

# Networking Course. Lesson 2. Layer 2 and Switching



Document version 0.1

This document was created, but not reviewed yet.  
You should use it very carefully.

- LAN Switching Foundation Technologies
  - OSI Model
  - Introducing Ethernet
  - Transparent Bridging
  - Broadcasts and Multicasts
  - Introducing Spanning Tree Protocol
- LAN Switch Architecture
  - Receiving Data-Switching Modes
  - Switching Data
  - Buffering Data
  - Oversubscribing the Switch Fabric
  - Congestion and Head-of-Line Blocking
  - Forwarding Data
- Layer 2 Fundamentals
  - Understanding Legacy LAN Segment
  - Introducing Virtual LANs
  - Trunking Methods
  - VLAN Trunking Protocol
  - Configuring VTP/VLAN/Trunk
  - VLAN Pruning
  - EtherChannel
  - Understanding VLAN 1
  - Private VLANs
- Spanning Tree Protocol
  - Root Bridge or Switch Port
  - Spanning Tree Protocol Configuration
- Virtual LANs
  - VLAN Overview
  - VLAN Topology
  - VLAN Operation
  - VLAN Trunking Protocol (VTP)
- Switching Security
  - Network Security Basic Rules
  - Port Security
  - Virtual LANs
  - VLAN-Based Network Attacks
  - Chapter Summary
- Configuring Switches
- Implementing and Tuning Spanning Tree
- Troubleshooting the LAN Switching Configuration

## Theory

- LAN Switching Foundation Technologies
  - OSI Model
  - Introducing Ethernet
  - Transparent Bridging
  - Broadcasts and Multicasts
  - Introducing Spanning Tree Protocol
- LAN Switch Architecture
  - Receiving Data-Switching Modes
  - Switching Data
  - Buffering Data
  - Oversubscribing the Switch Fabric
  - Congestion and Head-of-Line Blocking
  - Forwarding Data
- Layer 2 Fundamentals
  - Understanding Legacy LAN Segment

- Introducing Virtual LANs
- Trunking Methods
- VLAN Trunking Protocol
- Configuring VTP/VLAN/Trunk
- VLAN Pruning
- EtherChannel
- Understanding VLAN 1
- Private VLANs
- Spanning Tree Protocol
  - Root Bridge or Switch Port
  - Spanning Tree Protocol Configuration
- Virtual LANs
  - VLAN Overview
  - VLAN Topology
  - VLAN Operation
  - VLAN Trunking Protocol (VTP)
- Switching Security
  - Network Security Basic Rules
  - Port Security
  - Virtual LANs
  - VLAN-Based Network Attacks
  - Chapter Summary

## LAN Switching Foundation Technologies

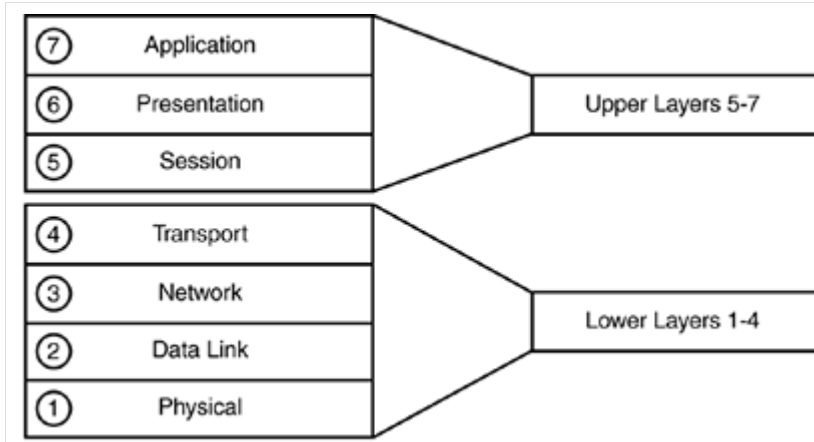
- OSI Model
  - OSI Upper Layers
  - OSI Lower Layers
- Introducing Ethernet
  - Types of Ethernet
  - Transmission Media
    - Ethernet over Twisted-Pair Cabling
    - Ethernet over Fiber Optics
    - Ethernet over Coax Cabling
  - Ethernet Cross-Over Cabling
  - Ethernet Topology
  - Ethernet Logical Addressing
  - CSMA/CD Operation
  - Full-Duplex Ethernet
  - Autonegotiation
- Transparent Bridging
  - Learning
  - Flooding
  - Filtering
  - Forwarding
  - Aging
- Broadcasts and Multicasts
- Introducing Spanning Tree Protocol
  - Spanning Tree Operations
  - Spanning Tree Port Transitions and Timers
  - Topology Changes in STP

## 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

## Ethernet over Coax Cabling

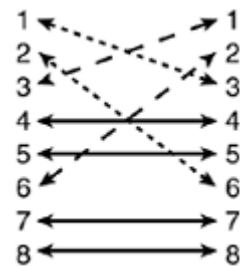
The use of coax cable for LANs is virtually nonexistent. One might run into it in an old abandoned building. Ethernet's eventual support of twisted pair cabling in a star topology virtually ended the use of coaxial cabling for Ethernet. Keep in mind that coax cable was not cheap either. Two major types of coax were used: thinnet (also called cheapernet) and thicknet. Thinnet uses 50 ohm coax 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.

Figure 1-2. Crossover Pinouts



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.

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](#).

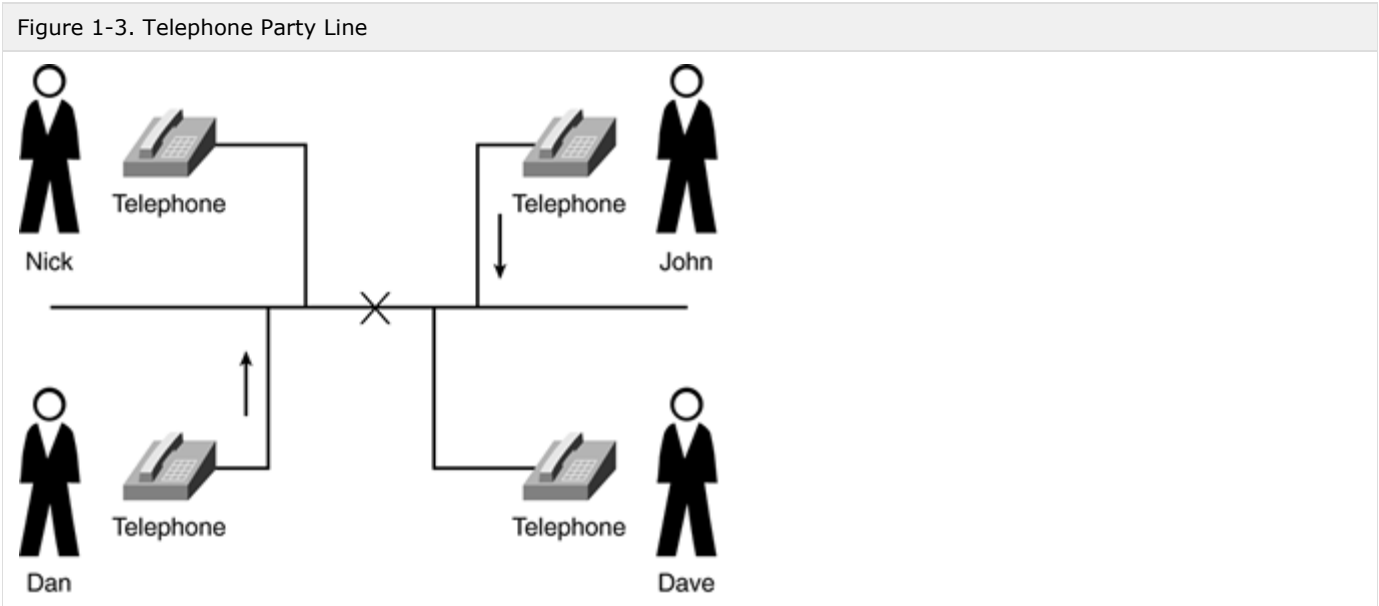


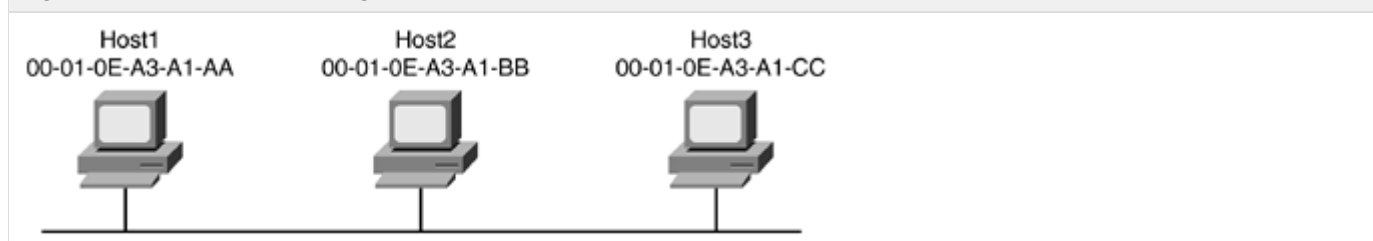
Table 1-2 lists each party line operation and compares it to Ethernet.

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
1 <sup>st</sup>	0-1	$(2^1)-1$	51.2us
2 <sup>nd</sup>	0-3	$(2^2)-1$	153.6us
3 <sup>rd</sup>	0-7	$(2^3)-1$	358.4us
4 <sup>th</sup>	0-15	$(2^4)-1$	768.0us
5 <sup>th</sup>	0-31	$(2^5)-1$	1587.2us
6 <sup>th</sup>	0-63	$(2^6)-1$	3225.6us
7 <sup>th</sup>	0-127	$(2^7)-1$	6502.4us
8 <sup>th</sup>	0-255	$(2^8)-1$	13056.0us
9 <sup>th</sup>	0-511	$(2^9)-1$	26163.2us

10 <sup>th</sup> - 15 <sup>th</sup>	0-1023	(2 <sup>10</sup> )-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
```

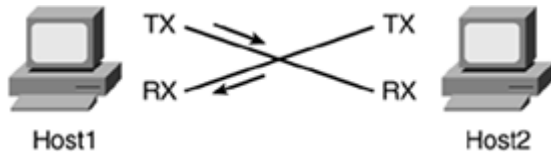
Port	Single-Col	Multi-Coll	Late-Coll	Excess-Col	Carri-Sen	Runts	Giants
10/3	37	3	24	0	0	0	0



## 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

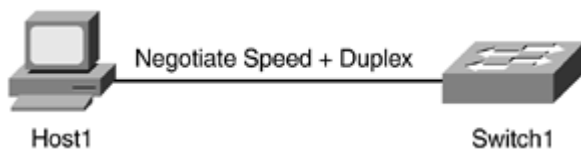


Table 1-4. Autonegotiation Priority Levels

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 ( $\pm 8$  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 ( $\pm 8$ ) 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

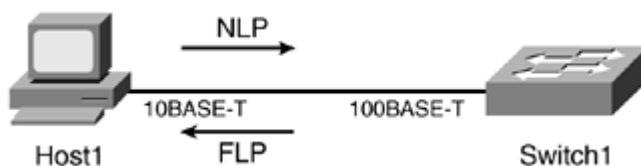


The following is the command that configures a switch to autonegotiate a port:

```
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.

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

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.

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.

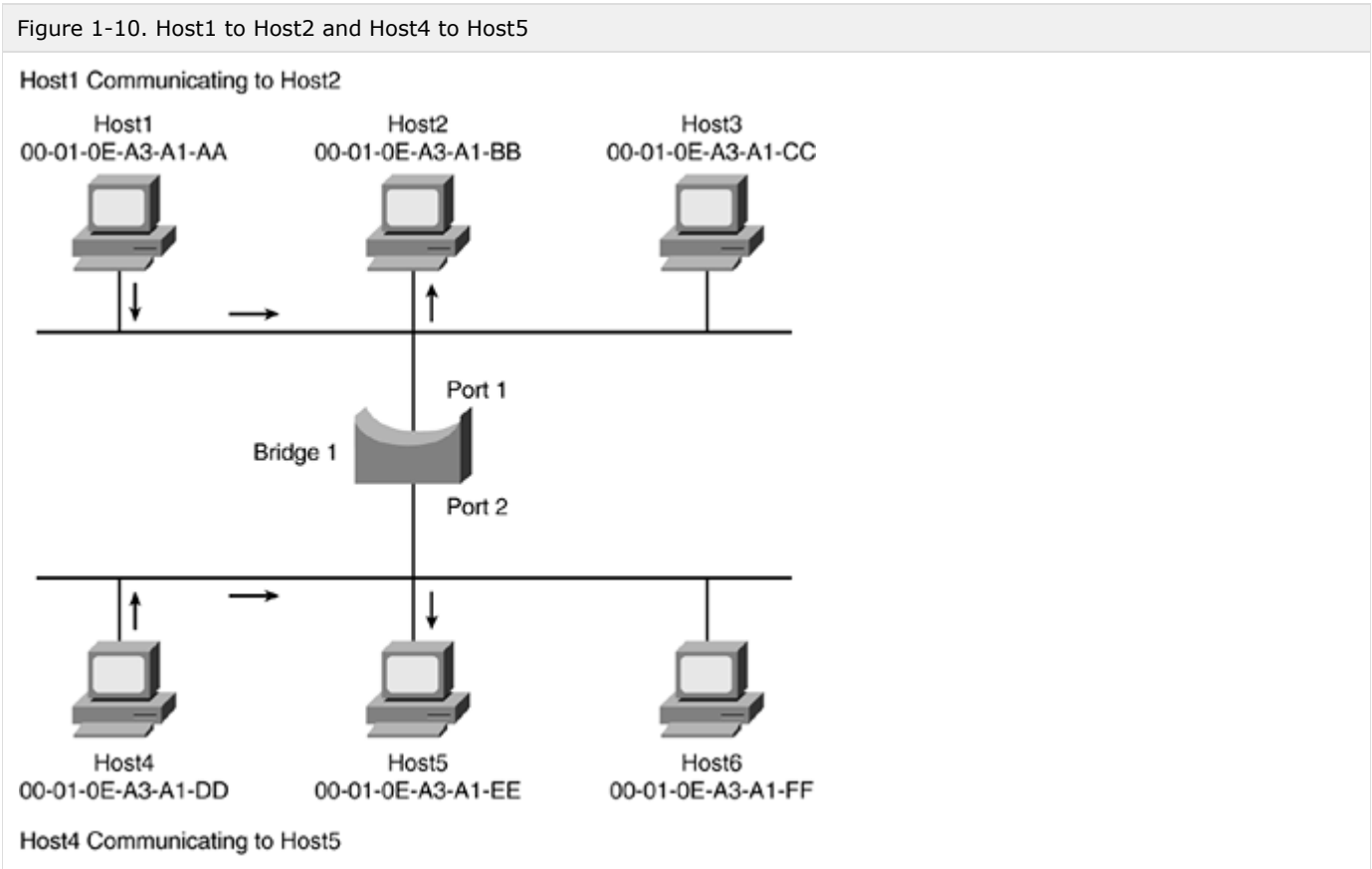
## Transparent Bridging

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.

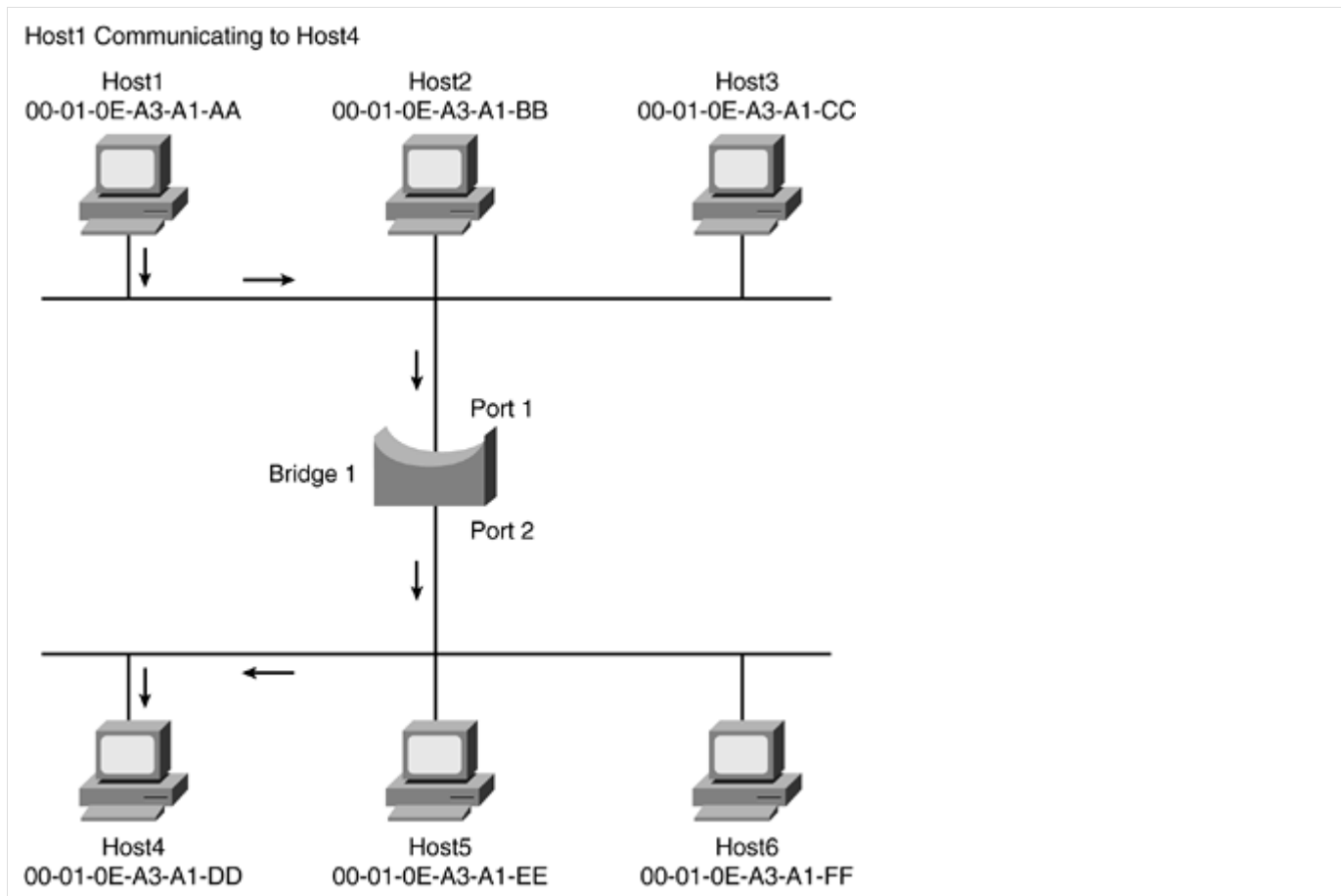
Transparent bridges are the predominant bridge type for Ethernet, and it is important to understand Ethernet switches essentially act as multiport transparent bridges.

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.

Figure 1-11. Host1 to Host4



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

The following sections describe each of these processes in more detail.

## Learning

Learning is the process of obtaining the MAC address of devices. When a bridge is first turned on, it has no entries in its bridge table. As traffic passes through the bridge, the sender's MAC addresses are stored in a table along with the associated port on which the traffic was received. This table is often called a bridge table, MAC table, or content addressable memory (CAM) table.

Table 1-7 shows a listing of all the devices on the sample network in Figure 1-10 and Figure 1-11.

