicate a required choice within an optional element. You do not need to enter what is in the brackets, but if you do, you have some required choices in the braces.

Introduction

Even as Cisco continues to expand its line of products into areas such as IP telephony, content networking, and security, its LAN switching products continue to account for more than half of the company's overall product sales. Because switching devices are considered by many to be the "plumbing" of every campus networking environment, LAN switching is an area of networking that virtually all network engineers need to understand, at least at a fundamental level. It is for this reason we set out to write the *Cisco LAN Switching Fundamentals* book for Cisco Press. Although a wide variety of books exists on traditional routing topics, we saw a need to write a book for those who might have had their hands on a router before, but did not have as much exposure to Cisco LAN switching products.

Goals and Objectives

Cisco LAN Switching Fundamentals will give you exposure to real-world networking and Catalyst switching best practices. Our desire is to provide much more than simply an overview of the various switch architectures, protocols, and features they support. As a result, we kept a few key objectives in mind when writing the book, including the following:

- Provide background on how the architecture and features of Cisco Catalyst switches evolved since their introduction
- Give enough depth of information for an engineer to become functional in each technology in a relatively short period of time
- Offer enough information on the range of LAN switching topics to serve as a reference for everyday use

Who Should Read This Book?

Cisco LAN Switching Fundamentals is intended for beginning and intermediate engineers looking to understand the architecture, configuration, and operation of Cisco Catalyst switches. You will learn about a wide range of topics including quality of service (QoS), multicast, spanning tree, and native and hybrid software. Many design and configuration examples are provided on a wide range of topics that will hopefully make the book a good reference, and a useful read for those wishing to master the fundamentals of LAN switching.

How This Book Is Organized

Cisco LAN Switching Fundamentals is made up of 12 chapters and 1 appendix, as follows:

- **Chapter 1, "LAN Switching Foundation Technologies"**—Chapter 1 creates the building blocks for the rest of the book. It is essential to have an understanding of the basics before tackling the more rigorous topics such as RSTP, QoS, multicast, configuration, and troubleshooting. Chapter 1 discusses the details of various Ethernet technologies and physical and data link layer protocols. Topics of discussion range from introducing Ethernet 10BASE-T to Gigabit Ethernet, and moving on from Transparent Bridging to Spanning Tree.

- **Chapter 2, "LAN Switch Architecture"**—Chapter 2 is dedicated to understanding switching modes, switching paths, and architectures common to many vendor's switches, not only Cisco. This chapter familiarizes you with the some of the challenges of various hardware implementations and their effects on the network. This chapter is a must read before delving into Chapter 3.
- **Chapter 3, "Catalyst Switching Architecture"**—Many of the newest Cisco switching platforms are discussed here in detail. Topics of discussion include the newest Catalyst 6000/6500, 4500, 3750 platforms and even the legacy Catalyst 5000/5500 platforms. There have been great improvements since the popular Catalyst 5000 product line. You will become intimately familiar with the architectures of these platforms.
- **Chapter 4, "Layer 2 Fundamentals"**—Chapter 4 introduces the concept behind virtual LANs (VLANs), types of trunks, VLAN Trunking Protocol (VTP), and private VLANs. This chapter introduces some very important best practices that have been gathered through experience of both authors and the Cisco community. This chapter clears up some confusion regarding VLAN designs, VTP misconfiguration, manual versus VTP pruning, and the significant role of private VLANs.
- **Chapter 5,** "**Using Catalyst Software"**—This chapter discusses issues such as the difference between native and hybrid code for the Catalyst 6000 product line, hybrid's equivalent command in native code, and the processes in migrating from one code to the other. This chapter also discusses software for the Catalyst 3750 and 4500.
- **Chapter 6,** "**Understanding Multilayer Switching"—**Multilayer Switching (MLS) is still a source of confusion for many beginner and intermediate engineers. This chapter takes you step by step through the MLS process and discusses the various hardware components involved in making MLS work. This chapter also serves as a good reference.
- **Chapter 7, "Configuring Switches"**—This chapter shows you how to configure Catalyst switches. This chapter provides many examples to familiarize you with the hardware. Configuration examples such as trunking, SNMP, and many others are given through this chapter.
- **Chapter 8,** "**Understanding Quality of Service on Catalyst 6500"**—In recent years, QoS has become more prominent in the enterprise network. QoS deployment is still in its infancy stage. The exception is with companies who have deployed Voice over IP (VoIP). In VoIP networks, QoS implementation is critical. This chapter introduces the fundamentals of QoS on the Catalyst switches. This chapter takes you through each crucial step as the packet enters a port, and how the packet travels in the switch, and eventually, out of an egress port.
- **Chapter 9,** "**Implementing Multicast on Catalyst Switches"**—Multicast is another important topic in this book. In this chapter, you are given ample examples and detailed explanations of the more popular multicast features and protocols supported on the Catalyst switches. Some of the more important topics that are introduced are Protocol Independent Multicast (PIM), Internet Group Management Protocol (IGMP) snooping, and Cisco Group Management Protocol (CGMP). This chapter also walks you through how the switch handles and forwards multicast traffic.
- **Chapter 10, "Implementing and Tuning Spanning Tree"**—This chapter describes data link layer features that are available on the Catalyst switches. Features such as BPDU Guard, BPDU Filter, Root Guard, Loop Guard, and so on are discussed. This chapter also covers Rapid Spanning Tree

Protocol (RSTP) and Multiple Spanning Tree (MST). Enterprise companies are currently looking closely at RSTP. Situations exist where spanning tree is necessary because of business or design requirements. RSTP can provide a robust solution in comparison to the legacy Spanning Tree Protocol (STP).