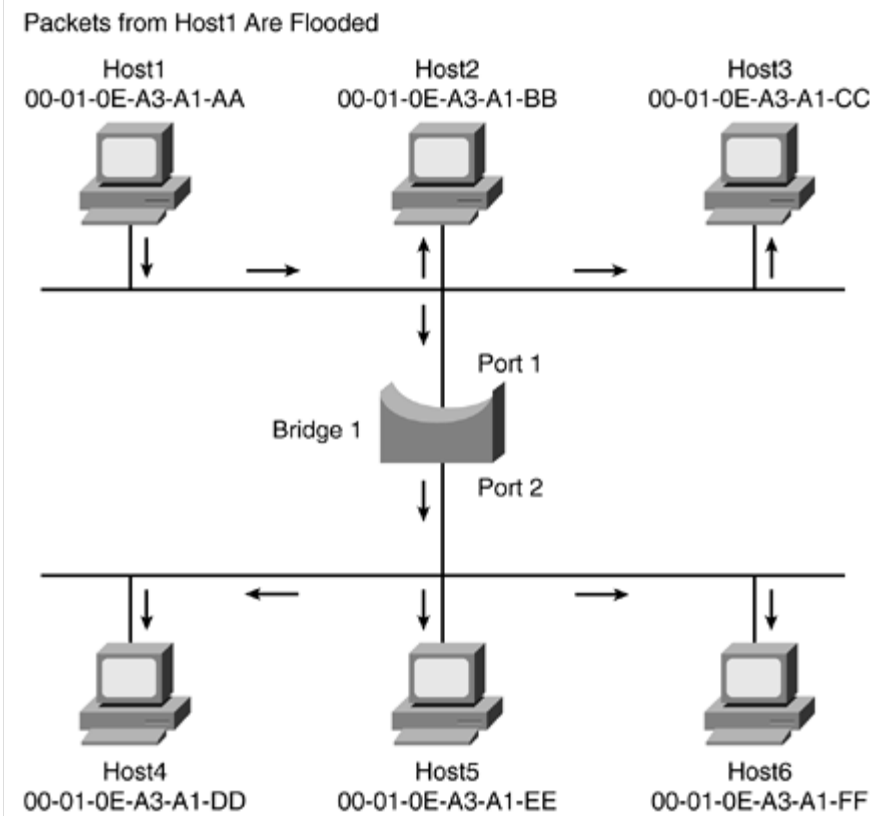
Table 1-7. Sample Bridge Table

Hosts	Port 1	Port 2
Host1/ 00-01-0E-A3-A1-AA	X	
Host2/ 00-01-0E-A3-A1-BB	X	
Host3/ 00-01-0E-A3-A1-CC	X	
Host4/ 00-01-0E-A3-A1-DD		X
Host5/ 00-01-0E-A3-A1-EE		X
Host6/ 00-01-0E-A3-A1-FF		X

## Flooding

When a bridge does not have an entry in its bridge table for a specific address, it must transparently pass the traffic through all its ports except the source port. This is known as flooding. The source port is not "flooded" because the original traffic came in on this port and already exists on that segment. Flooding allows the bridge to learn, as well as stay transparent to the rest of the network, because no traffic is lost while the bridge is learning. Figure 1-12 shows how the bridge forwards the traffic on all its ports.

Figure 1-12. Bridge1 Floods Traffic



## Filtering

After the bridge learns the MAC address and associated port of the devices to which it is connected, the benefits of transparent bridging can be seen by way of filtering. Filtering occurs when the source and destination are on the same side (same bridge port) of the bridge. In Figure 1-10, filtering occurs each time Host1 and Host2 talk, as well as when Host4 and Host5 talk.

## Forwarding

Forwarding is simply passing traffic from a known device located on one bridge port to another known device located on a different bridge port. Again, referring back to Figure 1-11, after the initial devices were learned, forwarding occurs when Host1 and Host4 talk.

## Aging

In addition to the MAC address and the associated port, a bridge also records the time that the device was learned. Aging of learned MAC addresses allows the bridge to adapt to moves, adds, and changes of devices to the network. After a device is learned, the bridge starts an aging timer. Each time the bridge forwards or filters a frame from a device, it restarts that device's timer. If the bridge doesn't hear from a device in a preset period of time, the aging timer expires and the bridge removes the device from its table.

Aging ensures that the bridge tracks only active systems, and ensures that the MAC address table does not consume too much system memory.

## Broadcasts and Multicasts

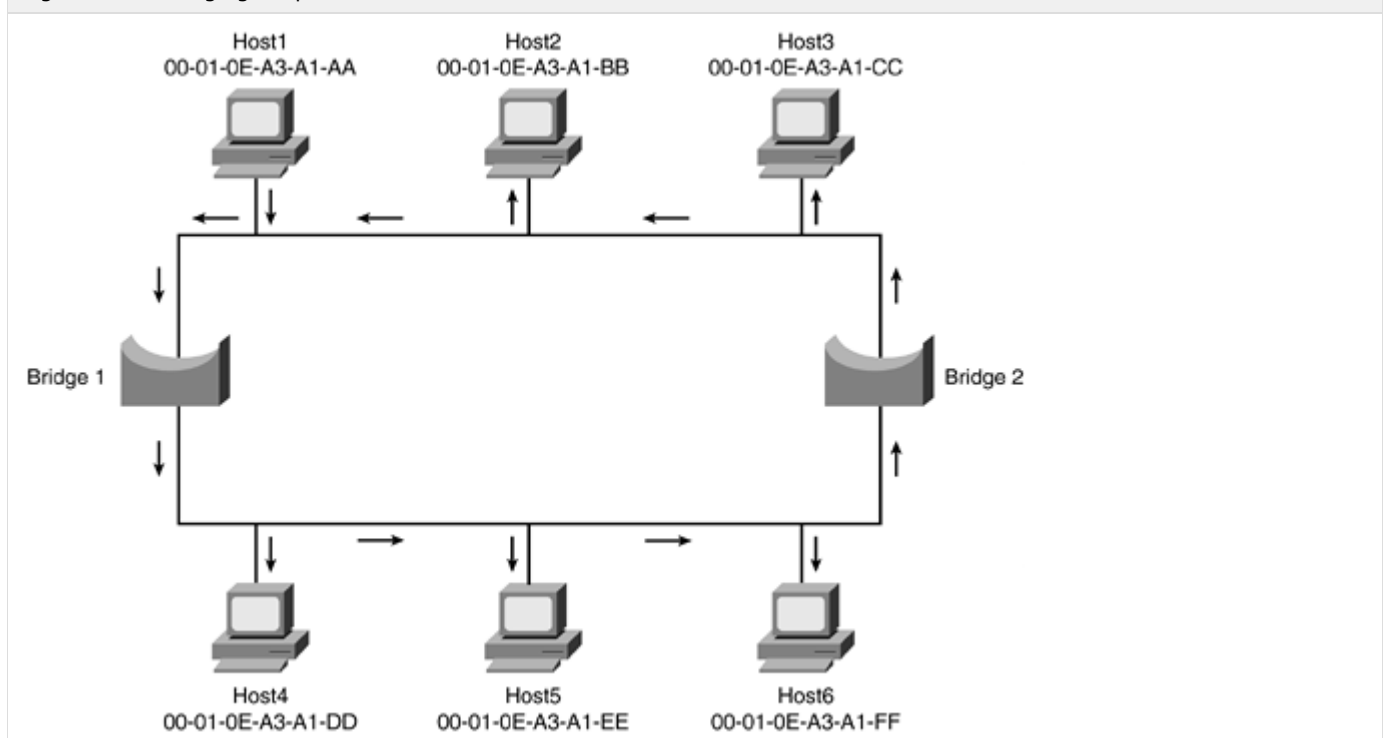
With Ethernet, broadcasts are specialized frames that are destined for all devices on an Ethernet network. Broadcasts use a MAC address of FF-FF-FF-FF-FF-FF. This is a special MAC address because it is the highest number allowed in the 48-bit schema of MAC addresses. In binary, all 48 bits are set to 1.

Multicasts are specialized broadcasts. Multicasts are used by higher layer protocols to direct traffic to more than one select destination, rather than a broadcast, which is sent to all destinations. Application layer multicasts start with 01-00-5E prefix. The rest of the digits are assigned by the application layer protocol handling the multicast. However, other Layer 2-only multicast addresses do not have the prefix of 01-00-5E; for example, STP with MAC address of 01-00-0c-cc-cc-cd. For the most part, Ethernet networks treat multicasts like broadcasts by default. Several higher layer protocols, such as IGMP (Internet Group Messaging Protocol), can be used by switches to differentiate the traffic and forward only multicast out specific ports.

## Introducing Spanning Tree Protocol

As with traditional shared Ethernet, transparent bridges inherently lack the capability to provide redundancy. The Spanning Tree Protocol (STP) inserts a mechanism into the Ethernet transparent bridge environment to dynamically discover the network topology and ensure only one path through the network. Without STP, there is no way to make a transparent bridge environment redundant. STP also protects a network against accidental miscablings because it prevents unwanted bridging loops in the transparent bridging environment. A bridge loop is similar to a wrestling match. At first, everything appears orderly, but pandemonium soon ensues. The normal referee and rules do not work. The pandemonium does not stop until someone comes in and shuts the match down. Bridge loops in Ethernet and transparent bridging also cause pandemonium. [Figure 1-13](#) shows a bridge loop in Ethernet.

Figure 1-13. Bridging Loop

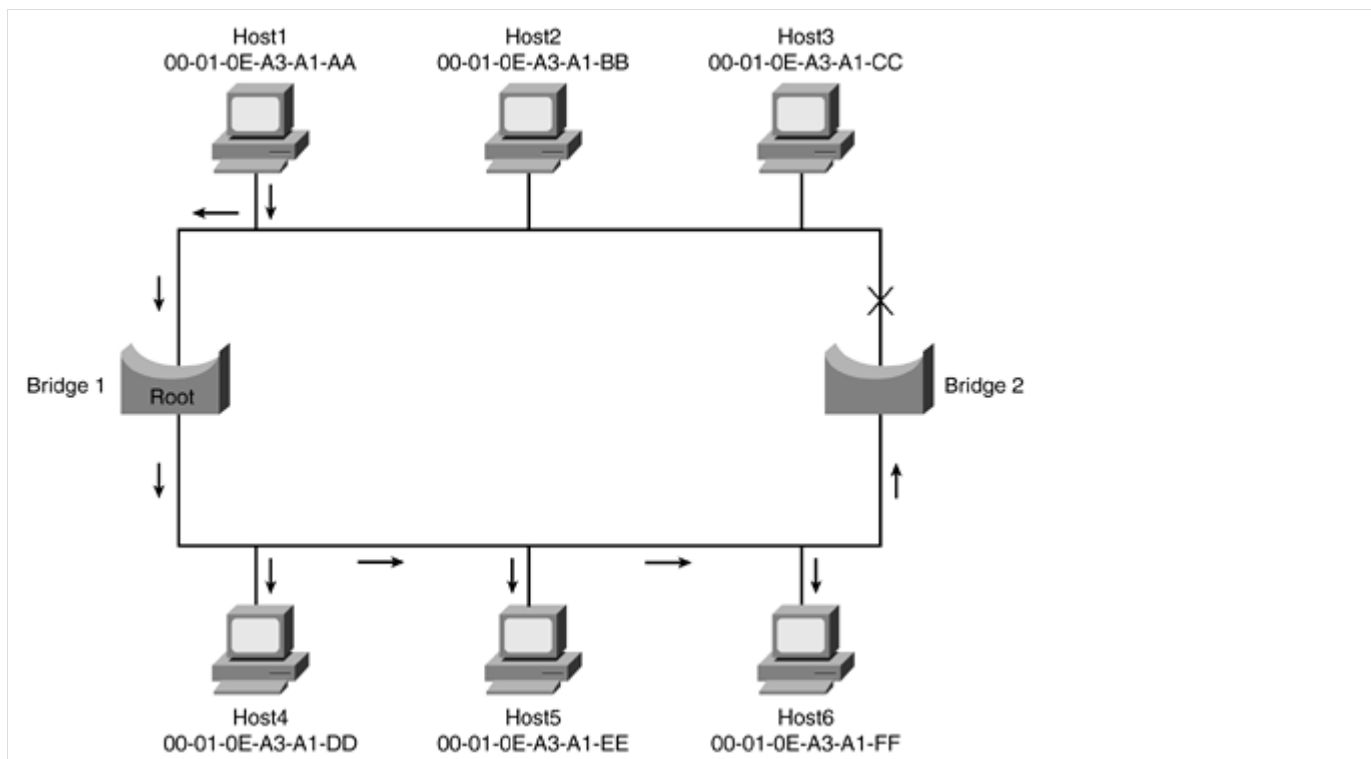


In this basic example, there are redundant links without STP. This creates a bridged loop. In this case, the redundant links cause the Ethernet data frame to have more than one path. Because the bridges are transparent, a copy of the data frame is sent across both paths. Bridge1 and Bridge2 both receive a copy of the data frame that was sent by the other. Then, each bridge sees alternating data frames, assumes that the source host is on the wrong side of the bridge, and updates the bridge table. The data frames then start to be recopied on each side of the bridge again and again. Think about how many data frames are needed for a simple e-mail message. With this bridging loop, the frames would be copied over and over again until they timed out. However, because the upper layer protocols are generating many requests, the process keeps happening. The entire network gets overwhelmed and legitimate traffic cannot pass.

[Figure 1-14](#) revisits the example again, this time adding STP, which blocks one of the redundant links, eliminating the bridging loop. If the first link or Bridge1 were to fail, STP would re-examine the network and enable the shutdown connection. This is how STP provides redundancy in a transparent bridging environment.

Figure 1-14. Spanning Tree Blocking





NOTE: Be aware that the spanning-tree algorithm is implemented in other media types such as Token Ring. STP has a different purpose and function in Token Ring than in Ethernet because bridging loops can be desirable in Token Ring.

## Spanning Tree Operations

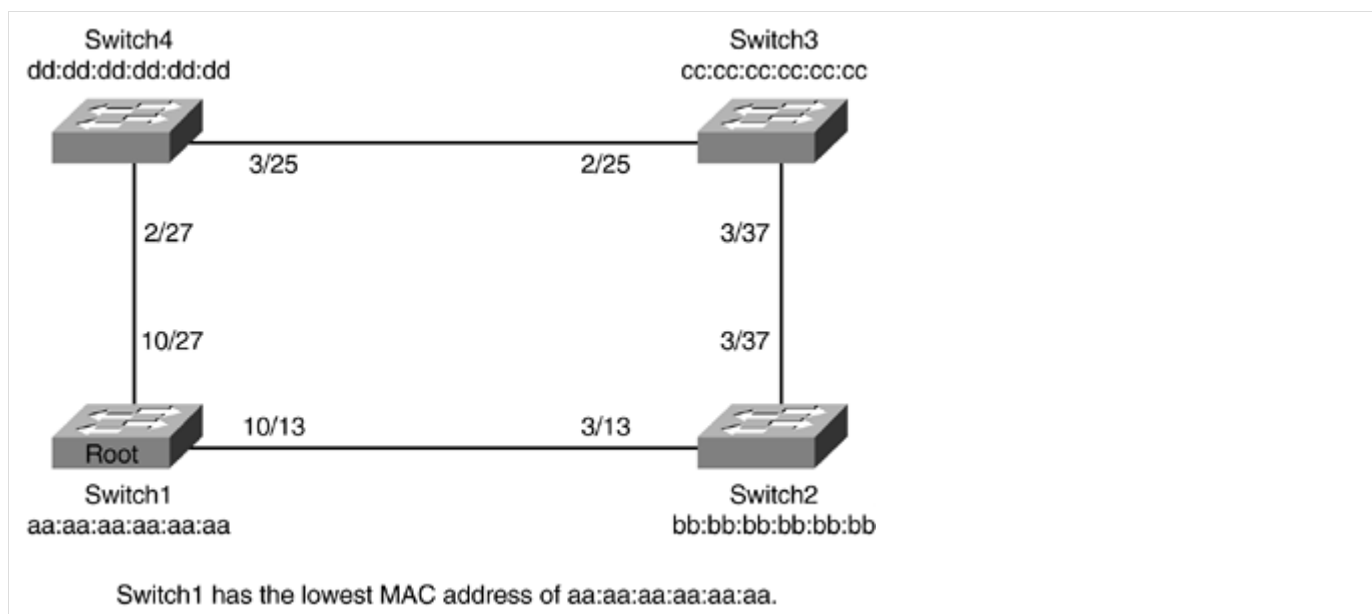
STP operation for each bridge can be broken down into three main steps:

- Root bridge selection
- Calculation of the shortest path to the root bridge
- Type of role an active port plays in STP

The main information to be concerned with is the Root ID (bridge that the transmitting bridge thinks is the root), Bridge ID, and cost (which is the cost to the root bridge). The STP topology is considered converged after a root bridge has been selected and each bridge has selected its root port, designated bridge, and which ports will participate in the STP topology. STP uses these configuration messages (BPDUs) as it transitions port states to achieve convergence.

Spanning tree elects one bridge on the LAN to be the master bridge. This bridge is called the root bridge. The root bridge is special because all the path calculation through the network is based on the root. The bridge is elected based on the Bridge ID (BID), which is comprised of a 2-byte Priority field plus a 6-byte MAC address. In spanning tree, lower BID values are preferred. In a default configuration, the Priority field is set at 32768. Because the default Priority field is the same for all the bridges, the root selection is based on the lowest MAC address. One method of selecting a specific bridge to be the root is to manually alter the Priority field to a lower value. Regardless of what the MAC address is, the Priority field decides what bridge is going to be the root, assuming that all bridges do not have the same priority value. For the remainder of this chapter, the figures depict a switch, which at its fundamental level is a glorified bridge. (See [Figure 1-15.](#))

Figure 1-15. Switch1 Becomes Root



Each bridge calculates all the paths from itself to the root. It then selects the shortest path. The next-hop bridge toward the root is the designated bridge. The port that leads to the designated bridge is selected to be the root port because it is closest from this bridge to the root bridge. The metric that STP uses for this determination is cost, which is based on the interface speed. [Table 1-8](#) compares bandwidth to STP interface costs.

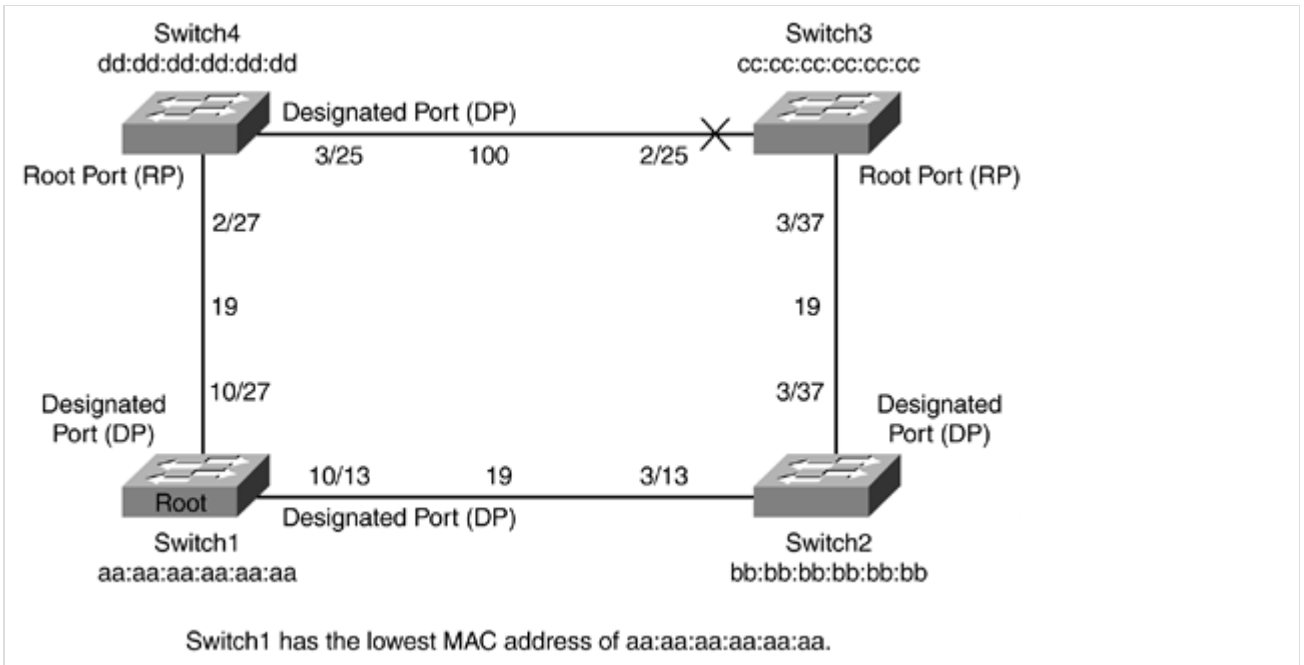
Table 1-8. Spanning Tree Interface Costs

Bandwidth	Cost
4 Mbps	250
10 Mbps	100
16 Mbps	62
45 Mbps	39
100 Mbps	19
155 Mbps	14
622 Mbps	6
1 Gbps	4
10 Gbps	2

As shown in [Figure 1-16](#), Switch3 has two paths to the root. To prevent a loop on the network, it must decide to block one of its ports. The algorithm used to make the decision is based on three choices:

- Lowest path cost to the root
- Lowest sender BID
- Lowest port ID

Figure 1-16. Converged Spanning-Tree Topology



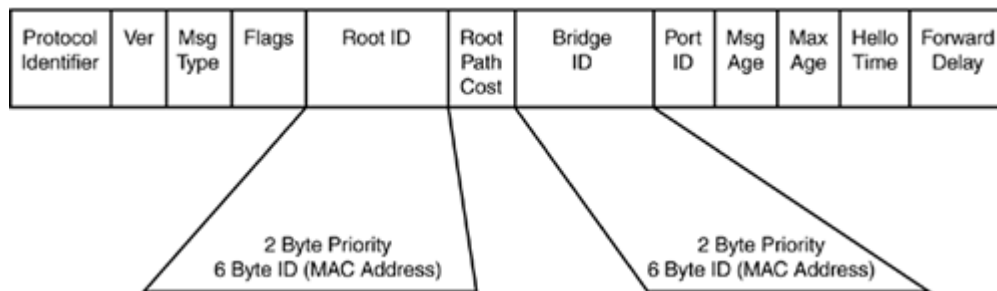
In this example, the lowest path cost to the root will decide which port will be forwarding and which one will be blocking. Because the cost is less through Switch2 path, 38, Switch3 will be forwarding out of this port and blocking on the other. This behavior of blocking a port allows the spanning tree to be loop free and provide redundancy should one of the ports go down.

Each active port can have a specific role to play in the spanning-tree algorithm:

- Designated Port (DP) - The port responsible for sending BPDUs on the segment
- Non-Designated Port (N-DP) - Does not send BPDUs on the segment
- Root Port (RP) - The closest port to the root

STP sends configuration messages out every port of the bridge. These messages are called bridge protocol data units (BPDUs). BPDUs contain the appropriate information for STP configuration. The Type field for BPDU message is 0x00, and it uses the multicast MAC address 01-80-C2-00-00-00. The BPDU packet is shown in Figure 1-17.

Figure 1-17. Bridge Protocol Data Unit Format



[View full size image](#)

## Spanning Tree Port Transitions and Timers

Part of the STP algorithm and process of building a loop-free network, as well as reconfiguration on a topology change, is to cycle the bridge ports through several states, as follows:

- Blocking - A port is placed in blocking mode upon startup and when STP determines it is a suboptimal path to the root bridge. Blocked ports do not forward traffic.
- Listening - When a port is transitioned from blocking to listening, it starts to listen for other bridges. It does not send out configuration messages, learn MAC addresses, or forward traffic.
- Learning - The bridge continues to listen for other bridges; however, it can now also learn MAC addresses of network devices.
- Forwarding - This is normal operation. Data and configuration messages are passed through the port.

STP uses timers to determine how long to transition ports. STP also uses timers to determine the health of neighbor bridges and how long to cache MAC addresses in the bridge table.

The explanation of the timers is as follows:

- Hello timer - 2 seconds. This timer is used to determine how often root bridge sends configuration BPDUs.
- Maximum Age (Max Age) - 20 seconds. This timer tells the bridge how long to keep ports in the blocking state before listening.
- Forward Delay (Fwd Delay) - 15 seconds. This timer determines how long to stay in the listening state before learning, and the learning state before forwarding.

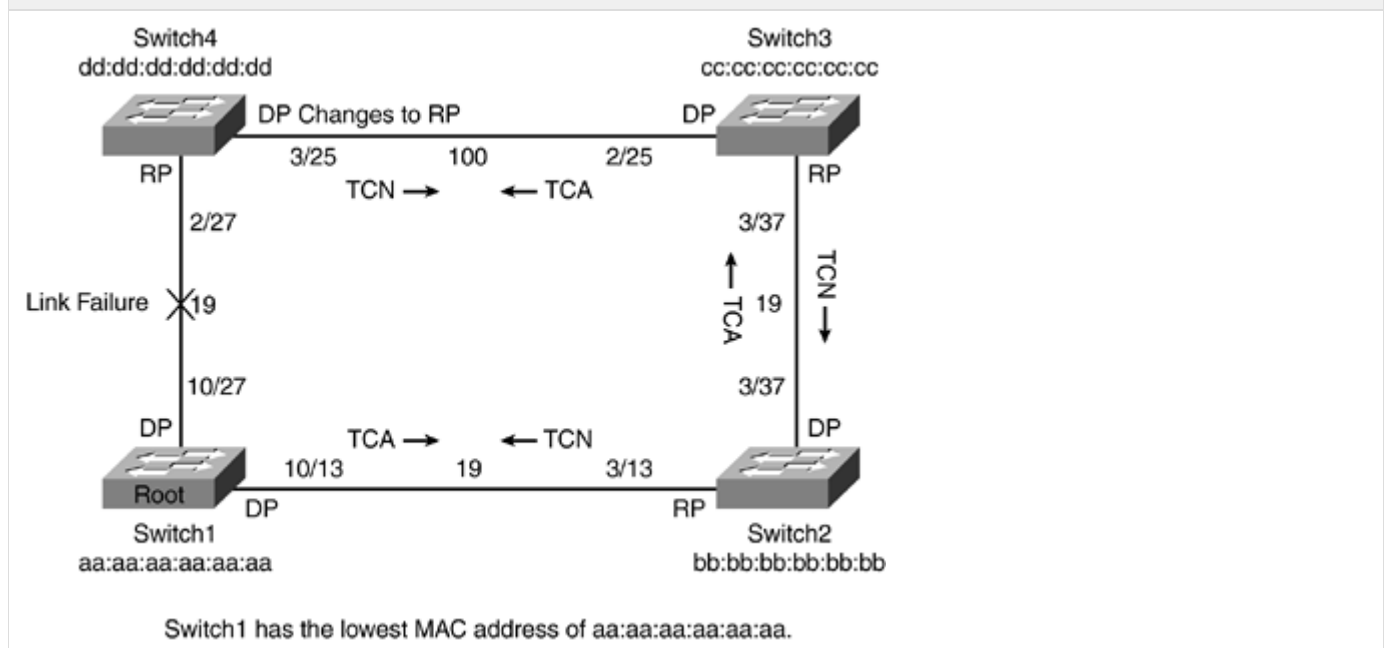
The STP timers can be tuned based on network size. These parameters are designed to give STP ample opportunity to ensure a loop-free topology. Mistuning these parameters can cause serious network instability. Tuning these parameters will be discussed in [Chapter 10](#), "Implementing and Tuning Spanning Tree." When a bridge sees BPDUs with a better path to the root, it recalculates STP. This allows ports to transition when appropriate.

## Topology Changes in STP

The other type of STP BPDU that needs to be discussed is Topology Change Notification (TCN). TCNBPDUs are generated when a bridge discovers a change in topology, usually because of a link failure, bridge failure, or a port transitioning to forwarding state. The TCN BPDU is set to 0x80 in the Type field and is subsequently forwarded on the root port toward the root bridge. The upstream bridge responds back with acknowledgment of the BPDU in the form of Topology Change Acknowledgment (TCA). The least significant bit is for TCN, and the most significant bit is for TCA in the Flag field.

[Figure 1-18](#) shows the flow of topology change BPDUs. The bridge sends this message to its designated bridge. Remember, the designated bridge is a particular bridge's closest neighbor to the root (or the root, if it is directly connected). The designated bridge acknowledges the topology change back to the sending neighbor and sends the message to its designated bridge. This process repeats until the root bridge gets the message. The root learns about the topology changes in the network in this way.

Figure 1-18. Topology Change Because of a Link Failure



By default, bridges keep MAC addresses in the bridge table for 5 minutes. When a topology change occurs, the bridge temporarily lowers this timer to the same as the forward delay timer (default: 15 seconds). This allows the STP network to react to changes in topology by having the bridges quickly relearn the MAC address changes that occur when links change state. Without this, network devices could be unreachable for up to 5 minutes while the bridge ages the MAC address out. This is typically called a black hole because data is forwarded toward a bridge that no longer can reach the network device. Topology change BPDUs are a mechanism to overcome this. A common misconception is that topology change BPDUs cause STP to recalculate. The purpose of topology change BPDUs is to avoid black holes and allow the bridges to have up-to-date bridge tables. STP recalculations only occur only when the bridge sees BPDUs with better paths through the bridged network or when the bridge no longer receives configuration BPDUs from the root bridge.

This section offers a simple introduction to spanning tree. Later chapters include examples of the complexities of spanning tree and the various enhancement features available.

## LAN Switch Architecture

- [Receiving Data-Switching Modes](#)
  - [Cut-Through Mode](#)

- Fragment-Free Mode
- Store-and-Forward Mode
- Switching Data
  - Shared Bus Switching
  - Crossbar Switching
- Buffering Data
  - Port Buffered Memory
  - Shared Memory
- Oversubscribing the Switch Fabric
- Congestion and Head-of-Line Blocking
- Forwarding Data

## Receiving Data-Switching Modes

The first step in LAN switching is receiving the frame or packet, depending on the capabilities of the switch, from the transmitting device or host. Switches making forwarding decisions only at Layer 2 of the OSI model refer to data as frames, while switches making forwarding decisions at Layer 3 and above refer to data as packets. This chapter's examination of switching begins from a Layer 2 point of view. Depending on the model, varying amounts of each frame are stored and examined before being switched.

Three types of switching modes have been supported on Catalyst switches:

- Cut through
- Fragment free
- Store and forward

These three switching modes differ in how much of the frame is received and examined by the switch before a forwarding decision is made. The next sections describe each mode in detail.

### Cut-Through Mode

Switches operating in cut-through mode receive and examine only the first 6 bytes of a frame. These first 6 bytes represent the destination MAC address of the frame, which is sufficient information to make a forwarding decision. Although cut-through switching offers the least latency when transmitting frames, it is susceptible to transmitting fragments created via Ethernet collisions, runts (frames less than 64 bytes), or damaged frames.

### Fragment-Free Mode

Switches operating in fragment-free mode receive and examine the first 64 bytes of frame. Fragment free is referred to as "fast forward" mode in some Cisco Catalyst documentation. Why examine 64 bytes? In a properly designed Ethernet network, collision fragments must be detected in the first 64 bytes.

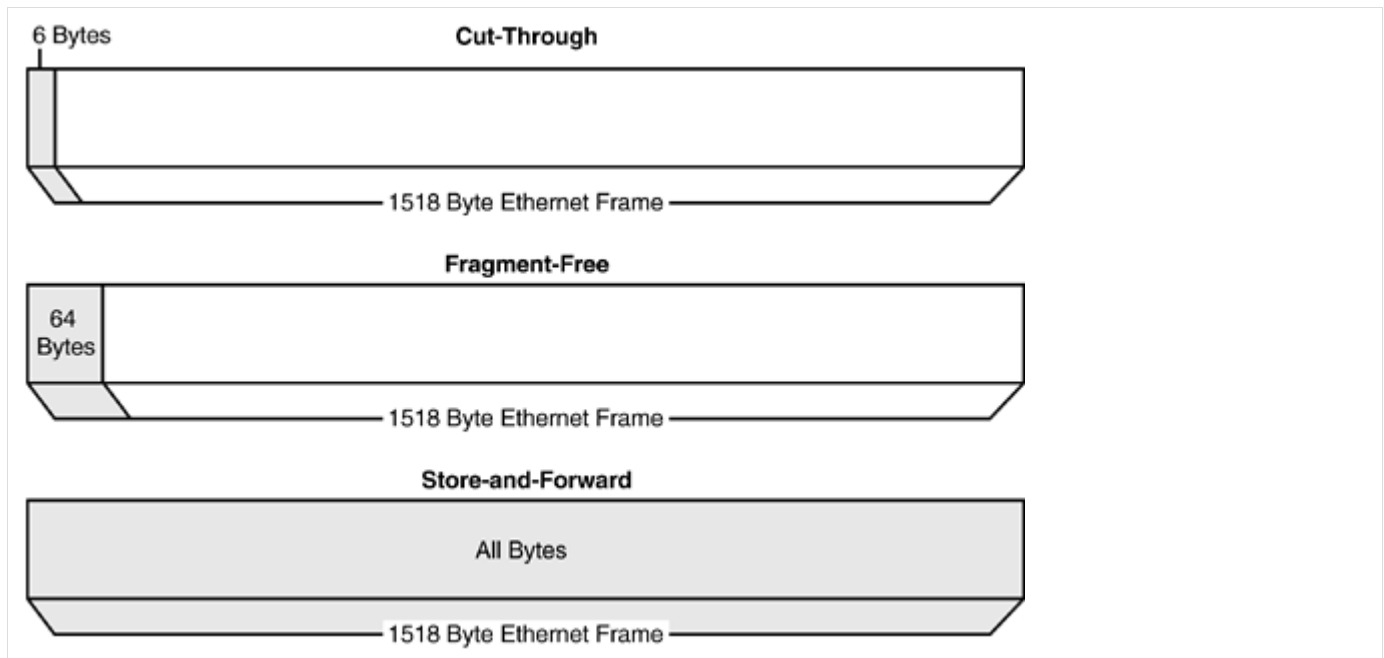
### Store-and-Forward Mode

Switches operating in store-and-forward mode receive and examine the entire frame, resulting in the most error-free type of switching.

As switches utilizing faster processor and application-specific integrated circuits (ASICs) were introduced, the need to support cut-through and fragment-free switching was no longer necessary. As a result, all new Cisco Catalyst switches utilize store-and-forward switching.

Figure 2-1 compares each of the switching modes.

Figure 2-1. Switching Modes



## Switching Data

Regardless of how many bytes of each frame are examined by the switch, the frame must eventually be switched from the input or ingress port to one or more output or egress ports. A switch fabric is a general term for the communication channels used by the switch to transport frames, carry forwarding decision information, and relay management information throughout the switch. A comparison could be made between the switching fabric in a Catalyst switch and a transmission on an automobile. In an automobile, the transmission is responsible for relaying power from the engine to the wheels of the car. In a Catalyst switch, the switch fabric is responsible for relaying frames from an input or ingress port to one or more output or egress ports. Regardless of model, whenever a new switching platform is introduced, the documentation will generally refer to the "[transmission](#)" as the switching fabric.

Although a variety of techniques have been used to implement switching fabrics on Cisco Catalyst platforms, two major architectures of switch fabrics are common:

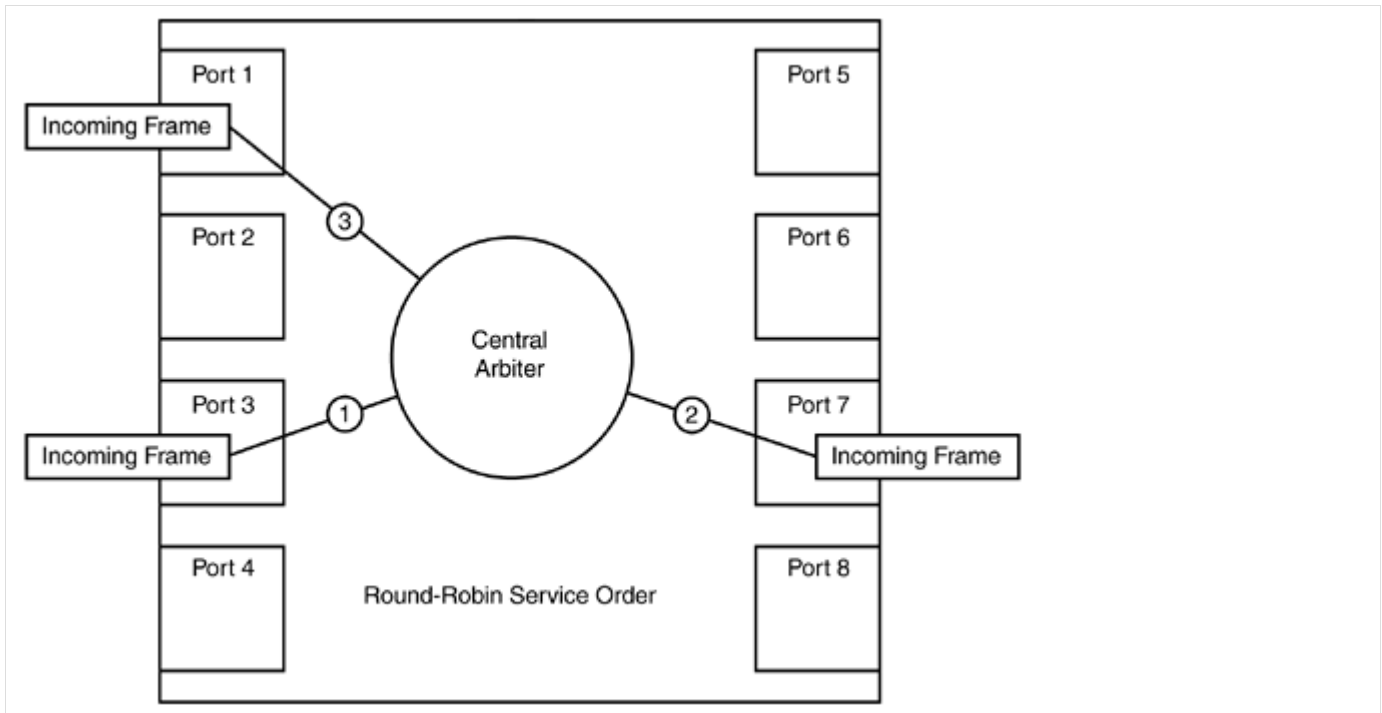
- Shared bus
- Crossbar

### Shared Bus Switching

In a shared bus architecture, all line modules in the switch share one data path. A central arbiter determines how and when to grant requests for access to the bus from each line card. Various methods of achieving fairness can be used by the arbiter depending on the configuration of the switch. A shared bus architecture is much like multiple lines at an airport ticket counter, with only one ticketing agent processing customers at any given time.

[Figure 2-2](#) illustrates a round-robin servicing of frames as they enter a switch. Round-robin is the simplest method of servicing frames in the order in which they are received. Current Catalyst switching platforms such as the Catalyst 6500 support a variety of quality of service (QoS) features to provide priority service to specified traffic flows. [Chapter 8](#), "Understanding Quality of Service on Catalyst 6500," will provide more information on this topic.

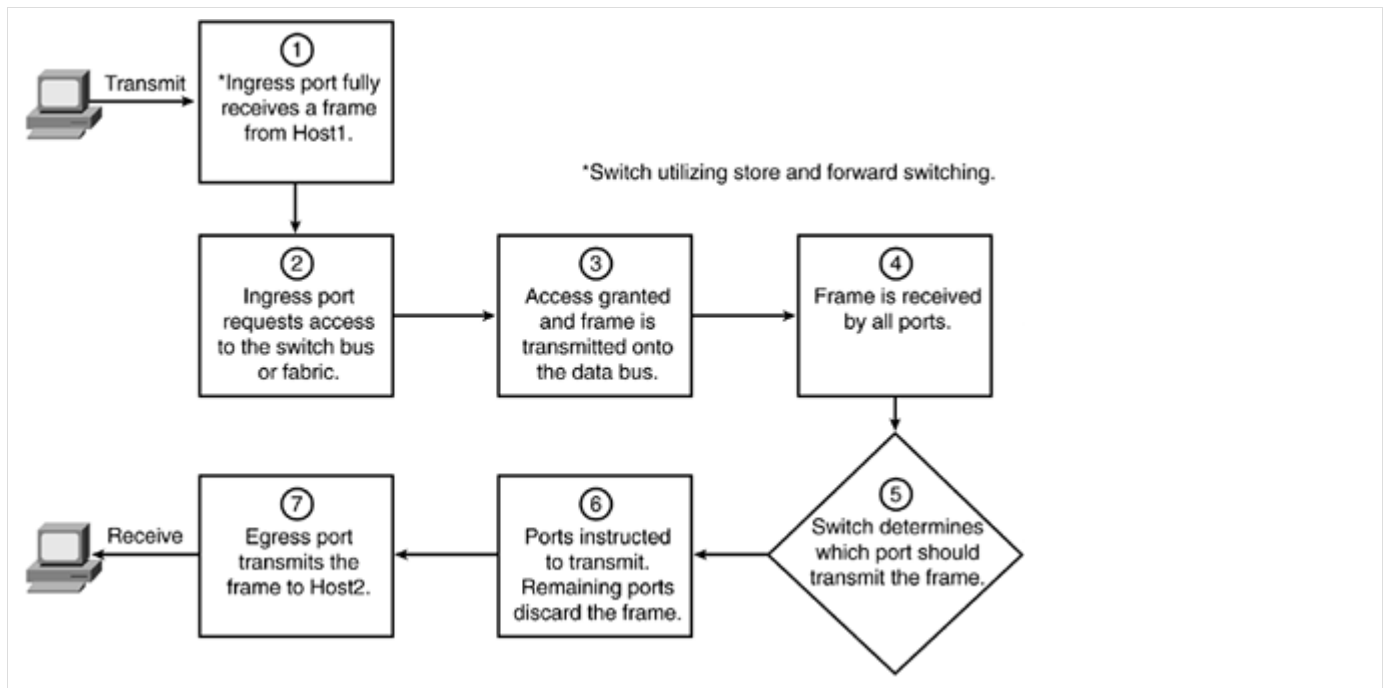
Figure 2-2. Round-Robin Service Order



The following list and [Figure 2-3](#) illustrate the basic concept of moving frames from the received port or ingress, to the transmit port(s) or egress using a shared bus architecture:

- Frame received from Host1 - The ingress port on the switch receives the entire frame from Host1 and stores it in a receive buffer. The port checks the frame's Frame Check Sequence (FCS) for errors. If the frame is defective (runt, fragment, invalid CRC, or Giant), the port discards the frame and increments the appropriate counter.
- Requesting access to the data bus - A header containing information necessary to make a forwarding decision is added to the frame. The line card then requests access or permission to transmit the frame onto the data bus.
- Frame transmitted onto the data bus - After the central arbiter grants access, the frame is transmitted onto the data bus.
- Frame is received by all ports - In a shared bus architecture, every frame transmitted is received by all ports simultaneously. In addition, the frame is received by the hardware necessary to make a forwarding decision.
- Switch determines which port(s) should transmit the frame - The information added to the frame in step 2 is used to determine which ports should transmit the frame. In some cases, frames with either an unknown destination MAC address or a broadcast frame, the switch will transmit the frame out all ports except the one on which the frame was received.
- Port(s) instructed to transmit, remaining ports discard the frame - Based on the decision in step 5, a certain port or ports is told to transmit the frame while the rest are told to discard or flush the frame.
- Egress port transmits the frame to Host2 - In this example, it is assumed that the location of Host2 is known to the switch and only the port connecting to Host2 transmits the frame.