- **Chapter 11, "Design and Implementation Best Practices"—**This chapter focuses entirely on best practices for design and configuration. You are taken through a variety of design options with the goal of providing a stable and available network and avoiding common design pitfalls. The intent of this chapter is to put in practice the theory and practical knowledge gained from the previous chapters.
- **Chapter 12, "Troubleshooting the LAN Switching Configuration"**—This chapter delineates some of the more common steps necessary in troubleshooting Catalyst switches. This chapter introduces common problems that the Cisco Technical Assistance Center (TAC) receives and provides solutions. The strength of this chapter is that it offers you the tools necessary to effectively deal with network problems when they arise.
- **Appendix A, "Catalyst 6500 Series Software Conversion"**—This appendix walks through the options for converting from hybrid software to native software and back again for the Catalyst 6500 series. The appendix begins with an overview of the automated tool for hybrid software to native software conversion, and follows up with the step-by-step manual conversion process using a Supervisor 720. Because no automated tool exists for native software to hybrid software conversion, the appendix shows the step-by-step manual process for converting from native software to hybrid software.

This page intentionally left blank

This chapter covers the following topics:

- OSI model
- Ethernet
- Transparent bridging
- Broadcasts and multicasts
- Spanning Tree Protocol

CHAPTER 1

LAN Switching Foundation Technologies

Anyone responsible for implementing and supporting local-area networks (LANs) is increasingly challenged with understanding the fundamental concepts behind Ethernet switching. Almost all new local-area networking infrastructure is based on some type of Ethernet (10 Mbps/100 Mbps/1000 Mbps). It is also worth noting that newer Ethernet implementations include support for the 10 Gbps rate. This book focuses on the concepts, architecture, configuration, and troubleshooting of Cisco Ethernet switches.

Virtually every discussion of networking begins with at least a mention of the Open System Interconnection (OSI) reference model, and for good reason. The OSI model serves as a useful framework for classifying the characteristics and operation of networking devices into seven categories, or layers. While a detailed examination of the OSI model is beyond the scope of this book, a brief overview is useful in understanding the operation of Ethernet switching.

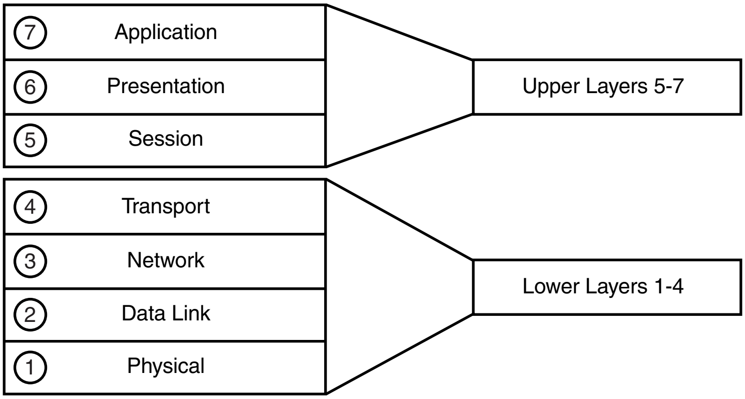
After learning the basic concepts of the OSI model, this chapter moves on to introduce the fundamentals of Ethernet, transparent bridging, and the Spanning Tree Protocol.

OSI Model

The application of a layered framework to networking allows individual layers to be modified, without affecting the layers above or below. The OSI model can be thought of as the networking community's application of the concept of interchangeable parts.

Figure 1-1 illustrates the seven layers of the OSI model. Each layer is tasked with specific functions that allow for the eventual communication of network devices. Note that the model is divided into upper layers and lower layers, which are described in the next sections.

Figure 1-1 *OSI Layers*



OSI Upper Layers

The upper OSI layers provide application level support such as the user interface, data formatting, and communication sessions. The upper layers are as follows:

- **Application**—The layer where applications and users interface with the network. Examples include web browsers, electronic mail, or a word processing program.
- **Presentation**—The layer that controls format translation and provides data encryption and compression. Examples include ASCII and JPEG.
- **Session**—The layer responsible for establishing, maintaining, and terminating sessions between presentation layer entities. Protocols that fall at this layer include NetBIOS and RPC.

OSI Lower Layers

The lower OSI layers define how data moves through the network. Because Ethernet itself and the switching of Ethernet frames are classified in the lower OSI layers, most of the discussion in this book focuses on the lower layers. The lower layers of the OSI model are as follows:

- **Transport**—The layer responsible for error detection and correction, flow control, and data sequencing; also determines the size of the packet. Examples include Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).
- **Network**—The layer responsible for the delivery of data packets. Network layer provides logical addressing and path determination. Examples include Internet Protocol (IP) and Internetwork Packet Exchange (IPX).

- **Data Link**—The layer responsible for access to media, hardware addressing, error detection, flow control, and encapsulation of data into frames. The two major components to Data Link layer are Logical Link Control (LLC) and Media Access Control (MAC). LLC handles error detection and flow control. MAC is responsible for communicating with the adapter card, and the type of media used. Examples include IEEE 802.3 CSMA/CD, 802.12 Demand Priority, and 802.5. Bridges and LAN switches also operate at this layer.
- **Physical**—The layer responsible for defining the electrical properties and physical transmission system. The physical layer is responsible in transmitting and receiving data. Examples include any type of cabling, hubs, repeaters, and fiber optics.

Introducing Ethernet

Ethernet's origins begin with the Aloha Radio System, a packet satellite system developed at the University of Hawaii. Beginning in the late 1960s, the Aloha Radio System was designed to facilitate communication between the university's IBM mainframe, located on the island of Oahu, with card readers located among different islands and ships at sea. Work on the Aloha Radio System proved to be the foundation for most modern packet broadcast systems including Ethernet.

Ethernet as it is known today took shape in the 1970s as a research project at Xerox's Palo Alto Research Center. Ethernet was eventually standardized by Digital, Intel, and Xerox in 1979, and harmonized with the international standard, IEEE 802.3, in 1982.

Modern LAN switched networks are based on the theory and operation of Ethernet. This section discusses the basic theory and operation of Ethernet. The initial version of Ethernet operated with a speed of 3 Mbps and used an algorithm called *carrier sense multiple access collision detect (CSMA/CD)* protocol to determine when a device could use the network. Ethernet is currently available in 10 Mbps, 100 Mbps, 1000 Mbps (1 Gbps), and 10000 Mbps (10 Gbps).

Types of Ethernet

As mentioned earlier, Ethernet provides various data rates with different physical layouts. A variety of Ethernet types have come and gone over the years, such as the following:

- 10BASE5 (Thicknet)
- 10BASE2 (Thinnet)
- 10BASE-FL
- 10BASE-T

In the mid 1990s, 100BASE-T (unshielded twisted-pair [UTP]) and 100BASE-FX (using fiber) were ubiquitous in the enterprise network, and they still are. Since the start of the millennium, enterprise networks have actively implemented Gigabit Ethernet, 1000BASE-T, in their network. The push for today is 10 Gbps in the core of the enterprise network.

Transmission Media

The more common transmission media are twisted pair and fiber optics. Coaxial cable is mentioned in this section for historical purpose. Categories defined under twisted pair support transmission over various distances and data rates. The most common UTP cable in the enterprise network is Category 5, which supports 100 Mbps and 1000 Mbps rates.

Ethernet over Twisted-Pair Cabling

Ethernet technology standards are the responsibility of the IEEE 802.3 working group. This group is responsible for evaluating and eventually approving Ethernet specifications as new Ethernet technologies are developed such as Gigabit and 10Gigabit Ethernet. Although this group defines the standards for Ethernet, it looks to other established standards organizations to define the specifications for physical cabling and connectors. These organizations include the American National Standards Institute (ANSI), Engineering Industry Association (EIA), and Telecommunications Industry Association (TIA). The TIA/EIA published specifications for twisted-pair cabling are found in the TIA/EIA-568-B specification document.

The more common forms of cabling are unshielded twisted-pair (UTP) and optical fiber. Twisted pair cable comes in a variety of forms. The most common categories in today's networks are the following:

- Category 3
- Category 5
- Category 5E
- Category 6

The categories represent the certification of the radio frequency capability of the cabling.

Category 3 was initially designed as voice grade cable and is capable of handling transmissions using up to 16 MHz. Category 5 is capable of handling transmissions up to 100 MHz. Category 5E is an improved version of Category 5; while still limited to 100 MHz, Category 5E defines performance parameters sufficient to support 1000BASE-T operation.

Category 6 provides the best possible performance specification for UTP cabling. Category 6 specifies much stricter requirements for cabling than Category 5 and 5E. The frequency range of Category 6 extends to 250 MHz, in contrast to Category 5 and 5E's 100 MHz. While new cabling installations typically install Category 5E or 6 cabling, Category 5 cabling can be utilized for 1000BASE-T applications. With few exceptions, if 100 Mbps Ethernet is operating without issues up to 100 meters on a Category 5 cable plant, 1000BASE-T will operate as well.

Although 10 Mbps and 100 Mbps Ethernet often use two pairs (pins 1, 2, 3, and 6) of twisted-pair cabling, Gigabit Ethernet over twisted pair uses all four pairs of wiring in the twisted-pair cable.

Even if the actual twisted pair is rated a specific category, it does not imply that a cabling infrastructure properly supports the category specification end-to-end. Installation and accessories (such as patch panels and wall plates) must meet the standard as well. Cable plants should be certified from end-to-end. When installing a cabling infrastructure, the installer should be able to use specialized equipment to verify the specifications of the cabling system from end-to-end.

Ethernet over Fiber Optics

Two major types of fiber used in Ethernet networks are multimode and single mode. Multimode fiber (MMF) is used for short haul applications (up to 2000 m). Examples include campus or building networks. MMF is usually driven by LED or low-power laser-based equipment. Single mode fiber (SMF) is used for longer haul applications (up to 10 km) and the equipment is laser based. SMF is generally used in metropolitan-area networks or carrier networks.

Table 1-1 compares Ethernet types over different transmission media.