Figure 2-3. Frame Flow in a Shared Bus



[View full size image](#)

One advantage of a shared bus architecture is every port except the ingress port receives a copy of the frame automatically, easily enabling multicast and broadcast traffic without the need to replicate the frames for each port. This example is greatly simplified and will be discussed in detail for Catalyst platforms that utilize a shared bus architecture in [Chapter 3, "Catalyst Switching Architecture."](#)

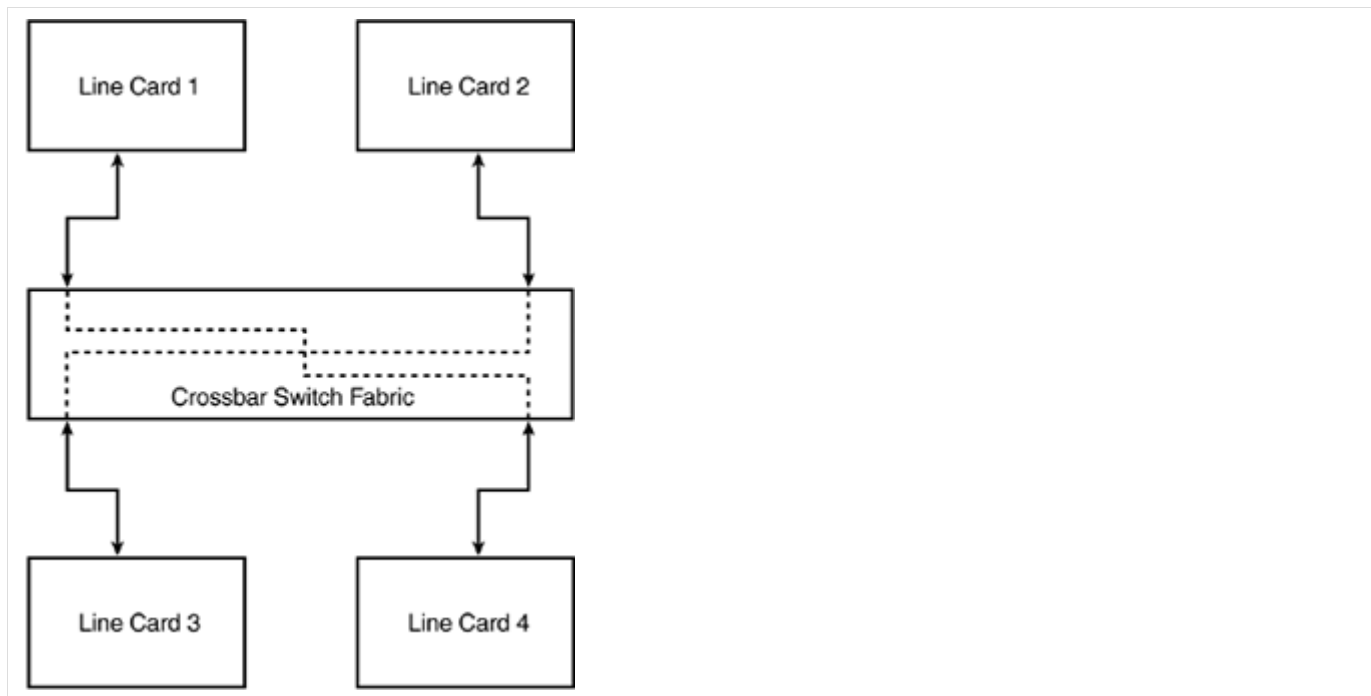
## Crossbar Switching

In the shared bus architecture example, the speed of the shared data bus determines much of the overall traffic handling capacity of the switch. Because the bus is shared, line cards must wait their turns to communicate, and this limits overall bandwidth.

A solution to the limitations imposed by the shared bus architecture is the implementation of a crossbar switch fabric, as shown in [Figure 2-4](#). The term crossbar means different things on different switch platforms, but essentially indicates multiple data channels or paths between line cards that can be used simultaneously.

Figure 2-4. Crossbar Switch Fabric





In the case of the Cisco Catalyst 5500 series, one of the first crossbar architectures advertised by Cisco, three individual 1.2-Gbps data buses are implemented. Newer Catalyst 5500 series line cards have the necessary connector pins to connect to all three buses simultaneously, taking advantage of 3.6 Gbps of aggregate bandwidth. Legacy line cards from the Catalyst 5000 are still compatible with the Catalyst 5500 series by connecting to only one of the three data buses. Access to all three buses is required by Gigabit Ethernet cards on the Catalyst 5500 platform.

A crossbar fabric on the Catalyst 6500 series is enabled with the Switch Fabric Module (SFM) and Switch Fabric Module 2 (SFM2). The SFM provides 128 Gbps of bandwidth (256 Gbps full duplex) to line cards via 16 individual 8-Gbps connections to the crossbar switch fabric. The SFM2 was introduced to support the Catalyst 6513 13-slot chassis and includes architecture optimizations over the SFM.

## Buffering Data

Frames must wait their turn for the central arbiter before being transmitted in shared bus architectures. Frames can also potentially be delayed when congestion occurs in a crossbar switch fabric. As a result, frames must be buffered until transmitted. Without an effective buffering scheme, frames are more likely to be dropped anytime traffic oversubscription or congestion occurs.

Buffers get used when more traffic is forwarded to a port than it can transmit. Reasons for this include the following:

- Speed mismatch between ingress and egress ports
- Multiple input ports feeding a single output port
- Half-duplex collisions on an output port
- A combination of all the above

To prevent frames from being dropped, two common types of memory management are used with Catalyst switches:

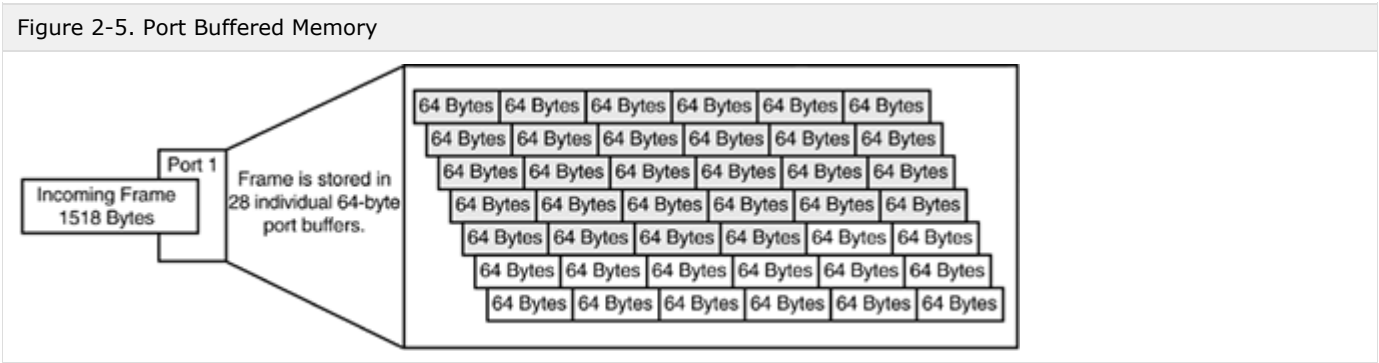
- Port buffered memory
- Shared memory

### Port Buffered Memory

Switches utilizing port buffered memory, such as the Catalyst 5000, provide each Ethernet port with a certain amount of high-speed memory to buffer frames until transmitted. A disadvantage of port buffered memory is the dropping of frames when a port runs out of buffers. One method of maximizing the benefits of buffers is the use of flexible buffer sizes. Catalyst 5000 Ethernet line card port buffer memory is flexible and can create frame buffers for any frame size, making the most of the available buffer memory. Catalyst 5000 Ethernet cards that use the SAINT ASIC contain 192 KB of buffer memory per port, 24 kbps for receive or input buffers, and 168 KB for transmit or output buffers.

Using the 168 KB of transmit buffers, each port can create as many as 2500 64-byte buffers. With most of the buffers in use as an output queue, the Catalyst 5000 family has eliminated head-of-line blocking issues. (You learn more about head-of-line blocking later in this chapter in the section "[Congestion and Head-of-Line Blocking](#).") In normal operations, the input queue is never used for more than one frame, because the switching bus runs at a high speed.

Figure 2-5 illustrates port buffered memory.



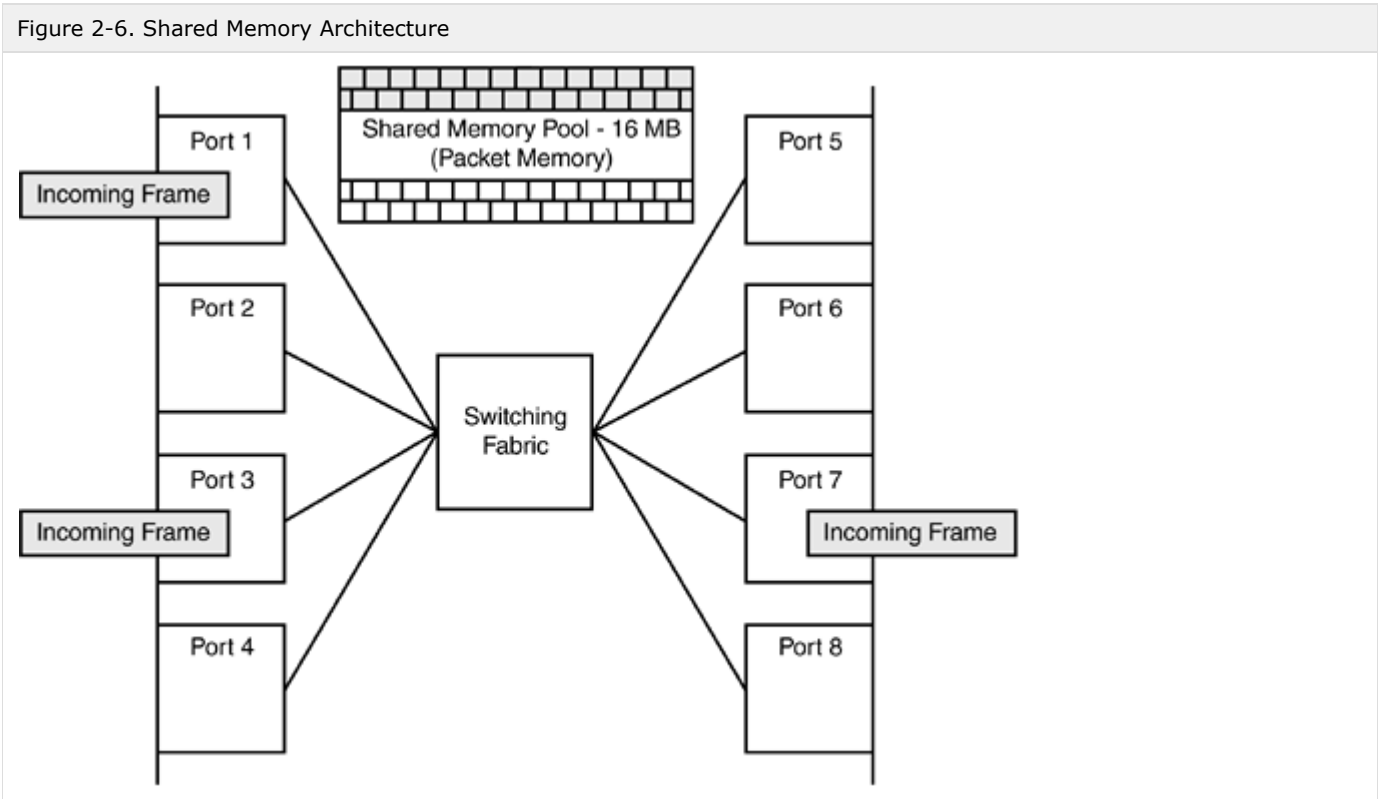
[View full size image](#)

### Shared Memory

Some of the earliest Cisco switches use a shared memory design for port buffering. Switches using a shared memory architecture provide all ports access to that memory at the same time in the form of shared frame or packet buffers. All ingress frames are stored in a shared memory "pool" until the egress ports are ready to transmit. Switches dynamically allocate the shared memory in the form of buffers, accommodating ports with high amounts of ingress traffic, without allocating unnecessary buffers for idle ports.

The Catalyst 1200 series switch is an early example of a shared memory switch. The Catalyst 1200 supports both Ethernet and FDDI and has 4 MB of shared packet dynamic random-access memory (DRAM). Packets are handled first in, first out (FIFO).

More recent examples of switches using shared memory architectures are the Catalyst 4000 and 4500 series switches. The Catalyst 4000 with a Supervisor I utilizes 8 MB of Static RAM (SRAM) as dynamic frame buffers. All frames are switched using a central processor or ASIC and are stored in packet buffers until switched. The Catalyst 4000 Supervisor I can create approximately 4000 shared packet buffers. The Catalyst 4500 Supervisor IV, for example, utilizes 16 MB of SRAM for packet buffers. Shared memory buffer sizes may vary depending on the platform, but are most often allocated in increments ranging from 64 to 256 bytes. [Figure 2-6](#) illustrates how incoming frames are stored in 64-byte increments in shared memory until switched by the switching engine.



### Oversubscribing the Switch Fabric

Switch manufacturers use the term non-blocking to indicate that some or all the switched ports have connections to the switch fabric equal to their line speed. For example, an 8-port Gigabit Ethernet module would require 8 Gb of bandwidth into the switch fabric for the ports to be considered non-blocking. All but the highest end switching platforms and configurations have the potential of oversubscribing access to the switching fabric.

Depending on the application, oversubscribing ports may or may not be an issue. For example, a 10/100/1000 48-port Gigabit Ethernet module with all ports running at 1 Gbps would require 48 Gbps of bandwidth into the switch fabric. If many or all ports were connected to high-speed file servers capable of generating consistent streams of traffic, this one-line module could outstrip the bandwidth of the entire switching fabric. If the module is connected entirely to end-user workstations with lower bandwidth requirements, a card that oversubscribes the switch fabric may not significantly impact performance.

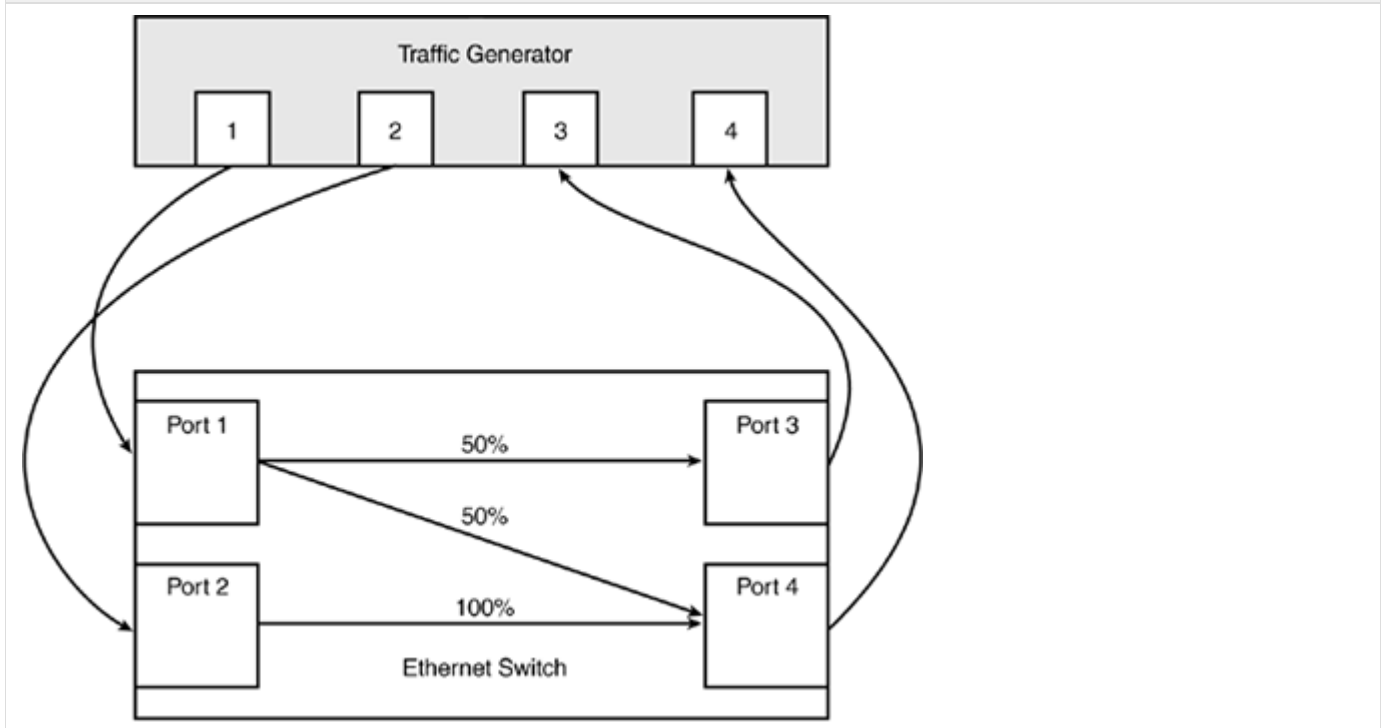
Cisco offers both non-blocking and blocking configurations on various platforms, depending on bandwidth requirements. Check the specifications of each platform and the available line cards to determine the aggregate bandwidth of the connection into the switch fabric.

## Congestion and Head-of-Line Blocking

Head-of-line blocking occurs whenever traffic waiting to be transmitted prevents or blocks traffic destined elsewhere from being transmitted. Head-of-line blocking occurs most often when multiple high-speed data sources are sending to the same destination. In the earlier shared bus example, the central arbiter used the round-robin service approach to moving traffic from one line card to another. Ports on each line card request access to transmit via a local arbiter. In turn, each line card's local arbiter waits its turn for the central arbiter to grant access to the switching bus. Once access is granted to the transmitting line card, the central arbiter has to wait for the receiving line card to fully receive the frames before servicing the next request in line. The situation is not much different than needing to make a simple deposit at a bank having one teller and many lines, while the person being helped is conducting a complex transaction.