Table 1-1 *Comparisons of Ethernet over Various Transmission Media*

Ethernet Type	Media Type	Distance Limitations (meters)	Speed (megabits)	Data Encoding
10BASE-T	UTP Category 3 or better	100	10	Manchester
10BASE-FX – MMF	MMF	2000	10	Manchester
100BASE-TX	UTP Category 5 or better	100	100	4B/5B
100BASE-FX – MMF	MMF	2000	100	4B/5B
100BASE-FX – SMF	SMF	10000	100	4B/5B
1000BASE-SX	MMF	2000	1000	8B/10B
1000BASE-LX	SMF	5000*	1000	8B/10B
1000BASE-T	UTP Category 5 or better	100	1000	PAM 5x5

* The Cisco implementation of 1000BASE-LX doubles this distance over the standard to 10,000 meters.

Ethernet over Coax Cabling

The use of coax cable for LANs is virtually nonexistent. One might run into it in an old abandoned building. Ethernet's eventual support of twisted pair cabling in a star topology virtually ended the use of coaxial cabling for Ethernet. Keep in mind that coax cable was not cheap either. Two major types of coax were used: thinnet (also called cheapernet) and thicknet. Thinnet uses 50 ohm coaxi cable (RG-58 A/U) with a maximum length of 185 meters when used for Ethernet. This cable is thinner and more flexible than thicknet, which is also 50 ohm coax cable. It is packaged and insulated differently than thinnet. It requires a specialized tool, a vampire tap, to pierce into and has a maximum length of 500 meters for Ethernet. The vampire tap was used to pierce the outer shielding of the cable, creating an electrical connection between the device and the shared media. Traditionally, thicknet was used as a backbone technology because of its additional shielding. Both thinnet and thicknet are virtually extinct in production networks today.

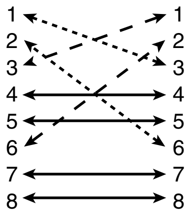
Ethernet Cross-Over Cabling

Network devices can be categorized as either data circuit equipment (DCE) or data terminating equipment (DTE). DCE equipment connects to DTE equipment, similar to the male and female end of a garden hose. DCE equipment usually is a type of concentrator or repeater, like a hub. DTE equipment is usually equipment that generates traffic, like a workstation or host.

Sometimes, it is necessary to connect like equipment. Connecting like devices can be accomplished by altering the twisted-pair media, and taking transmit and receive wires and reversing them. This is commonly called a "cross-over" cable. Figure 1-2 shows an RJ-45 connector with its pinouts. Pins 4, 5, 7, and 8 are not used.

The pinouts are a bit different in a Gigabit scenario because all the pins are used. In addition to the pinouts for 10 Mbps/100 Mbps aforementioned, two additional changes are necessary: pin 4 to 7, and 5 to 8.

Figure 1-2 *Crossover Pinouts*



A crossover cable can link DCE to DCE, and DTE to DTE. The exception to connecting like devices is that some devices are manufactured to be connected together. An example would be that some hubs and switches have an uplink or Media Dependent Interface (MDI) port. There is typically a selector that allows the user to toggle between MDI and MDI-X (X for crossover), with MDI-X intentionally reversing the pin out of transmit and receive similar to a crossover cable. A setting of MDI-X allows two DCE devices, such as two hubs or switches, to connect to each other using a typical straight through wired twisted-pair cable.

Ethernet Topology

Ethernet is defined at the data link layer of the OSI model and uses what is commonly referred to as a bus topology. A bus topology consists of devices strung together in series with each device connecting to a long cable or bus. Many devices can tap into the bus and begin communication with all other devices on that cable segment. This means that all the network devices are attached to a single wire and are all peers, sharing the same media.

Bus topology has two very glaring faults. First, if there were a break in the main cable, the entire network would go down. Second, it was hard to troubleshoot. It took time to find out where the cable was cut off. The star topology has been deployed for a long time now and is the standard in the LAN environment. Star topologies link nodes directly to a central point. The central point is either a hub or a LAN switch. Ethernet hubs are multiport repeaters, meaning they repeat the signal out each port except the source port.

The advantages of a physical star topology network are reliability and serviceability. If a point-to-point segment has a break, in the star topology, it will affect only the node on that link. Other nodes on the network continue to operate as if that connection were nonexistent. Ethernet hubs and LAN switches act as the repeaters that centralize the twisted-pair media. Twisted-pair media can also be used to join like devices. Following the OSI model and the concept of interchangeable parts, even Token Ring, which is a logical ring, can use a physical star topology with twisted pair.

Ethernet Logical Addressing

In Ethernet, LAN devices must have a unique identifier on that specific domain. LAN devices use a Media Access Control (MAC) address for such purpose. MAC addresses are also referred to as *hardware addresses* or *burned-in addresses* because they are usually programmed into the Ethernet adapter by the manufacturer of the hardware.

The format of a MAC address is a 48-bit hexadecimal address. Because hexadecimal uses the digits 0-9 and the letters a-f (for numbers 10-15), this yields a 12-digit address. MAC addresses are represented in any one of four formats. All the formats properly identify a MAC address and differ only in the field separators, as follows:

- Dashes between each two characters: 00-01-03-23-31-DD
- Colons instead of dashes between each two characters: 00:01:03:23:31:DD
- Periods between each fourth character: 0001.0323.31DD
- The digits without dashes, periods, or colons: 0001032331DD

Cisco routers typically use the 0001.0323.31DD formatting, while Cisco switches running Catalyst Operation System (Catalyst OS) images use 00:01:03:23:31:DD to represent the same address.

CSMA/CD Operation

Ethernet operates using CSMA/CD. By definition, CSMA/CD is half-duplex communication. Half duplex implies that only one device on the Ethernet LAN "talks" at a time, and devices connected to the same Ethernet network are considered to be part of the same collision domain. Devices sending traffic in the same collision domain have the potential of their packets colliding with each other when two devices attempt to transmit at the same time. The logical definition of this range of devices is called a *domain*, hence the term *collision domain*.

The old style telephone party line example best illustrates the concept of a collision domain, as shown in Figure 1-3.

Figure 1-3 *Telephone Party Line*

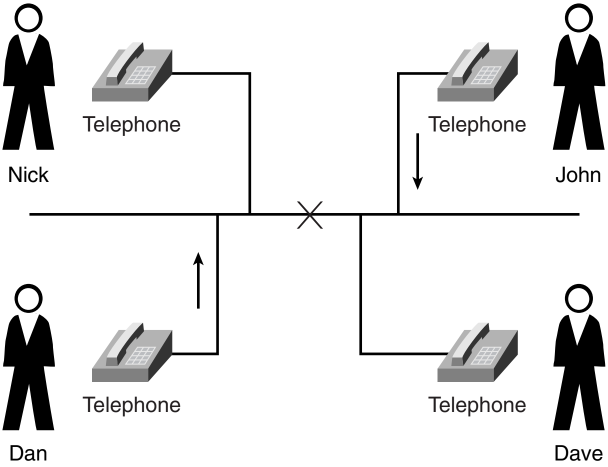


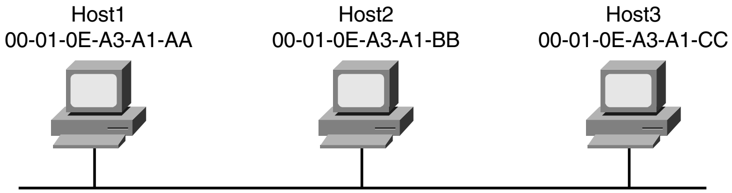
Table 1-2 lists each party line operation and compares it to Ethernet.

Table 1-2 *Comparing Party Line and Ethernet Operations*

Step	Telephone Party Line Operation	Ethernet Operation
1	I pick up the phone. Is anyone talking?	The LAN device listens to the Ethernet network to sense the carrier signal on the network.
2	If no one is speaking, I can start talking. I'll keep listening to make sure no one speaks at the same time as me.	If the LAN device does not detect a carrier signal on the network, it can begin transmitting. The LAN device listens to the carrier signal on the network and matches it to the output.
3	If I can't hear myself speak, I'll assume someone is trying to speak at the same time.	If there is a discrepancy between input and output, another LAN device has transmitted. This is a collision.
4	I'll then yell out to tell the other person to stop talking.	The LAN device sends out a jamming signal to alert the other LAN devices that there has been a collision.
5	I will then wait a random amount of time to start my conversation again.	The LAN device waits a random amount of time to start transmitting again. This is called the *backoff* algorithm. If multiple attempts to transmit fail, the backoff algorithm increases the amount of time waited.

In a party line, people occasionally speak over each other. When the party line is loaded with more callers, the more often people attempt to speak at the same time. It is the same with Ethernet collisions. Because users share Ethernet bandwidth and are part of the same collision domain, it is often referred to as *shared media* or *shared Ethernet.* (See Figure 1-4.) The efficiency of shared Ethernet is proportional to the number of devices attempting to communicate at the same time. As more devices are added, the efficiency decreases.

Figure 1-4 *Shared Ethernet Segment*



The algorithm in CSMA/CD used after a collision is Truncated Binary Exponential Backoff algorithm. When a collision occurs, the device must wait a random number of slot times before attempting to retransmit the packet. The slot time is contingent upon the speed of the link. For instance, slot time will be different for 10 Mbps Ethernet versus 100 Mbps Ethernet. Table 1-3 shows an example for a 10 Mbps Ethernet link. Cisco switches uses a more aggressive Max Wait Time than what is illustrated in this example. The purpose of the example is to give you a feel for how Truncated Binary Exponential Backoff works.

Table 1-3 *CSMA/CD Collision Backoff Ranges*

Retry	Range	Max Number	Max Wait Time
1st	0-1	(2^1)-1	51.2us
2nd	0-3	(2^2)-1	153.6us
3rd	0-7	(2^3)-1	358.4us
4th	0-15	(2^4)-1	768.0us
5th	0-31	(2^5)-1	1587.2us
6th	0-63	(2^6)-1	3225.6us
7th	0-127	(2^7)-1	6502.4us
8th	0-255	(2^8)-1	13056.0us
9th	0-511	(2^9)-1	26163.2us
10th – 15th	0-1023	(2^10)-1	52377.6us

Cisco switches monitor various collision counters, as follows:

- Single
- Multiple
- Late
- Excessive

Of the four types, be wary of late and excessive collisions. Late collisions occur when two devices send data at the same time. Unlike single and multiple collisions, late collisions cause packets to be lost. Late collisions are usually indicative of the cable exceeding IEEE specifications. Cascading hubs (connecting two or more hubs to each other) can also cause the length of the collision domain to increase above specification. You can use a Time Delay Reflectometer (TDR) to detect cable fault and whether the cable is within the IEEE standard. Other factors that cause late collisions include mismatched duplex settings and bad transceivers. Example 1-1 shows the output from a switch that has detected a late collision on one of its ports.