In [Figure 2-7](#), a congestion scenario is created using a traffic generator. Port 1 on the traffic generator is connected to Port 1 on the switch, generating traffic at a 50 percent rate, destined for both Ports 3 and 4. Port 2 on the traffic generator is connected to Port 2 on the switch, generating traffic at a 100 percent rate, destined for only Port 4. This situation creates congestion for traffic destined to be forwarded by Port 4 on the switch because traffic equal to 150 percent of the forwarding capabilities of that port is being sent. Without proper buffering and forwarding algorithms, traffic destined to be transmitted by Port 3 on the switch may have to wait until the congestion on Port 4 clears.

Figure 2-7. Head-of-Line Blocking



Head-of-line blocking can also be experienced with crossbar switch fabrics because many, if not all, line cards have high-speed connections into the switch fabric. Multiple line cards may attempt to create a connection to a line card that is already busy and must wait for the receiving line card to become free before transmitting. In this case, data destined for a different line card that is not busy is blocked by the frames at the head of the line.

Catalyst switches use a number of techniques to prevent head-of-line blocking; one important example is the use of per port buffering. Each port maintains a small ingress buffer and a larger egress buffer. Larger output buffers (64 Kb to 512 k shared)

allow frames to be queued for transmit during periods of congestion. During normal operations, only a small input queue is necessary because the switching bus is servicing frames at a very high speed. In addition to queuing during congestion, many models of Catalyst switches are capable of separating frames into different input and output queues, providing preferential treatment or priority queuing for sensitive traffic such as voice. [Chapter 8](#) will discuss queuing in greater detail.

## Forwarding Data

Regardless of the type of switch fabric, a decision on which ports should forward a frame and which should flush or discard the frame must occur. This decision can be made using only the information found at Layer 2 (source/destination MAC address), or on other factors such as Layer 3 (IP) and Layer 4 (Port). Each switching platform supports various types of ASICs responsible for making the intelligent switching decisions. Each Catalyst switch creates a header or label for each packet, and forwarding decisions are based on this header or label. [Chapter 3](#) will include a more detailed discussion of how various platforms make forwarding decisions and ultimately forward data.

## Layer 2 Fundamentals

- [Understanding Legacy LAN Segment](#)
- [Introducing Virtual LANs](#)
- [Trunking Methods](#)
  - [Inter-Switch Link Protocol](#)
  - [IEEE 802.1Q](#)
  - [Configuration Best Practices](#)
- [VLAN Trunking Protocol](#)
  - [Summary Advertisement](#)
  - [Subset Advertisement](#)
  - [Advertisement Request](#)
  - [Join](#)
  - [VTP Example 1](#)
  - [VTP Example 2](#)
  - [VTP Mode Best Practices](#)
- [Configuring VTP/VLAN/Trunk](#)
- [VLAN Pruning](#)
- [EtherChannel](#)
- [Understanding VLAN 1](#)
  - [Management VLAN](#)
  - [Management VLAN Best Practices](#)
- [Private VLANs](#)

## Understanding Legacy LAN Segment

In a VLAN environment, a user can be situated physically anywhere in the network, and still be part of a group of users that have the same requirements such as IP addressing and specific network privileges. In legacy networks, hosts with similar network requirements had to be on the same LAN segment. On the other hand, VLANs are logical and not bound to any physical location.

In legacy networks, users connected to a hub, which may have been connected to some kind of bridge or router if wide-area connectivity was required (see [Figure 4-1](#)). Multiple hubs were interconnected to allow more users on the same segment.

Figure 4-1. Multiple Hubs Connected to a Router or Bridge



This type of network enabled a department with a large number of users to remain part of the same segment. This setup was easy for a network engineer to implement, because generally one network configuration was supported for that department. For example, the department likely had local file servers on the same Layer 2 segment allowing most, if not all, traffic to stay

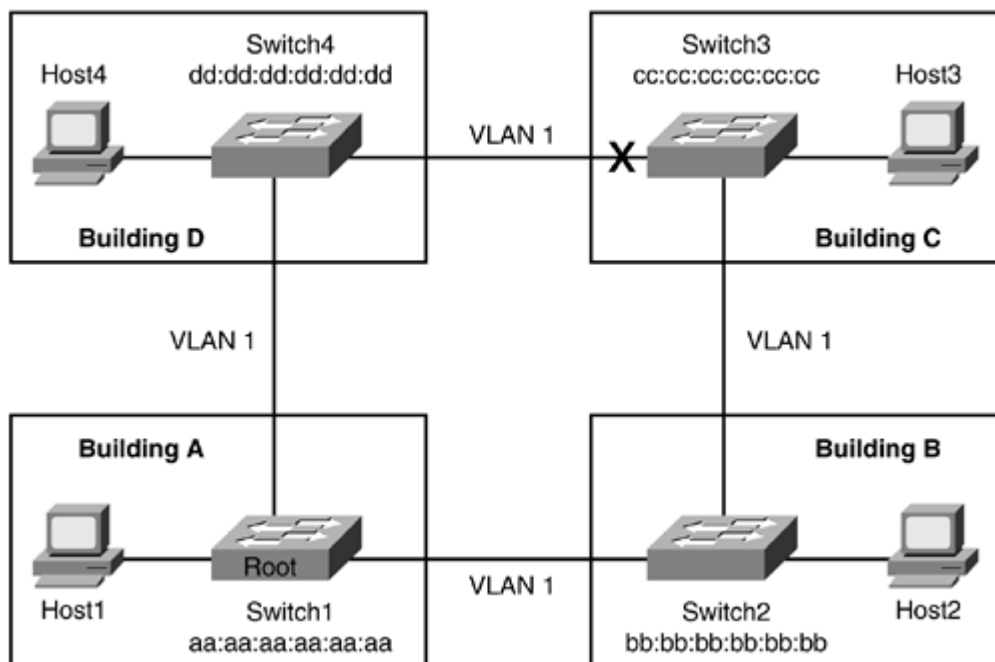
local.

Cascading hubs allow more users to be on the same segment at the cost of causing more contention among hosts trying to access the network. A bigger collision domain translates to more packet collisions on the network, and, therefore, more retransmissions and possible packet drops. By cascading multiple hubs, the possibility of exceeding the cabling requirements of IEEE 802.3 standard exists, which translates to late collisions, causing packet loss. A defective network interface card (NIC) that "jabbers" can cause severe performance issues on this type of network because it continuously sends garbled data, ignoring the carrier sense multiple access collision detect (CSMA/CD) rule that you learned about in [Chapter 1](#), "LAN Switching Foundation Technologies."

## Introducing Virtual LANs

With the advent of VLANs, the same users in the legacy example are no longer restricted by physical cabling to be on the same logical segment or have the same access and privileges on the network. VLAN implementation makes the network more flexible as shown in [Figure 4-2](#). Notice that one of the ports off Switch3 is crossed out in the drawing. Spanning tree has put that specific port into blocking state to prevent loops in the network as discussed in "Introducing Spanning Tree Protocol" section of [Chapter 1](#).

Figure 4-2. Users Located in Multiple Buildings



In the legacy network, when a host moved to a different location on another router port, the IP address of the host had to be changed. Possible network changes were necessary to accommodate the user; for example, a change in a router access list allowing the user access back to a department server. The main point is that it was not easy to move users in the network without some type of change in the network or host. In an environment that supports VLANs, such network changes are not necessary because of flat Layer 2 network infrastructure. If Host1 in [Figure 4-2](#) is moved to Building B, there is no need to change the configuration of the user's machine or the network. The user simply plugs the host into the jack and is ready to go.

**NOTE:** Keep in mind that the discussion thus far has been strictly focused on Layer 2, where a single VLAN is extended to multiple switches and with one instance of spanning tree.

Initially, there was big push to extend VLANs across the network. In fact, most universities implemented this technology because the implementation was relatively simple, and many applications at the time had a requirement to be on the same Layer 2 network because of their communications protocols. Network engineers simply configured a VLAN with a large IP range. They pushed security and other network policies on these VLANs on the fly. This was great in saving time and money.

The risks associated with such an implementation were quickly noticed. Extending VLANs has a dark side, enlarging the broadcast domain. If a single host sends out a broadcast message, every machine on that VLAN, regardless of the number of buildings and switches involved, receives that broadcast message. The result is excessive traffic on the network. The greater penalty is a broadcast storm, occurring when a host sends an incorrect broadcast message that is received by all hosts on that VLAN, and all those hosts broadcast as well. This process can eventually bring a flat Layer 2 network to its knees.

Spanning trees can also bring the network down when VLANs are extended across the switched network. Too much traffic on the network or some partial or complete hardware failure can cause a spanning-tree outage. In a spanning-tree outage, spanning tree is unable to calculate a loop-free topology correctly, and a loop occurs in the network. Similar to the example of

a loop in transparent bridging, traffic exponentially increases causing a network meltdown until the loop is broken, many times requiring manual intervention.

A VLAN is tagged with a user-defined number to differentiate it from another VLAN. For instance, users on VLAN 4 are members of the same subnet and are on the same broadcast domain, whereas VLAN 5 has its own users and broadcast domain. Typically an enterprise switch has no more than 30 VLANs configured on a switch. Depending on the trunking mechanism used, the number of VLANs configured on a switch can be as high as 4096 minus some reserved VLANs. The "[Trunking Methods](#)" section later in this chapter discusses trunking further.

[Table 4-1](#) provides the valid range of VLANs that can be configured on a switch. The Catalyst 5500 switch does not support the extended VLANs that fall in the 1025 - 4096 range. The trunking mechanism used might limit the number of VLANs available for use. For example, Inter-Switch Link (ISL) does not support extended VLAN range. The "[VLAN Trunking Protocol](#)" section later in this chapter will discuss VTP further.

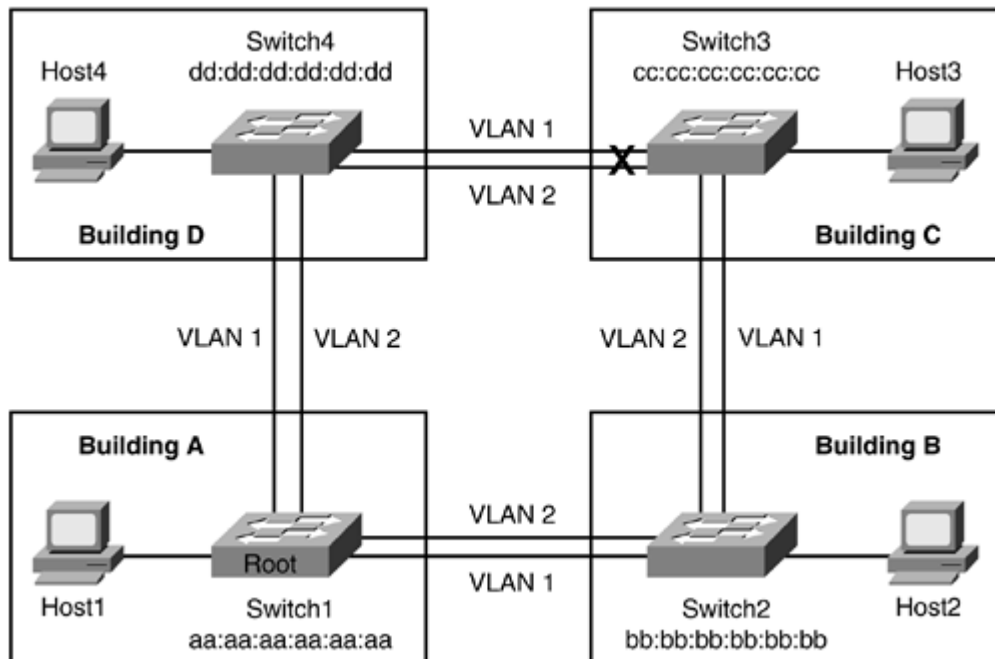
Table 4-1. Valid VLAN Range

VLANs	Range	Usage	Propagated by VTP (Y/N)
0, 4095	Reserved range	For system use only. You cannot see or use these VLANs.	N/A
1	Normal range	Cisco default. You can use this VLAN but you cannot delete it.	Yes
2 - 1000	Normal range	Used for Ethernet VLANs. You can create, use, and delete these VLANs.	Yes
1001	Normal range	You cannot create or use this VLAN. May be available in the future.	Yes
1002 - 1005	Reserved range	Cisco defaults for FDDI and Token Ring. Not supported on Catalyst 6000 family switches. You cannot delete these VLANs.	N/A
1006 - 1009	Reserved range	Cisco defaults. Not currently used but might be used for defaults in the future. Nonreserved VLANs may be mapped to these reserved VLANs when necessary.	N/A
1010-1024	Reserved range	These VLANs might not be seen, but can be mapped to nonreserved VLANs when necessary.	N/A
1025-4094	Extended range	For Ethernet VLANs only. These may be created, used, and deleted with the following exception: FlexWAN modules and routed ports automatically allocate a sequential block of internal VLANs starting at VLAN 1025. If the devices are used, the required number of VLANs must be allowed for.	No

## Trunking Methods

The examples so far in this chapter demonstrate one VLAN spanning multiple switches. In the real world, typically, a great number of VLANs are configured, which are extended to multiple switches. [Figure 4-3](#) shows two VLANs in the switched network. Each VLAN has its own STP-topology, IP range, and network requirements.

Figure 4-3. Multiple Switches with Two VLANs

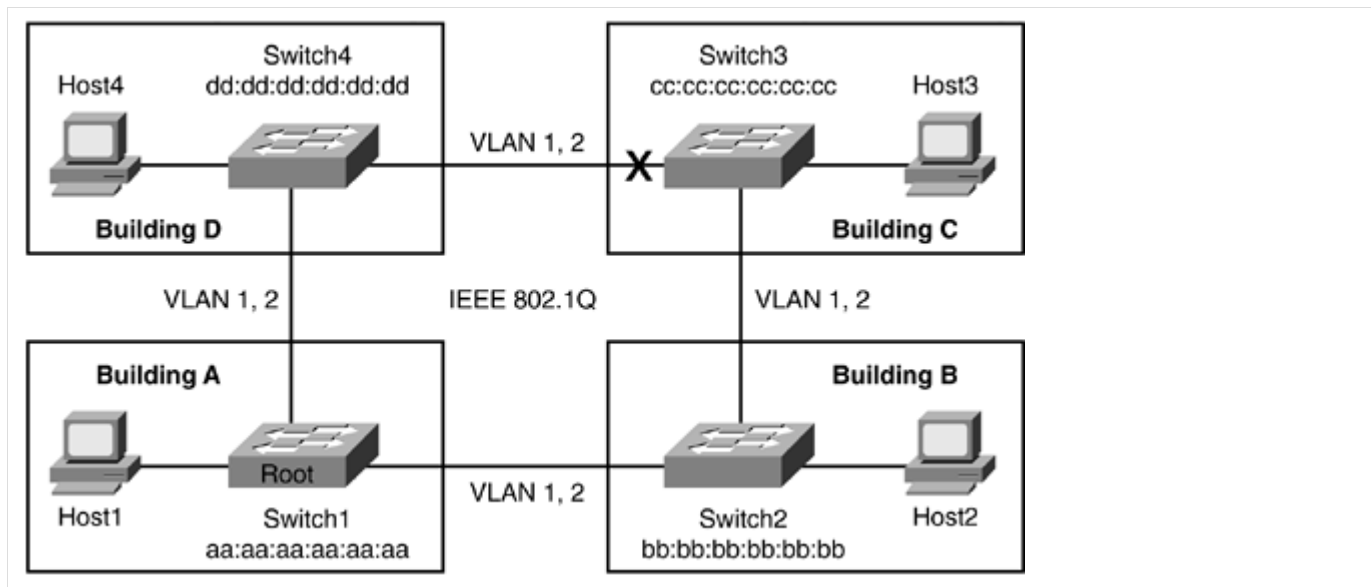


Now imagine a third VLAN is added, which requires another physical link between the switches. Because it may not be efficient to continue to add links as the number of VLANs grows, the solution is trunking. A trunk can be configured between two or more switches, between a router and a switch, or between a switch and a host such as a server. Check the hardware to find out what type of trunking capabilities a device has, if any. This section will primarily concentrate on trunking between Cisco switches.

A trunk multiplexes multiple VLANs over a single physical connection. This kind of multiplexing is conceptually similar to the way many television signals are multiplexed onto the airwaves using different frequencies. In this analogy, each VLAN acts like a different television station, while sharing the same physical wire.

Cisco supports only two types of trunks using Fast, Gigabit, and 10 Gigabit Ethernet ports: Cisco Inter-Switch Link Protocol (ISL) and IEEE 802.1Q. The Dynamic Trunking Protocol (DTP) allows a port to negotiate which method to use for trunking. DTP will first attempt to form an ISL trunk if both switches support it; if not, DTP will attempt IEEE 802.1Q. DTP uses the address 01-00-0C-CC-CC-CC with a SNAP value of 0x2004. DTP sends messages every 1 second, and after formation of the trunk, every 30 seconds. The ports negotiating the trunk will not participate in spanning tree until the negotiation is complete. Recently, IEEE 802.1Q is being implemented in networks because it is an IEEE standard, whereas ISL is proprietary to Cisco. Many Cisco routers and some older Cisco Catalyst switches do not support dynamic trunking. In these cases, a static configuration is required. [Figure 4-4](#) shows two VLANs 1 and 2 running over the same cable because of the use of trunking, whereas before, an extra connection was required between the switches in the diagram as depicted in [Figure 4-3](#).

Figure 4-4. Trunks Used Between Switches



Trunking is an integral part of networking, and it is worth going over Cisco ISL and IEEE 802.1Q methods in detail. This section will also provide some best practices that will help with properly configuring the switches.

## Inter-Switch Link Protocol

ISL encapsulates the Ethernet frame with a 26-byte header and a 4-byte frame check sequence (FCS) for a total of 30 bytes of overhead. ISL requires a minimum Fast Ethernet connection between the two devices. The 15-bit VLAN field in the ISL header allows for the multiplexing of the VLANs on a single wire. ISL supports up to 1024 VLANs because Cisco switches use the lower 10 bits of the 15-bit field. The range of ISL packet sizes is 94 bytes (64-byte minimum Ethernet frame + 30-byte ISL overhead) to 1548 bytes (1518-byte maximum Ethernet frame + 30-byte ISL overhead). Each VLAN will have its own spanning-tree topology in an ISL trunking configuration. For instance, if there are two VLANs configured on an ISL trunk, each VLAN will have its own root and spanning-tree topology layout.

The following describes the fields of the ISL encapsulation frame shown in [Figure 4-5](#):

- DA - The destination address uses the multicast MAC address 01-00-0C-00-00-00.
- Type - The type of frame encapsulated: Ethernet (0000), Token Ring (0001), FDDI (0010), and ATM (0011).
- User - This field is used as an extension for the technologies covered under the Type field. The User field can also be used to define priority of the frame. The default value is 0000 for Ethernet with low-priority traffic.
- SA - Source address of the switch transmitting the ISL frame.
- Len - The length of the packet.
- AAAA03 - Standard SNAP 802.2 LLC header. This value is constant.
- HSA - High bits of SA.
- VLAN - VLAN ID.
- BPDU - STP bridge protocol data unit/Cisco Discovery Protocol (BPDU/CDP) for control traffic.
- Index - The port index of the source of the packet.
- Res - Reserved field for additional information, for instance, Token Ring or FDDI Frame Check Sequence field. For Ethernet, this field should be zero.
- Encap Frame - The actual Ethernet frame.
- ISL CRC - Four-byte check on the ISL packet to ensure it is not corrupted.