Example 1-1 *Late Collision Error Messages*

```
%LANCE-5-LATECOLL: Unit [DEC], late collision error
%PQUICC-5-LATECOLL: Unit [DEC], late collision error
```

The slot time, 51.2 microseconds, used to detect and report collisions is based on the round trip time between the furthest points on the Ethernet link. The value is calculated by taking the smallest Ethernet frame size of 64 bytes and multiplying it by 8 bits, which gives 512 bits. This number is then multiplied by .1 microseconds. The farthest distance between the end points of the cable should be reached within half of this slot time, 25.6 microseconds.

Excessive collisions typically occur when too much traffic is on the wire or too many devices are in the collision domain. After the fifteenth retransmission plus the original attempt, the excessive collisions counter increments, and the packet gets dropped. In this case, too many devices are competing for the wire. In addition, duplex mismatches can also cause the problem. A syslog message is generated by the switch, as depicted in Example 1-2, when excessive collision occurs on the port.

Example 1-2 *Excessive Collisions Error Message*

```
%PQUICC-5-COLL: Unit [DEC], excessive collisions. Retry limit [DEC] exceeded
```

On the switch, the **show port** *mod/port* command provides information about collisions, multiple collisions, and so on. Example 1-3 is an excerpt from the **show port** command that is useful. This example was taken from a switch that was running Catalyst OS software.

Example 1-3 *Sample of* ***show port*** *Command*

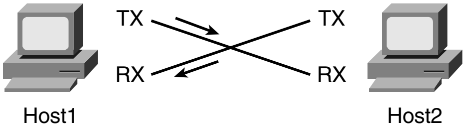
```
Switch1 (enable) show port 10/3

Port  Single-Col Multi-Coll Late-Coll  Excess-Col Carri-Sen Runts     Giants
----- ---------- ---------- ---------- ---------- --------- --------- ---------
10/3          37          3         24          0         0         0         0
```

Full-Duplex Ethernet

In the party line scenario, congestion occurs when more than two people attempt to talk at the same time. When only two people are talking, or only two devices, virtually all the bandwidth is available. In cases where only two devices need to communicate, Ethernet can be configured to operate in full-duplex mode as opposed to the normal half-duplex operation. Full-duplex operation allows a network device to "talk" or transmit and "listen" or receive at the same time. (See Figure 1-5.)

Figure 1-5 *Full-Duplex Directly Connected Hosts*



Because Ethernet is based on CSMA/CD, full-duplex devices either need to be directly connected to each other or be connected to a device that allows full-duplex operation (such as a LAN switch). Ethernet hubs do not allow full-duplex operation, as they are only physical layer (Layer 1) signal repeaters for the logical bus (Layer 2). Ethernet still operates as a logical bus under full duplex.

Autonegotiation

Autonegotiation is a mechanism that allows two devices at either end to negotiate speed and duplex settings at physical layer. The benefits of autonegotiation include minimal configuration and operability between dissimilar Ethernet technologies.

In today's networks, 10BASE-T and 100BASE-T are ubiquitous. Newer Cisco modules such as the WS-X6548-GE-TX have ports capable of 10/100/1000BASE-T. Most existing network interface cards (NICs) operate at 10/100 speeds, with newer NICs offering 10/100/1000BASE-T operation. NICs capable of autonegotiating speed and duplex are beneficial because more and more users are becoming mobile. One day, a user might be connected to the office Catalyst switch at 100 Mbps, and the next day, a remote site that supports only 10 Mbps. The primary objective is to ensure that the user not only has easy access to the network but also has network reliability. If the user's laptop NIC is hard coded at 100BASE-T full duplex, the user connectivity might be impacted because the two switches might have different types of modules that operate at different speeds. For instance, the module in the office building is WS-X5225 (24 port 10/100BASE-TX), and the remote site has WS-X5013 (24 port 10BASE-T). In this case, because the switches are set by default to autonegotiate, a user with a NIC hard coded to 100BASE-T full duplex will not get any connectivity. Setting up autonegotiation on both the switch and laptop gets rid of this problem. The user no longer has to worry about the laptop NIC settings because the NIC automatically negotiates the proper physical layer configuration with the end device to which it connects.

The actual mechanics behind autonegotiation are straightforward, as depicted in Figure 1-6. Autonegotiation attempts to match speed and duplex mode at the highest priority with its link partner. Since the introduction of 1000BASE-T, the priorities have been readjusted. Table 1-4 describes each priority level.

Figure 1-6 *Ethernet Autonegotiation*

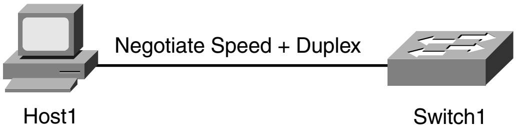


Table 1-4 *Autonegotiation Priority Levels*

Priority	Ethernet Specification	Type of Duplex
1	1000BASE-T	Full duplex
2	1000BASE-T	Half duplex
3	100BASE-T2	Full duplex
4	100BASE-TX	Full duplex
5	100BASE-T2	Half duplex
6	100BASE-T4	---
7	100BASE-TX	Half duplex
8	10BASE-T	Full duplex
9	10BASE-T	Half duplex

The 10BASE-T specification does not include autonegotiation between devices. Autonegotiation was first introduced in IEEE 802.3u Fast Ethernet specification as an optional parameter. In a 10BASE-T environment, a single pulse, called the Normal Link Pulse (NLP), is sent every 16 ms (±8 ms) on an idle link. The NLP performs a link integrity test for 10BASE-T. When no traffic is on the link, the 10BASE-T device generates a NLP on the wire to keep the link from going down. The 10BASE-T device stops generating pulses when it receives data packets. A link failure occurs under conditions when the 10BASE-T device does not receive NLPs or a single data packet within a specified time slot.