Figure 4-5. Frame Using ISL Encapsulation

Bits	40	4	4	48	16	24	24	15	1	16	16	Various	32
Frame Field	DA	Type	User	SA	Len	AAAA03	HSA	VLAN	BPDU	Index	Res	Encap Frame	FCS

[View full size image\](#)

## IEEE 802.1Q

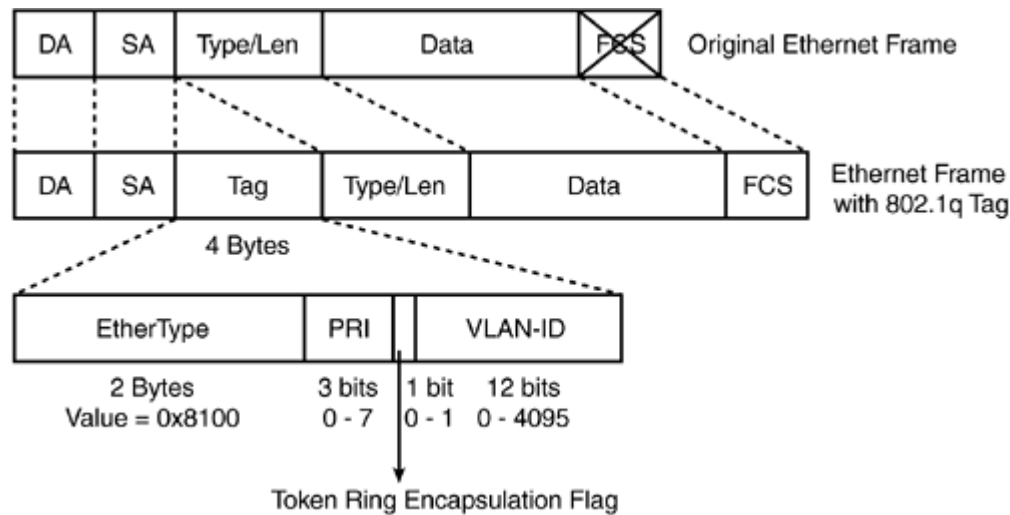
While ISL encapsulates an Ethernet frame with a 30-byte header, IEEE 802.1Q simply adds an additional 4-byte Tag field to the Ethernet frame (EtherType 0x8100). The Tag field has three components in addition to the EtherType:



- Priority (3 bits) - The Priority field is used by 802.1p to implement Layer 2 quality of service (QoS).
- Canonical Format Identifier (CFI) (1 bit) - The CFI bit is used for compatibility purposes between Ethernet and Token Ring.
- VLAN ID (VID) (12 bits) - The VID field is used to distinguish between VLANs on the link.

FCS is recomputed after the 4-byte tag is inserted. IEEE 802.1Q supports up to 4096 VLANs because of the 12-bit length. The IEEE 802.1Q tag is not inserted on the native VLAN, which is the VLAN that the port was assigned to before becoming a trunk port. Figure 4-6 illustrates the IEEE 802.1Q tag format.

Figure 4-6. IEEE 802.1Q Tag Format



If the adjoining trunk port's native VLAN is different from the local port on the switch, a native mismatch VLAN error occurs. A mismatched native VLAN scenario will bridge VLAN STP information, which translates to having one single STP rather than STP for each VLAN defined. Example 4-1 shows an asterisk on the remote switch's port 1/1, which has a different native VLAN.

#### Example 4-1. Detecting a Native VLAN Mismatch

```
%CDP-4-NVLANMISMATCH:Native vlan mismatch detected on port 1/2
```

```
Switch1 (enable) show cdp neighbor
```

```
* - indicates vlan mismatch.
```

```
# - indicates duplex mismatch.
```

Port	Device-ID	Port-ID	Platform
1/2	Switch#2	1/1*	WS-C6506

## Configuration Best Practices

Trunking has five modes in which it can operate:

- On
- Off
- Desirable
- Auto
- Nonegotiate

In nonegotiate mode, the switch will form a trunk, but will not send DTP frames. The other end switch has to be in On or

Nonegotiate mode for nonegotiate to work. Typically, this type of setup is used for connecting a third-party switch that does not support DTP. Table 4-2 provides the details on the various trunking modes.

Table 4-2. Summary of Five Trunking Mode

Mode	Description
On	Forces the port to become a trunk port and persuades the neighboring port to become a trunk port. The port becomes a trunk port even if the neighboring port does not agree to become a trunk.
Off	Forces the port to become a non-trunk port and persuades the neighboring port to become a non-trunk port. The port becomes a non-trunk port even if the neighboring port does not agree to become a non-trunk port.
Desirable	Causes the port to negotiate actively with the neighboring port to become a trunk link.
Auto	Causes the port to become a trunk port if the neighboring port tries to negotiate a trunk link.
Nonegotiate	Forces the port to become a trunk port but prevents it from sending DTP frames to its neighbor.

Cisco recommends Desirable-Desirable mode for all trunk ports.

## VLAN Trunking Protocol

In Figure 4-4, two VLANs extend over multiple switches using trunking. Because each switch sharing trunks must support common VLAN information for the trunks to function correctly, Cisco created the VLAN Trunking Protocol (VTP) for creating and managing that VLAN information. It should be noted that any VLAN created on a switch is in an inactive state until VTP is configured.

A collection of switches that are under the same administrative control and will support the same range of configured VLANs are said to be in the same VTP domain. A domain name is simply a unique identifier up to 32 characters long used to identify the switches that will share the same VTP information. The domain name is also case sensitive.

VTP packets are sent to destination address 01-00-0C-CC-CC-CC with a SNAP type of 0x2003. Each switch can operate in one of three modes:

- Server (default)
- Client
- Transparent

In server mode, the switch has a list of all the VLANs for that domain. It can add, delete, or rename any VLAN, and the configuration information is stored in nonvolatile random-access memory (NVRAM). In client mode, the switch obtains its information for the VLAN database from a VTP server, and it cannot make any modifications to it. The information learned by the client switch is not stored in NVRAM. If the client switch is rebooted, the switch must dynamically learn all the VLAN information again from a VTP server. In transparent mode, a switch does not participate in VTP; it merely passes the VTP advertisements to other switches. In transparent mode, the switch can be configured to add, delete, and modify, and the information is stored in NVRAM.

Certain requirements must be met before VTP can be used to manage a domain and distribute VLAN information. Each switch must have a configured trunk port, use the same domain name, and be directly connected. As noted earlier, the trunk port is used to send the VTP information to the adjacent switch. VTP can automatically distribute VLAN information to all other switches in the same domain through a trunk port, or allow manually for each switch to be configured. The dynamic process using server/client mode is administratively palatable because it is easy to implement; a server switch is configured with VLANs, and the rest of the switches in that domain receive that information. On the other hand, server/client mode can pose potential risks on the network, which will be discussed in this section shortly. Transparent mode requires manually configuring each switch.

VTP has four types of messages:

- Summary advertisements (0x01)
- Subset advertisement (0x02)
- Advertisement requests (0x03)
- Join (0x04)

The two types of VTP versions, version 1 and version 2, have some major differences. Version 2 has support for Token Ring. In version 2, switches running in transparent mode forward VTP advertisements they receive regardless of VTP version or domain

name; switches configured for VTP version 1 ignore VTP advertisements with a different VTP domain name than the one configured. Cisco switches default to version 1.

## Summary Advertisement

A switch configured as a VTP server sends a summary advertisement every 5 minutes to inform other connected switches of the domain name and revision number. The revision number is tied to changes in VLAN information and increments each time a modification is made on the VTP server switch. When a switch receives a revision number, it compares it to its own. If the number is the same or lower, the switch ignores the summary advertisement.

In [Example 4-2](#), the debug output shows that the switch received a summary advertisement that has a lower revision number than the one that is currently on the switch. Therefore, the switch will ignore the VTP message.

### Example 4-2. Debug Output of Summary Advertisement

```
VTP: domain Cisco, current rev = 6 found for summary pkt
```

```
VTP: summary packet rev 2 lower than domain Cisco rev 6
```

If the revision number is higher, it will update the VLAN database with the information received. The VTP revision number is extremely important because a higher value revision number always wins. Imagine a situation where a switch used only for testing is accidentally connected to a production network. If the test switch is configured with the same VTP domain name as the production network and has a higher revision number, all production switches in that domain will synchronize to it. All previously used VLAN information is overwritten in favor of the VLAN database on the test switch. If the test switch has not been configured for the same VLANs as the production environment, switched ports will revert back to being members of VLAN 1, resulting in loss of connectivity. Always check the revision number of a new switch before bringing it on the network regardless if the switch is going to operate in client or server mode. Make sure the revision number is lower than the production server mode switch. An easy way to ensure that a new switch does not affect the operation of the other switches in the VTP domain is to simply change the domain name of the new switch to something bogus and back to the valid domain name. At this point, it is safe to bring the new switch to the production network, because any time a VTP domain name is changed, the revision number is reset. A reboot will also reset the revision number.

## Subset Advertisement

Subset advertisement sends the list of VLANs to the client and server switches. This is the actual database that is being pushed to the switches. The subset advertisement gives information about the name of the VLAN, its status, type, and so on. More than one switch can be configured as a VTP server, and VTP servers will negotiate VLAN information until their databases are synchronized using subset advertisement messages. In [Example 4-3](#), the switch receives information about VLANs 12, 30, 34, 100, and notification of a new VLAN, 111. This output can be collected using the `set trace vtp` command on the switch. Only during networking troubleshooting and as a last measure should the `set trace vtp` command be used because the command taxes the resources of the switch.

#### Example 4-3. Debug of Subset Advertisement

```
VTP/Active: Opening vlan_EVENT_ET event - vlan=vlan12 mode=3
VTP/Active: Closing event
VTP/Active: Opening vlan_EVENT_ET event - vlan=vlan30 mode=3
VTP/Active: Closing event
VTP/Active: Opening vlan_EVENT_ET event - vlan=vlan34 mode=3
VTP/Active: Closing event
VTP/Active: Opening vlan_EVENT_ET event - vlan=vlan0100 mode=3
VTP/Active: Closing event
VTP/Active: Opening vlan_EVENT_ET event - vlan=vlan0111 mode=1
vtp_vlan_change_notification: vlan = 111, mode = 1
2003 Sep 04 10:44:16.110 setVtpVlanInformation: vlanNo [111], mode [1], remoteSpan [0], remote_span [0] primary[0] PType[0], mistp[0]
2003 Sep 04 10:44:16.250
```

### Advertisement Request

An advertisement request is sent when a switch has rebooted, the domain name has been changed, or the VTP summary revision number is higher than what is locally on the switch. As noted in [Example 4-4](#), the switch is requesting VTP database information from its directly connected neighbor.

#### Example 4-4. VTP Advertisement Request

```
VTP: tx vtp request, domain Cisco, start value
```

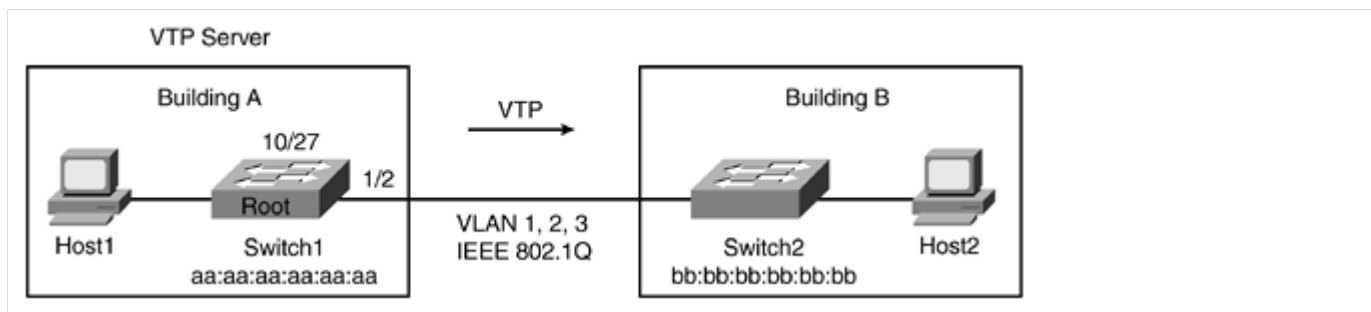
### Join

VTP join messages prevent the upstream switches from pruning a VLAN on a trunk. The "[VLAN Pruning](#)" section later in this chapter will expand the role of this message type.

### VTP Example 1

[Figure 4-7](#) shows two switches participating in VTP domain. The server switch will propagate its VLAN information to the client switch. Any VLAN changes must occur on Switch1. The client, Switch2, will not lose its VLAN information if its connection is severed to the VTP server. However, VLAN information will be lost if the client switch is rebooted.

Figure 4-7. VTP Between Two Switches



## VTP Example 2

Switch1 is a VTP server that is configured with VLANs 2 and 3 (see [Example 4-5](#)). Switch2, a new device on the network, is connected on the same VTP domain as Switch1, as shown in [Figure 4-7](#).

Normally, bringing a new switch on the network is a rudimentary process, but in this case, the revision number of Switch2 is higher than Switch1. The higher VTP revision number will cause Switch1 to synchronize to Switch2. Switch1 believes that Switch2 has newer information than it. Using the set trace command, the router will generate a log message (see [Example 4-6](#))

### Example 4-5. Output of show vlan Command on Switch1

Switch1 (enable) show vlan

VLAN Name	Status	IfIndex	Mod/Ports, Vlans
-----			
1 default	active	5	1/1 2/1-2 6/1-48 10/7-48
2 vlan2	active	157	10/1-3
3 Vlan3	active	173	10/4-6

### Example 4-6. Debug of VTP on Switch1 Learning of a Higher Revision Number

VTP: i summary, domain = Cisco, rev = 4, followers = 1

VTP: domain Cisco, current rev = 1 found for summary pkt

VTP: summary packet rev 4 greater than domain Cisco rev 1

As a result, Switch1 loses VLANs 2 and 3, and any ports associated with those VLANs default back to VLAN 1. Remember, the highest revision number wins regardless of the mode of the switch. The output from [Example 4-7](#) shows all ports are once again associated with VLAN 1.

### Example 4-7. Output of show vlan Command on Switch1 After Synchronizing with Switch2

Switch1 (enable) show vlan

VLAN Name	Status	IfIndex	Mod/Ports, Vlans
-----			
1 default	active	5	1/1 2/1-2 6/1-48 10/1-48
1002 fddi-default	active	6	
1003 token-ring-default	active	9	
1004 fddinet-default	active	7	
1005 trnet-default	active	8	

## VTP Mode Best Practices

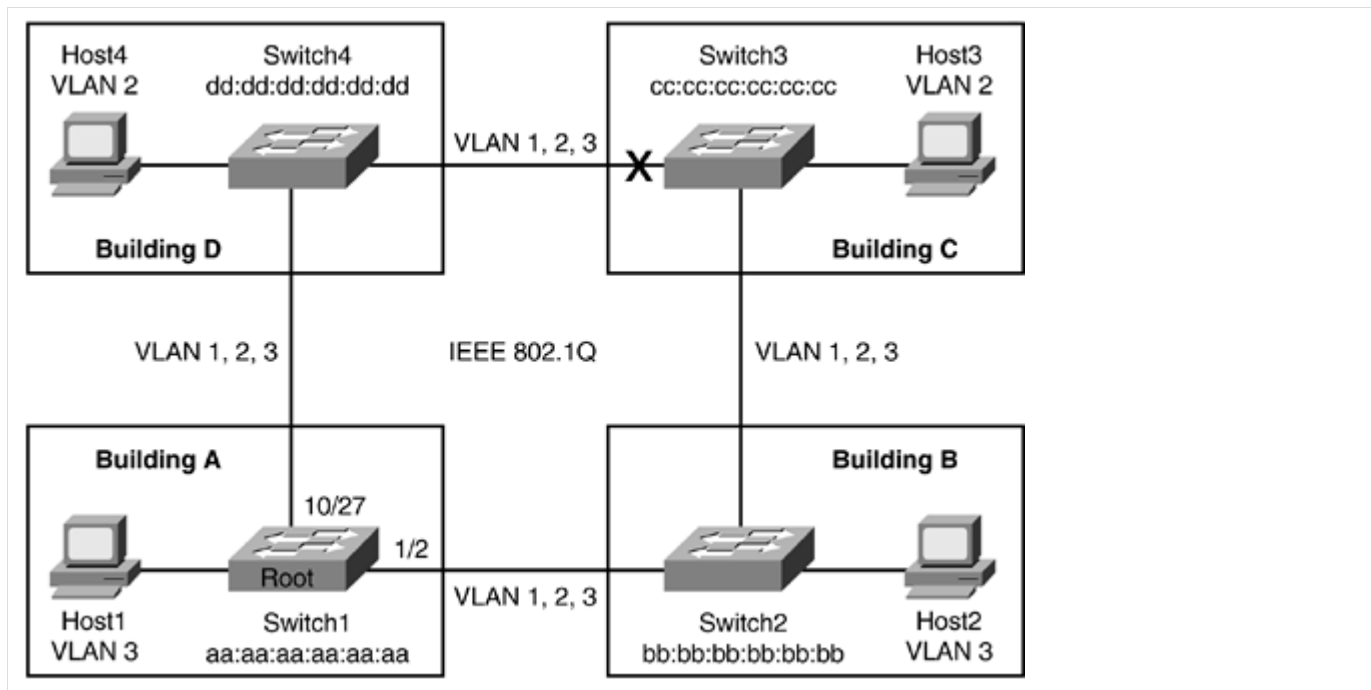
Some environments still deploy VTP server/client mode, while others stick with transparent mode. It is recommended that you configure VTP for transparent mode for a number of reasons, aside from the revision number issue. If an engineer accidentally erases a VLAN, the switch through the VTP mechanism will propagate that information to the rest of the domain. In addition, VTP server/client mode currently only supports VLANs 1-1024. The extended VLAN range, 1024-4096, requires the switch to be configured in transparent mode. The rule of thumb is keep everything simple, because in the long run it can save time and money.

Thus far, the discussion throughout this chapter has been on theory and design considerations. The following section introduces some rudimentary examples on configuring the aforementioned topics. Cisco provides quite a bit of information on its site about how to configure various protocols, features, and so forth. [Chapter 5](#), "Using Catalyst Software," is exclusively dedicated to providing configuration examples that will familiarize the reader on the more common configurations seen in the enterprise network.

## Configuring VTP/VLAN/Trunk

[Figure 4-8](#) illustrates multiple switches across the campus network. IEEE 802.1Q is used for trunking method. VLANs 2 and 3 are configured as well. VTP is used to propagate the VLAN information to the rest of the switches.

Figure 4-8. Configure Switch1 for VTP/VLAN/Trunk



The first thing to do is define the VTP domain, which allows VTP to manage switches that are under its domain. This command is programmed on all four switches:

```
| set vtp domain cisco
```

VTP domain names and passwords are case sensitive. A VTP domain name is only required for server/client mode. By default, Cisco switches will operate in server mode. For redundancy purposes, the recommendation is to have multiple VTP servers in the switch network. Therefore, only Switch2 and Switch3 will be configured for client mode. Switch1 and Switch4 will operate in server mode so that they may provide redundancy. Use the following command to change Switch2 and Switch3 to client mode:

```
| set vtp mode client
```

If a VTP domain name is not configured for server/client mode, or if the switch is not programmed to be in transparent mode, any VLANs created will be in inactive state.

The next step is to create VLANs 2 and 3. The set vlan command is used to create VLAN 3 on Switch1:

```
| set vlan 3
```

Host1 is connected to port 10/3 and is configured to be a member of VLAN 3:

```
| set vlan 3 10/3
```

By default, all ports are configured with VLAN 1. Hence, any trunks created will have their native VLAN as VLAN 1. Port 1/2 of Switch1 will be used to form a trunk with Switch2 so that Switch1 can pass VTP information to Switch2. The following command will force the native VLAN for trunk 1/2 to be VLAN 2:

```
| set vlan 2 1/2
```

Typically, it is recommended to assign native VLANs rather than using VLAN 1. The "[Understanding VLAN 1](#)" section of this chapter provides greater insight as to why this is necessary.

The port is configured to be an IEEE 802.1Q trunk port with the next command. Dot1q is the equivalent of IEEE 802.1Q on Cisco switches:

```
| set trunk 1/2 dot1q
```

Recall from earlier discussions of trunking that trunking mode must also be configured. The recommended method is to set trunking mode to desirable:

```
| set trunk 1/2 desirable
```

Port 1/2 will attempt to actively form a trunk with its directly connected link because it has the desirable parameter configured. Port 10/27 will similarly be configured to form a trunk to Switch4.

## VLAN Pruning

One of the major problems with extending Layer 2 architecture is excessive unwanted traffic on the network. A pruning method can be implemented on Cisco switches to prune VLANs from going to switches that do not have any hosts for that VLAN. It is important to note that although pruning can prevent some unnecessary traffic from being circulated across the network, pruning VLANs does not simplify the spanning-tree topologies.

By default, a trunk port allows all VLANs through the trunk as shown in [Example 4-8](#). Trunk 10/27 goes from Switch1 to Switch4.