As mentioned earlier, the IEEE 802.3u specification has an optional programmable field for autonegotiation. Within autonegotiation, there are various other optional operations, such as Remote Fault Indication and Next Page Function. Remote Fault Indication detects and informs the link partner of physical layer errors. The Next Page Function provides more verbose information about the negotiation process. One of the more appealing features of autonegotiation is compatibility with dissimilar Ethernet technologies. For example, Fast Ethernet is backward-compatible with 10BASE-T through a Parallel Detection mechanism. Essentially, the Fast Ethernet switches to NLP to communicate with a 10BASE-T device. Parallel Detection is when only one of the two link partners is capable of autonegotiation.

Fast Ethernet uses the same pulse structure as 10BASE-T. In 10BASE-T, there is only a single pulse every 16 ms, whereas in Fast Ethernet, there are bursts of pulses in intervals of 16 (±8) ms. In these pulses, or groups of pulses, the capability of the device is encoded in a 16-bit word called a Link Code Word (LCW), also known as Fast Link Pulse (FLP). The length of the burst is approximately 2 ms.

NOTE Fast Ethernet vendors used their discretion whether to add autonegotiation capabilities to their devices. As a result, Fast Ethernet NICs without autonegotiation capabilities were once found in the marketplace.

Gigabit Ethernet implementation requires that all IEEE 802.3z compliant devices have autonegotiation capability. Autonegotiation can, however, be disabled through a software feature. From the actual hardware perspective, the 802.3z specification requires autonegotiation capabilities on the device. On Cisco Catalyst switches, autonegotiation can be disabled with the following command. Note that this command must be configured on both link partners:

```
set port negotiation <mod/port> enable | disable
```

The parameters that 802.3z devices negotiate are

- Duplex setting
- Flow control
- Remote fault information

Although duplex setting can be negotiated, Cisco switches operate Gigabit Ethernet in full-duplex mode only. With the introduction of the newer 1000/100/10 blades, a port can operate at various speeds and duplex settings. However, it is unlikely that Cisco will support Gigabit half duplex in any point-to-point configurations with even the aforementioned blades. Use the **show port capabilities** command that is available in Catalyst OS to view the features supported by the line module, as shown in Example 1-4.

Example 1-4 *Output from* **show port capabilities** *Command*

```
Switch1 (enable) show port capabilities 1/1
Model                   WS-X6K-SUP2-2GE
Port                    1/1
Type                    1000BaseSX
Speed                   1000
Duplex                  full
```

Flow control is an optional feature that is part of the 802.3x specification. The concept behind flow control is to help reduce the burden on the port that is overwhelmed with traffic. It does this by creating back-pressure on the network. If the volume of traffic is such that a port runs out of buffers, it drops subsequent packets. The flow control mechanism simply tells the transmitter to back off for a period of time by sending an Ethernet Pause Frame (MAC address of 01-80-c2-00-00-01) to the transmitter. The transmitter receives this frame and buffers the outgoing packets in its output buffer queue. This mechanism provides needed time for the receiver to clear the packets that are in its input queue. The obvious advantage is that packets are not dropped. The negative aspect to this process is latency. Certain multicast, voice, and video traffic are sensitive to latency on the network. It is recommended that flow control should be implemented with care. Typically, this feature is implemented as a quick fix. Not all Cisco switches support this feature.

```
set port flowcontrol <mod/port>
```

Remote fault information detects and advertises physical layer problems such as excessive noise, wrong cable types, bad hardware, and so on to the remote peer. The switch is programmed to take a proactive approach when excessive physical layer problems exist. A port that is generating errors can potentially disrupt a network. For instance, it can cause spanning-tree problems and traffic black holing, and drain system resources. As a result, the switch error disables the port.

Looking at some examples will solidify the concept and function of autonegotiation. In Figure 1-7, Host1 and the hub are link partners over a 10BASE-T connection. 10BASE-T has no knowledge of autonegotiation, and therefore, the devices must statically be configured. NLPs are sent by both devices when they come online. In this example, these devices operate over a 10BASE-T half-duplex connection.

Figure 1-7 *10BASE-T Autonegotiation*

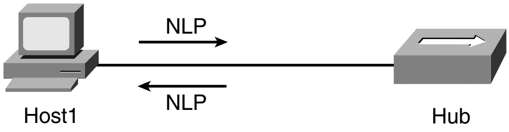
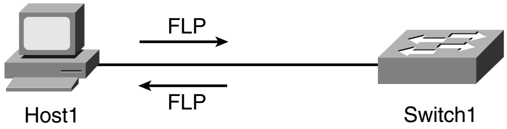


Figure 1-8 shows a straight 100BASE-T connection with both devices enabled for autonegotiation. FLP bursts are sent to advertise the device's capabilities and negotiate a maximum highest bandwidth connection. The highest connection negotiated is priority 4, which is 100BASE-TX full duplex.

Figure 1-8 *100BASE-T Autonegotiation*

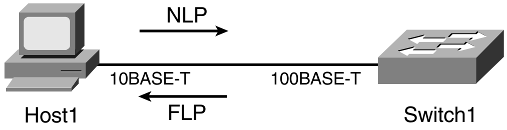


The following is the command that configures a switch to autonegotiate a port:

```
set port speed <mod/port> auto
```

In Figure 1-9, Host1 has a 10BASE-T card. The switch has a capability to operate in both 10BASE-T and 100BASE-T mode. The 10/100 modules are common in a switching environment. Cisco has various 10/100 modules with various features and functionalities. In this example, there is a mismatch between the pulses sent by the Host1 and the switch. Because Host1 has a 10BASE-T card, it can send only NLPs. Initially, when the switch comes online, it generates only FLP bursts. When the switch detects NLPs from its link partner, it ceases to generate FLP bursts and switches to NLP. Depending on the static configuration on Host1, the switch chooses that priority. In this instance, the connection is 10BASE-T operating at half duplex.

Figure 1-9 *10/100BASE-T Autonegotiation*



The finer points of autonegotiation have been discussed; however, some drawbacks need to be discussed. Numerous network problems resulted when the autonegotiation feature was first deployed. The issues ranged from degradation in performance to connectivity loss. The cause of some of these problems included advanced software features that came with the NIC, vendors not fully conforming to 802.3u standard, and buggy code. These days, now that manufacturers have resolved these issues, misconfiguration is the biggest remaining problem. Table 1-5 and Table 1-6 show various consequences from misconfigurations. For instance, a duplex mismatch can degrade performance on the wire and potentially cause packet loss.

Network engineers still have heated discussions about whether to enable autonegotiation in the network. As mentioned earlier, autonegotiation is a big advantage for mobile users. A user should not have to worry about configuring his laptop every time he goes to a different location.

Table 1-5 *Autonegotiation Configurations for 10/100 Ethernet*

Configuration NIC (Speed/ Duplex)	Configuration Switch (Speed/ Duplex)	Resulting NIC Speed/ Duplex	Resulting Catalyst Speed/Duplex	Comments
AUTO	AUTO	100 Mbps, Full duplex	100 Mbps, Full duplex	Assuming maximum capability of Catalyst switch and NIC is 100 full duplex.
100 Mbps, Full duplex	AUTO	100 Mbps, Full duplex	100 Mbps, Half duplex	Duplex mismatch.
AUTO	100 Mbps, Full duplex	100 Mbps, Half duplex	100 Mbps, Full duplex	Duplex mismatch.
100 Mbps, Full duplex	100 Mbps, Full duplex	100 Mbps, Full duplex	100 Mbps, Full duplex	Correct manual configuration.
100 Mbps, Half duplex	AUTO	100 Mbps, Half duplex	100 Mbps, Half duplex	Link is established, but switch does not see any autonegotiation information from NIC and defaults to half duplex.
10 Mbps, Half duplex	AUTO	10 Mbps, Half duplex	10 Mbps, Half duplex	Link is established, but switch will not see FLP and will default to 10 Mbps half duplex.
10 Mbps, Half duplex	100 Mbps, Half duplex	No Link	No Link	Neither side will establish link because of speed mismatch.
AUTO	100 Mbps, Half duplex	10 Mbps, Half duplex	10 Mbps, Half duplex	Link is established, but NIC will not see FLP and default to 10 Mbps half duplex.

Table 1-6 *Autonegotiations Configurations for Gigabit Ethernet*

Switch Port Gigabit	Autonegotiation Setting	NIC Gigabit Autonegotiation Setting	Switch Link/NIC Link
Enabled	Enabled	Up	Up
Disabled	Disabled	Up	Up
Enabled	Disabled	Down	Up
Disabled	Enabled	Up	Down

The rule of thumb is to enable autonegotiation on access ports that connect to users. Mission-critical devices should be statically configured to protect the network from possible outages and performance hits. Therefore, connections between routers and switches, or servers and switches should be hard coded with the appropriate speed and duplex settings.

Transparent Bridging

The inability to allow more than one device to transmit simultaneously presents a major challenge when attempting to connect dozens or hundreds of users together through Ethernet.

Transparent bridging is the augmentation of Ethernet allowing partial segmentation of the network into two or more collision domains. The IEEE-defined transparent bridging is an industry standard in 802.1D. Transparent bridges improve network performance by allowing devices in the same segmented collision domain to communicate without that traffic unnecessarily being forwarded to the other collision domain.

Transparent bridges are the predominant bridge type for Ethernet, and it is important to understand Ethernet switches essentially act as multiport transparent bridges.

Figure 1-10 shows a transparent bridge supporting Ethernet segments or collision domains. If Host1 and Host2 are talking to each other, their conversation will use bandwidth only on their side of the bridge. This allows Host4 and Host5 to also hold a conversation. If all devices were in the same collision domain, only one conversation would be possible.

However, if Host1 wants to talk to Host4, as shown in Figure 1-11, the bandwidth will be utilized on both sides of the bridge, allowing only the one conversation.

How does the transparent bridge determine which users are connected to which side of the bridge? Well, transparent bridging has a little more "under the hood" than the example illustrates. The 802.1D specification for transparent bridging defines five unique processes as part of transparent bridging:

- Learning
- Flooding
- Filtering
- Forwarding
- Aging