Example 4-8. Output of show trunk Command Connected to Switch4

Switch1 (enable) show trunk 10/27

\* - indicates vtp domain mismatch

Port	Mode	Encapsulation	Status	Native vlan
10/27	auto	n-dot1q	trunking	1
Port	Vlans allowed on trunk			
10/27	1-1005,1025-4094			
Port	Vlans allowed and active in management domain			
10/27	1-3			
Port	Vlans in spanning tree forwarding state and not pruned			
10/27	1-3			

An example helps clarify this material. Looking back at [Figure 4-8](#), the trunk ports are permitting all VLAN traffic across the Layer 2 network. Host1 and Host2 are part of VLAN 3. Host3 and Host4 are in VLAN 2. Any broadcast, multicast, or unicast traffic generated by Host2 is received by all the switches. There is absolutely no need for Switch3 and Switch4 to receive these packets because these switches do not have any hosts that are part of VLAN 3. Switch3 and Switch4 will simply drop these packets upon receiving them. Therefore, this exercise is going to demonstrate how to filter or prune the unnecessary traffic from ever hitting Switch3 and Switch4.

Pruning VLAN 3 from Switch3 and Switch4 can happen in one of two ways. The first method discussed is VTP pruning, which is a dynamic process that VTP handles. The second method involved is manually pruning VLANs. Enterprise customers have used both methods to prune VLANs. However, manual pruning is preferred because VTP pruning requires VTP client/server mode operation.

VTP pruning is a global command and affects all the switches in the VTP domain. It only needs to be configured on one switch. All VLANs by default are prune eligible, which means that all VLANs can be pruned. To block specific VLANs from the pruning mechanism, use the clear vtp pruneeligible command. [Example 4-9](#) demonstrates how to configure Switch1 so that it does not forward VLAN 3 traffic to switches that do not have hosts that are part of VLAN 3.



#### Example 4-9. Enabling VTP Pruning on Switch1

Switch1 (enable) set vtp pruning enable

This command will enable the pruning function in the entire management domain.

All devices in the management domain should be pruning-capable before enabling.

Do you want to continue (y/n) [n]? y

VTP domain Cisco modified.

After turning pruning on, port 10/27, which is connected to Switch4, now only receives traffic from VLANs 1 and 2 from Switch1 as the output from [Example 4-10](#) shows. Furthermore, because VTP pruning is a global command, Switch1 sends only VLAN 1 and 3 traffic to Switch2.

#### Example 4-10. Output of show trunk Command to Switch4 After Pruning Is Enabled

Switch1 (enable) show trunk 10/27

\* - indicates vtp domain mismatch

Port	Mode	Encapsulation	Status	Native vlan
10/27	auto	n-dot1q	trunking	1

Port Vlans allowed on trunk

10/27 1-1005,1025-4094

Port Vlans allowed and active in management domain

10/27 1-3

Port Vlans in spanning tree forwarding state and not pruned

10/27 1-2

Now, consider a situation where Host4 on Switch4 is now part of VLAN 3. Switch4 will be forced to send VTP Join messages back to Switch1 for VLAN 3. As a result, Switch1 will once again start sending VLAN 3 traffic toward Switch4. VTP pruning is a dynamic process that allows or blocks VLAN traffic from the directly connected switches. VTP statistics can be gathered through the show vtp statistics command as shown in [Example 4-11](#). The command shows the number of VTP Join messages transmitted and received. The command can also be used for troubleshooting if any VTP pruning errors occur.

#### Example 4-11. Output of show vtp statistics Command on Switch2

Switch4 (enable) show vtp statistics VTP pruning statistics:

Trunk	Join Transmitted	Join Received	Summary advts received from GVRP PDU non-pruning-capable device	Received
10/27	777	780	0	0

!output omitted for brevity

The second method of pruning involves manually filtering VLANs from trunks. Manual pruning explicitly requires configuring the switch to filter specific VLANs on a trunk. In [Figure 4-8](#), Switch1 must clear the VLAN 3 off the trunk to prevent VLAN 3 traffic from hitting Switch4. In VTP pruning, trunks dynamically allow and prune VLANs based on VTP Join messages. In the manual process, this is not the case. Typically, manual pruning is configured on trunks that will not have any hosts associated with the filtered VLAN. Pruning also affects spanning-tree topology. Using the clear trunk command, manual pruning removes the VLAN from the spanning-tree topology on that switch. [Example 4-12](#) demonstrates removing VLAN 3 from trunk 10/27.

#### Example 4-12. Removing VLAN 3 from Trunk 10/27

Switch1 (enable) clear trunk 10/27 3

Removing Vlan(s) 3 from allowed list.

Port 10/27 allowed vlans modified to 1-2,4-1005,1025-4094.

The output from [Example 4-13](#) shows the changes after manually pruning VLAN 3. The only active VLANs now on port 10/27 are VLANs 1 and 3.

### Example 4-13. VLAN 3 Is Removed from Trunk 10/27

Switch1 (enable) show trunk 10/27

\* - indicates vtp domain mismatch

Port	Mode	Encapsulation	Status	Native vlan
10/27	desirable	dot1q	trunking	2

Port	Vlans allowed on trunk
------	------------------------

10/27	1-2,4-1005,1025-4094
-------	----------------------

Port	Vlans allowed and active in management domain
------	---

10/27	1-2
-------	-----

Port	Vlans in spanning tree forwarding state and not pruned
------	--

10/27	1-2
-------	-----

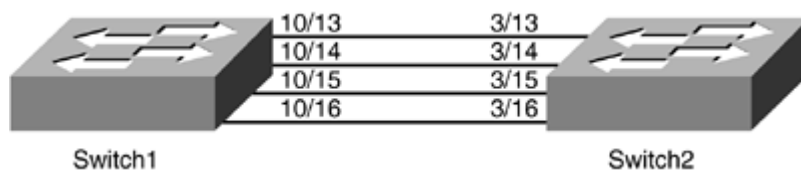
Traffic from VLANs 1, 2, and 3 are now going through a single connection, as shown earlier in [Figure 4-8](#). At some point, more bandwidth is needed to help deal with the volume of traffic passing through these switches. Assuming that altering the design of the network is not an option, you can either upgrade to a faster port such as Gigabit or bundle the existing ports into one, thereby, creating a bigger bandwidth connection.

## EtherChannel

Companies require greater and cheaper bandwidth to run their networks. Users are becoming more impatient with any sort of latency that occurs in the network. The insatiable appetite of customers for faster networks and higher availability of the networks has made the competition intense between vendors. A few years ago, Cisco came up with a method to not only provide substantially higher bandwidth but with lower cost overhead.

EtherChannel is a technology originally developed by Cisco Systems as a LAN switch-to-switch technique of inverse multiplexing of multiple Fast or Gigabit Ethernet switch ports into one logical channel. Its benefit is that it is effectively cheaper than higher-speed media while using existing switch ports, as shown in [Figure 4-9](#).

Figure 4-9. 4-Port EtherChannel



EtherChannel has developed into a cross-platform method of load balancing between servers, switches, and routers. EtherChannel can bond two, four, or eight ports (Catalyst 6500) to develop one logical connection with redundancy. The three major aspects to EtherChannel are

- Frame distribution
- Management of EtherChannel
- Logical port

EtherChannel does not do frame-by-frame forwarding on a round-robin fashion on each of the links. The load-balancing policy or frame distribution used is contingent upon the switch platform used. For instance, in a Catalyst 5500 switch platform, the load-balancing operation performs an X-OR calculation on the two lowest-order bits of the source and destination MAC address. An X-OR operation between a given pair of addresses will use the same link for all frames. One of the primary benefits of the X-OR operation is to prevent out-of-order frames on the downstream switch. The other advantage is redundancy. If the active channel used by a connection is lost, the existing traffic can traverse over another active link on that EtherChannel. The one

disadvantage to X-OR operation is the load on the channels might not be equal because the load-balancing policy is done on a specific header as defined by the platform or user configuration. On a Catalyst 6500 switch, the load-balancing operation can be performed on MAC address, IP address, or IP + TCP/UDP depending on the type of Supervisor/PFC used. Use the show port capabilities command to check the module for EtherChannel feature.

The default frame distribution behavior for the Catalyst 6500 is IP. [Example 4-14](#) is based off a Supervisor II/PFC2 card (WS-X6K-SUP2-2GE/WS-F6K-PFC2). This specific Supervisor can support the load-balancing policy up to Layer 4. Some of the older Catalyst 6500 Supervisors do not have this feature available. It is worth noting that most enterprise customers have deployed Catalyst 6500 Supervisor IIs in their networks. Load-balancing policies cannot be configured on Catalyst 5000, and the older Catalyst 4000 Supervisors. The newer Catalyst 4000s with Supervisor II-plus and higher do support load-balancing policies.

Example 4-14. Output of the set port channel Command

```
Switch1 (enable) set port channel all distribution ?
ip                Channel distribution ip
mac              Channel distribution mac
session          Channel distribution session
```

The management of the EtherChannel is done by Port Aggregation Protocol (PAgP). PAgP packets are sent every 30 seconds using multicast group MAC address 01-00-0C-CC-CC-CC with protocol value 0x0104. PAgP checks for configuration consistency and manages link additions and failures between two switches. It ensures that when EtherChannel is created that all ports have the same type of configuration. In EtherChannel, it is mandatory that all ports have the same speed, duplex setting, and VLAN information. Any port modification after the creation of the channel will also change all the other channel ports.

Finally, the last component of EtherChannel is creation of the logical port. The logical port, or Agport, is composed of all the links that make up the EtherChannel. The actual functionality and behavior of the Agport is no different than any other port. For instance, the spanning-tree algorithm treats Agport as a single port.

[Example 4-15](#) shows the recommended steps for configuring EtherChannel on a Catalyst 6500, as shown in [Figure 4-9](#). The ports used by Switch1 for EtherChannel are 10/13-16, and are configured with desirable mode. The desirable mode stipulates that a port actively initiate a channel setup.

Example 4-15. Output of the show port channel group Command

```
Switch1 (enable) show channel group

Admin Group  Ports
-----
272          10/1-4
273          10/5-8
274          10/9-12
275          10/13-16
276          10/17-20
277          10/21-24
278          10/25-28
279          10/29-32
280          10/33-36
281          10/37-40
282          10/41-44
283          10/45-48
```

The Admin Group defines the range of the ports that are going to be used by the EtherChannel. In [Example 4-15](#), the Admin Group is 275, which covers the ports that fall in the range of 13-16. Ports 10/9-12 are configured as a channel using Admin Group 274. If a new EtherChannel needs to be configured on ports 10/11-14, two separate channels would form because the ports are part of two different Admin Groups. The Admin Group for an EtherChannel needs to be the same. The Admin Group can be reassigned with the following command in [Example 4-16](#) to allow ports 10/11-14 to be in a single EtherChannel.

#### Example 4-16. Output of the set port channel Command

Switch1 (enable) set port channel 10/10-14 ?

<admin_group>	Admin group
mode	Channel mode

Example 4-17 shows how to configure ports 13-16 on module 10 for EtherChannel.

#### Example 4-17. Enabling a 4-Port EtherChannel

Switch1 (enable) set port channel 10/13-16 mode desirable

The show port channel command shown in Example 4-18 shows ports that are configured for channeling. The Admin Group is 275, and the Channel ID for the EtherChannel is 871.

#### Example 4-18. Viewing EtherChannel Configuration

Switch1 (enable) show port channel

Port	Status	Channel Mode	Admin Group	Ch Id
10/13	connected	desirable silent	275	871
10/14	connected	desirable silent	275	871
10/15	connected	desirable silent	275	871
10/16	connected	desirable silent	275	871

Port	Device-ID	Port-ID	Platform
10/13	TBA04081025(Switch#2)	3/13	WS-C6506
10/14	TBA04081025(Switch#2)	3/14	WS-C6506
10/15	TBA04081025(Switch#2)	3/15	WS-C6506
10/16	TBA04081025(Switch#2)	3/16	WS-C6506

Channel ID distinguishes between different EtherChannels on the switch. Silent/Non-Silent modes are involved with unidirectional link failures. It is possible for a fiber connection to remain up even if one of its transceivers has become faulty. Non-Silent mode detects faulty RX/TX transceivers on a fiber port. PAgP will reset the port for 1.6 seconds to force the other side to shut down as well. A unidirectional link can cause black holing of traffic because the return traffic is not received by the RX transceiver. The detection of unidirectional link by a PAgP is about 3.5 \* 30 seconds. On the other hand, UniDirectional Link Detection (UDLD) can detect the failure less than 50 seconds versus PAgP. UDLD will be discussed later in Chapter 10, "Implementing and Tuning Spanning Tree." In Silent mode, PAgP does not look for faulty transceivers. The recommendation is to leave Silent/Non-Silent modes at their default values because UDLD better addresses this problem. However, for devices that do not support UDLD, configure Non-Silent mode. Example 4-19 provides useful information about the channel that was created.

#### Example 4-19. Output from the show port channel information Command

Switch1 (enable) show port channel information

Switch Frame Distribution Method: ip both

Port	Status	Channel mode	Admin Channel group id	Channel Speed	Duplex	VLAN
10/13	connected	desirable silent	275	871 a-100 a-full		1
10/14	connected	desirable silent	275	871 a-100 a-full		1
10/15	connected	desirable silent	275	871 a-100 a-full		1
10/16	connected	desirable silent	275	871 a-100 a-full		1

Port	ifIndex	Oper-group	Neighbor Oper-group	Method	Oper-Distribution	PortSecurity/ Dynamic port
10/13	132	49	65	ip both		
10/14	132	49	65	ip both		
10/15	132	49	65	ip both		
10/16	132	49	65	ip both		

The key point here is the load-balancing policy used by the switch. According to [Example 4-19](#), the method used for load balancing is IP. To find out which IP address pairing is using a specific link on the EtherChannel, use the hidden command, as shown in [Example 4-20](#). The show bundle hash provides the same information in the newer codes.

#### Example 4-20. Output of show bundle hash Command

Switch1 (enable) show bndlhash 871 10.1.11.3 10.1.34.4

Selected port: 10/14

As noted, for source 10.1.11.3 to get to destination 10.1.34.4, it must use the 10/14 link of the EtherChannel.

The show channel traffic command shown in [Example 4-21](#) provides utilization information on each of the EtherChannel links.

#### Example 4-21. Output of the show channel traffic Command

Switch1 (enable) show channel traffic

ChanId	Port	Rx-Ucst	Tx-Ucst	Rx-Mcst	Tx-Mcst	Rx-Bcst	Tx-Bcst
869	10/11	26.08%	0.00%	20.10%	51.70%	0.00%	26.13%
869	10/12	17.39%	40.00%	19.57%	25.69%	0.00%	5.68%
869	10/13	30.43%	60.00%	40.21%	11.14%	100.00%	64.77%
869	10/14	26.10%	0.00%	20.12%	11.45%	0.00%	3.40%

For troubleshooting purposes, it is important to note if the switch is sending and receiving PAgP packets on the wire as revealed in [Example 4-22](#). This is one of the first commands that needs to be looked at to ensure that adjacent devices configured for EtherChannel support PAgP, and/or the devices are configured correctly.

#### Example 4-22. Output of the show port channel statistics Command

Switch1 (enable) show port channel statistics

Port	Admin	PAgP Pkts Group Transmitted	PAgP Pkts Received	PAgP Pkts InFlush	PAgP Pkts RetnFlush	PAgP Pkts OutFlush	PAgP Pkts InError
10/13	275	180	149	0	0	0	0
10/14	275	181	150	0	0	0	0
10/15	275	148	130	0	0	0	0
10/16	275	152	133	0	0	0	0

Example 4-23 is a hidden command on the Catalyst switch. The show agport command provides the assignment of the logical port, 14/39.

#### Example 4-23. Output of the show agport Command

Switch1 (enable) show agport

--- 14/39 ---

old\_mem\_cnt = 0; path\_cost = 8; path\_VLAN\_cost = 0  
trunk\_id = 870, time\_stamp = 233242506, agifindex = 132  
chnl\_list = 10/13-16  
agport\_list = 10/13-16  
bndlctrl: prtcnt = 4, num\_map = f0, dist\_req = 2, dist\_port = 0

mod	port	bndl_port	bndl_sel	bndl_sel*	act_flag	no_bits
10	13	0	c0	c0	2	2
10	14	1	30	30	2	2
10	15	2	0c	0c	2	2
10	16	3	03	03	2	2
0	0	0	00	00	0	0
0	0	0	00	00	0	0
0	0	0	00	00	0	0
0	0	0	00	00	0	0

The agport\_list parameter shows the active ports on the channel. For instance, if the 10/13 link is lost, the agport\_list will take the port out from the list. Notice that in Example 4-24, agport\_list does not have port 10/13 as member of the channel.

#### Example 4-24. Output of show agport with Link 10/13 Nonfunctional

Switch1 (enable) show agport

--- 14/39 ---

old\_mem\_cnt = 0; path\_cost = 8; path\_VLAN\_cost = 0  
trunk\_id = 870, time\_stamp = 233242506, agifindex = 132  
chnl\_list = 10/13-16  
agport\_list = 10/14-16  
bndlctrl: prtcnt = 4, num\_map = f0, dist\_req = 2, dist\_port = 0

mod	port	bndl_port	bndl_sel	bndl_sel*	act_flag	no_bits
10	13	0	00	00	0	0
10	14	1	30	30	2	2
10	15	2	8c	8c	2	3
10	16	3	43	43	2	3
0	0	0	00	00	0	0
0	0	0	00	00	0	0
0	0	0	00	00	0	0
0	0	0	00	00	0	0

In On mode, the agport\_list field is never adjusted because PAgP is disabled. Remember that PAgP is responsible for the addition and deletion of links on the channel.

From spanning tree's perspective, the EtherChannel is seen as a single port, as shown in [Example 4-25](#).

#### Example 4-25. Output of the show spantree Command Using an EtherChannel

Switch1 (enable) show spantree

VLAN 1

Spanning tree mode PVST+  
Spanning tree type ieee  
Spanning tree enabled

Designated Root 00-05-74-18-04-80  
Designated Root Priority 4097  
Designated Root Cost 0  
Designated Root Port 1/0  
Root Max Age 20 sec Hello Time 2 sec Forward Delay 15 sec

Bridge ID MAC ADDR 00-05-74-18-04-80  
Bridge ID Priority 4097 (bridge priority: 4096, sys ID ext: 1)  
Bridge Max Age 20 sec Hello Time 2 sec Forward Delay 15 sec

Port	VLAN	Port-State	Cost	Prio	Portfast	Channel_id
10/13-16	1	forwarding	8	32	disabled	871

The valid EtherChannel configurations are

- Desirable-Desirable
- Desirable-Auto



- On-On
- Off-Off

Cisco recommends Desirable-Desirable mode configuration for EtherChannel. This is beneficial because ports will actively negotiate setting up a channel and will allow the operation of PAGP. It is also recommended to leave Silent/Non-Silent parameters to their default values if UDLD is supported.

Table 4-3 describes the type of channel states that will develop depending on the configuration of the adjacent switches. Spanning tree shuts down (errdisable) channels that are misconfigured, as noted in Table 4-3.

Table 4-3. Channeling Modes Between Switches

Switch-A Channel Mode	Switch-B Channel Mode	Channel State
On	On	Channel
On	Off	Not Channel (errdisable)
On	Auto	Not Channel (errdisable)
On	Desirable	Not Channel (errdisable)
Off	On	Not Channel (errdisable)
Off	Off	Not Channel
Off	Auto	Not Channel
Off	Desirable	Not Channel
Auto	On	Not Channel (errdisable)
Auto	Off	Not Channel
Auto	Auto	Not Channel
Auto	Desirable	Channel
Desirable	On	Not Channel (errdisable)
Desirable	Off	Not Channel
Desirable	Auto	Channel
Desirable	Desirable	Channel

## Understanding VLAN 1

It is important to understand the significance of VLAN 1. By default, all switch ports are part of VLAN 1. VLAN 1 contains control plane traffic and can contain user traffic. It is recommended that user traffic be configured on VLANs other than VLAN 1, primarily to prevent unnecessary user broadcast and multicast traffic from being processed by the Network Management Processor (NMP) of the supervisor. Although VLAN 1 user traffic can be pruned from a trunk, it is not the case with control plane traffic. In fact, in older Cisco Catalyst Software versions (5.4 or earlier), VLAN 1 could not be removed at all from a trunk. Control plane traffic such as VTP, CDP, and PAGP protocols are tagged with VLAN 1 information and are forwarded on a trunk regardless if the trunk has pruned VLAN 1.

Management VLAN, discussed in the next section, is used to monitor and manage the switches on the network. This section also introduces best practices involving management VLAN.

### Management VLAN

The internal switch interface, sc0, is used for management of the switch. Management VLAN is used for purposes such as telnet, SNMP, and syslog. By default, VLAN 1 is the management VLAN. Ensure that there are no redundant links for the management VLAN. This practice eliminates the need to rely on the spanning-tree algorithm. This prevents the management VLAN from having any potential issues with spanning-tree loops. If the Layer 2 configuration does not provide an option to eliminate redundancy for the management VLAN, separate physical connections supporting only the management VLAN should be considered. Ensure that the management VLAN does not have any user traffic on it by only allowing switch management interfaces to be members of that VLAN.

### Management VLAN Best Practices

With the introduction of Cisco's powerful switches and VLAN feature, most companies started to deploy a switched network with VLANs extending throughout the LAN campus. Perhaps the strongest driving force to deploying a Layer 2 network was that Layer 3 devices could not keep up with Layer 2 switching engines. The phrase "the network is as fast as its slowest link" comes to mind. These days Layer 3 engines are no longer bottlenecks and can keep pace with Layer 2 engines. For example, the Catalyst 6500 is not the only Layer 3/Layer 2 switching device, but it has the most features and highest switching performance on the market today.

Perhaps the biggest issue with extending VLANs across multiple switches is spanning-tree loops. Spanning Tree Protocol (STP) is a loop-avoidance protocol designed to provide redundancy in a switch fabric network. Host3 will take the path via Switch2 to send traffic to the rest of the hosts on that segment that is not on Switch3 (see [Figure 4-8](#)). This works relatively well. If a failure occurs between Switch2 and Switch3, STP can bring up the redundant link, and traffic will be forwarded again after spanning tree converges.

Consider a situation where an STP loop occurs because of a bad transceiver that maintains a link but passes traffic unidirectional, or a hardware failure that results in missed STP BPDUs. This loop will degrade the performance on the switch network, users will have intermittent connectivity, and eventually, the network will be saturated. In a spanning-tree loop, an engineer at times has to console into the device because of slow telnet sessions because of excessive traffic on the network. Any time a VLAN is extended to various switches with redundant links, the network is vulnerable to such an event.

The other chronic issue with Layer 2 design is broadcast, multicast, and unicast flooding. A broadcast message is sent to MAC address FF-FF-FF-FF-FF-FF, which is received by all hosts on the VLAN. When Host1 sends an Address Resolution Protocol (ARP) for Host2, all other devices will also receive the broadcast message. In a huge network, with a great number of users and multiple switches involved, broadcast traffic can unnecessarily eat up bandwidth. Each device will look at the packet at Layer 3 to see if the packet belongs to it; if not, the packet is thrown away. The process of looking at the packet at Layer 3 requires CPU cycles, and as result, devices are functioning sub-optimally. Typically, ARP does not really cause that much trouble, but if in-house applications exist that communicate via broadcast, the application can adversely affect the network for the aforementioned reasons. If a broadcast storm occurs, it can and will bring the segment down completely. The real solution is to keep the segment small regardless whether the discussion is based on physical or logical segment. The rule of thumb is that the broadcast traffic should not be greater than 20 percent of the total traffic on the VLAN or segment.

To prevent excessive broadcasts on a segment, especially in a broadcast storm situation, Cisco switches have a mechanism to control the upper limits of broadcast traffic on a port. Cisco switches monitor the level of broadcast activity in 1-second intervals. They do this by looking at the individual/group destination address in the Ethernet frame. This value is compared with a predefined threshold set by the user. If the sample rate per second exceeds the threshold, the suppression mechanism is enabled, which filters broadcast packets on that port for a specified period of time. By default, broadcast suppression is disabled on Catalyst switches. In this example, the threshold is set to 50%, and anything higher will be dropped.

Broadcast suppression can easily be configured on the switch. According to the following command, any broadcast traffic that exceeds 50% on port 10/1 will be dropped:

```
set port broadcast 10/1 50%
```

The actual threshold value is contingent up the engineer's knowledge of the traffic on that segment. This feature does not allow broadcast storms to consume all available bandwidth and melt the network down.

Pruning can also be configured on the switches to reduce the diameter of the broadcast domain. Options are available to control the broadcast domain; these would require time and strategic planning to make Layer 2 somewhat resilient to a broadcast storm.

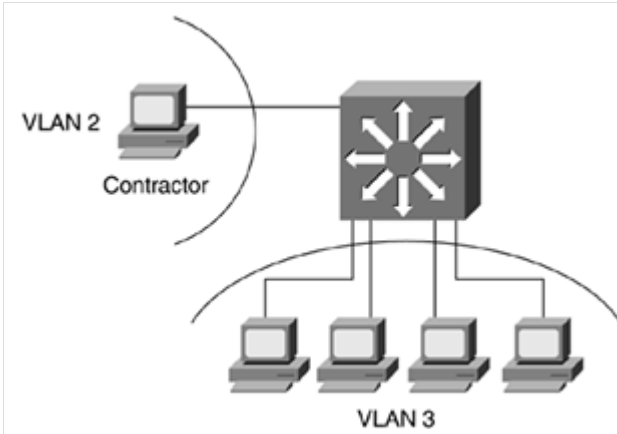
Some engineers believe that VLANs should never leave the box. In other words, keep Layer 2 small, which can help address issues with VTP client/server mode, and more importantly, spanning tree. The practice also means that trunking is not necessary. Essentially, the engineers want to push for a Layer 3 model. A Layer 3 design has many positive and negative attributes, but it does have one big advantage: Layer 3 does a better job of controlling the spread of the outage in the network.

Avoid extending VLANs, if at all possible. If this is not possible, keep the diameter of the Layer 2 switches small. Spanning tree recommends no more than seven switches between hosts. Avoid VTP client/server mode, and if pruning is required, use manual pruning.

## Private VLANs

Common VLAN implementation allows for any-to-any communication. Each host on the VLAN can communicate with any other host on that segment. Preventing communication between hosts on the same VLAN requires moving the users off the VLAN to their own separate VLANs. In the past, VLANs generally had a homogenous pool of users. The users in the VLAN had some type of commonality that allowed them to share the same resources and have the same access on the network. As a result, there was no need to filter traffic between users on the same segment. For instance, vendors or contractors who needed onsite access to the customer network were typically segregated in their own VLAN. Scalability was not an issue because the number of these groups was small and manageable (see [Figure 4-10](#)).

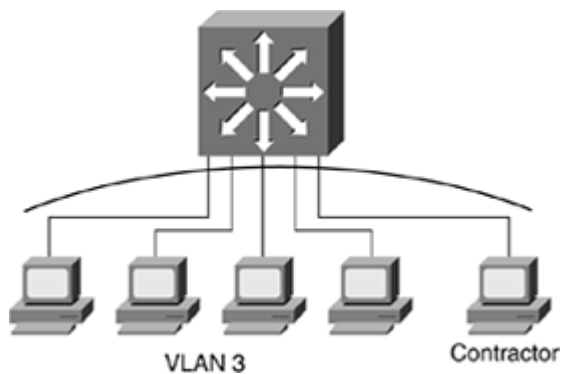
Figure 4-10. Contractors on a Separate VLAN



However, the numbers of these third-party groups have dramatically grown and are ubiquitous throughout the customer network, specifically government contractors. To isolate them in their own VLANs would require many IP addresses and VLANs. Figure 4-10 illustrates contractors in VLAN 2. VLAN 3 consists of company workers. Isolating contractors in their own VLANs is not practical and also would require some effort to maintain these VLANs. Private VLANs can help mitigate some of these issues. Private VLANs have the capability to restrict users on the same segment without the necessity of Layer 3 architecture. This translates to fewer IP address used and fewer new VLANs created on the network.

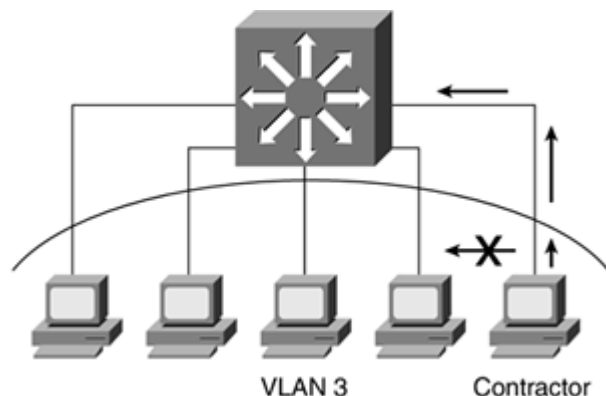
Figure 4-11 shows the contractors have been moved from VLAN 2 and are now members of VLAN 3.

Figure 4-11. Contractors on the Same VLAN



Private VLANs can also help protect hosts from each other on a segment (see Figure 4-12). Recently, corporate networks have been hit with various forms of worms. Typically, a worm infects a machine, and then it tries to form connections with other machines on the network through the infected machine. In a private VLAN environment, the infected machine has restrictions on it as to what ports it can communicate with. As a result, not all ports on that VLAN will be affected by this worm originated by the infected machine. Restrictions on a port might not give you all the protection in the world; however, as mentioned, restrictions can provide some benefits that cannot be overlooked.

Figure 4-12. Contractor Host Prevented from Communicating to Other Members on the Same VLAN



A private VLAN is an extension of the common VLAN to help restrict traffic from users on the same VLAN. It accomplishes this by assigning port designations, which include the following:

- Promiscuous
- Isolated
- Community
- Two-way community

All ports on the VLAN are assigned as part of the primary VLAN. Each port is also defined by a secondary VLAN.

The promiscuous port can communicate with any other host on the primary VLAN. It is usually the MSFC router port, Catalyst 6500 running Catalyst OS. The isolated port communicates only with the promiscuous port and no other host on the segment. It cannot communicate with other isolated ports. There can only be one isolated secondary VLAN in the primary VLAN. Community ports can communicate with the promiscuous port and other ports that are members of the community VLAN. Two different communities cannot communicate with each other.

Flows coming from the isolated or community ports are tagged internally on the switch with a secondary VLAN identifier. The identifier is used to forward the packet to the appropriate destination ports. For isolated ports, it is always going to be the promiscuous port, which internally tags all traffic destined to designated ports with primary VLAN information. If a Layer 2 access list is applied (VLAN Access List-**VACL**), the access list will only affect traffic coming from the secondary VLAN (isolated or community port) to the promiscuous port. The Layer 2 access list does not affect traffic going to the secondary VLAN because the promiscuous port will tag all traffic internally with a primary VLAN identifier. In other words, the promiscuous port cannot apply access lists going to a specific secondary VLAN. In a two-way community configuration, the router will remember the secondary VLAN information. As a result, it will be able to apply Layer 2 access list outbound to the secondary VLAN group.

Some caveats to private VLAN implementation include the following:

- VTP must be configured in transparent mode.
- Private VLANs can use VLANs 2-1000 and 1025-4096.
- Both primary and secondary VLANs can traverse a trunk, and will participate in spanning tree. Any modifications to the primary VLAN spanning tree will affect the secondary VLAN spanning-tree algorithm.
- Private VLANs cannot be configured on trunk, dynamic VLAN, or channel ports. By default, the configuration will automatically disable trunking on the port.
- BPDU guard gets enabled. The BPDU guard protects against a portfast enabled port sending BPDU messages.
- Internet Group Management Protocol (IGMP) snooping is not supported on the private VLANs.

The primary VLAN is 3 with two secondary VLANs, 13 and 15. These secondary VLANs have been associated with a primary VLAN with specific ports defined in [Example 4-26](#). Then the primary and its secondary VLANs have been mapped to the promiscuous port, 15/1.

### Example 4-26. Configuring Private VLANs

```
Switch1 (enable) set vlan 3 pvlan-type primary
Switch1 (enable) set vlan 13 pvlan-type isolated
Switch1 (enable) set vlan 15 pvlan-type twoway-community
Switch1 (enable) set pvlan 3 13 10/1

Vlan 13 configuration successful
Ports 10/1-12 trunk mode set to off.

Successfully set the following ports to Private Vlan 3,13:
10/1

Switch1 (enable) set pvlan 3 15 10/2-3

Vlan 15 configuration successful
Ports 10/1-12 trunk mode set to off.
Ports 10/1-12 trunk mode set to off.

Successfully set the following ports to Private Vlan 3,15:
10/2-3

Switch1 (enable) show pvlan

Primary Secondary Secondary-Type  Ports
-----
3      13      isolated      10/1
3      15      twoway-community 10/2-3

Switch1 (enable) set pvlan mapping 3 13 15/1

Successfully set mapping between 3 and 13 on 15/1

Switch1 (enable) set pvlan mapping 3 15 15/1

Successfully set mapping between 3 and 15 on 15/1

Switch1 (enable) show pvlan mapping

Port Primary Secondary
-----
15/1 3      13,15
```

Referring to [Figure 4-12](#), the contractor now cannot communicate with Hosts1, 2, and 3. The other hosts can communicate with each other, but not with the contractor. All hosts including the contractor can communicate with the promiscuous port. Any broadcast or unicast floods generated by any of these hosts including contractor will be contained in their secondary VLAN environment. If contractor sends a broadcast message, the other hosts will not receive the broadcast message.

Private VLANs are relatively new in the enterprise network. Private VLAN offers many features, and it will become popular in the near future, especially in parts of the network where devices need to be protected from other users and possible network attacks.

## Spanning Tree Protocol

- Root Bridge or Switch Port
  - Bridge Protocol Data Units (BPDUs)
- Spanning Tree Protocol Configuration
  - Spanning-Tree Port States
    - Blocking
    - Listening
    - Learning
    - Forwarding
    - Disabled
  - Spanning-Tree Operation
    - Designated Ports
    - Convergence

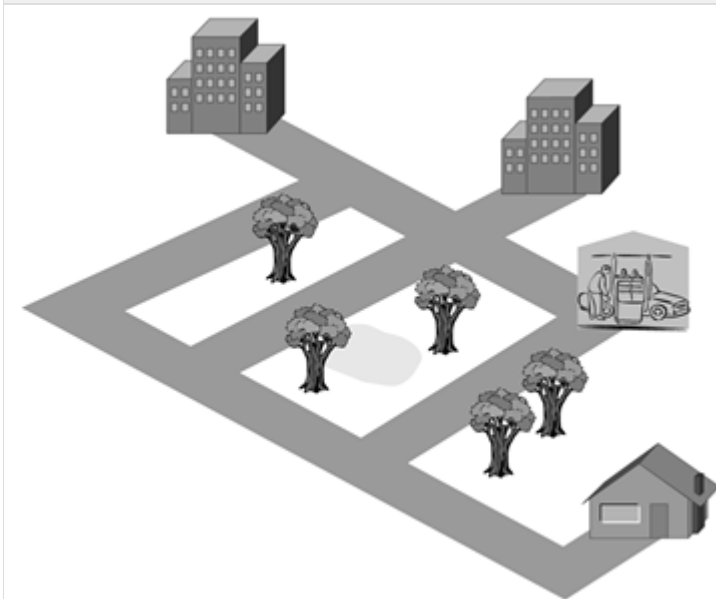
Recall that in the networking world, a protocol is a standard set of rules and formats for data transmission between computers, similar to the rules of grammar in the English language. If two people put commas and periods in different places and use them for different purposes, for example, communication between the two would be difficult, if not impossible. In this same way, communication is impossible if two computers use different protocols when trying to communicate with one another. This chapter explains the protocol, or grammar, of communication between switches - specifically the Spanning Tree Protocol (STP).

The Spanning Tree Protocol, or STP, is a link-management protocol that is part of the Institute of Electrical and Electronics Engineers (IEEE) 802.1 standard for bridges and switches. STP uses the spanning-tree algorithm and provides path redundancy in the network. STP also prevents network loops created by multiple active paths between stations.

note: A bridge loop occurs when two or more paths exist between network segments.

The spanning-tree algorithm is used in bridge-and switch-based networks and determines the best path for traffic to move across the network from source to destination. The algorithm creates a hierarchical tree spanning the entire network, including all bridges and switches. The spanning-tree algorithm determines all redundant paths and makes only one of them active at any given time, much as you might consider alternative routes from your home to your office before taking one.

Figure 7-1. Multiple Routes Between Home and Work



If you tried one morning using more than one route from your home to work, you could end up going around in circles and never getting to work. In a network, loops create broadcast storms and constant table changes, which cause damage to your network because your data will time out before it ever reaches its intended destination.

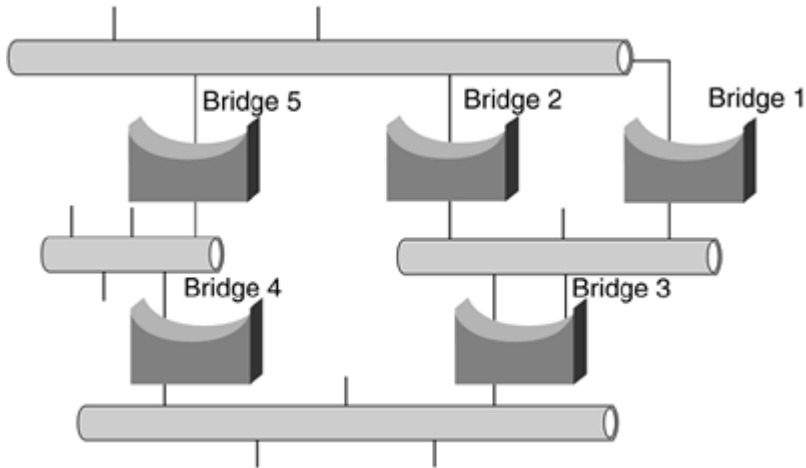
Loops occur when more than one route, or path, exists between nodes in a network. Establishing path redundancy, STP creates a tree spanning across all the switches in an extended network and forces redundant paths into a standby, or blocked, state. STP allows only one active path at a time between any two network devices, preventing loops, but establishes redundant links as a backup if the primary link fails. If a network segment becomes unreachable for whatever reason, the spanning-tree algorithm reconfigures the logical topology, reestablishing the link by activating the standby path. Without a spanning tree in place, it is possible that both connections might be considered the primary path, resulting in an endless loop of traffic on the local-area network (LAN).

## Root Bridge or Switch Port

STP has a root and branches like a tree. The primary decision-making switch in an STP environment is called the root bridge.

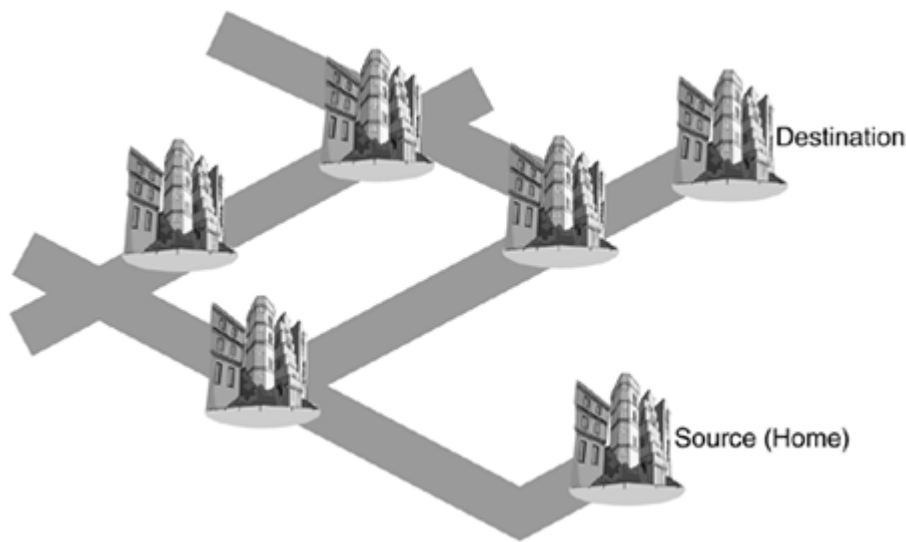
The network flows out from the root bridge to form a logical branched network. All switches in a LAN participating in STP branch from the root switch port or bridge. In [Figure 7-2](#), Bridge 1 is the root bridge, and all connected segments branch from this root.

Figure 7-2. Five-Bridge Network with Root Bridge (Bridge 1)



The first task of the STP is determining where the spanning tree begins - the root bridge or switch port. The root bridge is used to build a reference point in the network so that the spanning-tree algorithm can be calculated. All paths from all bridges and switches must be traceable back to the root bridge or root switch, much as all roads lead to or from your hometown, regardless of how many other towns you travel through to reach your destination, as illustrated in the [Figure 7-3](#).

Figure 7-3. Multiple Routes Between the Source and Destination



The root switch is elected as part of the STP and is necessary to build a reference point for the spanning-tree algorithm calculations. All paths not needed to reach the root switch network are placed in backup mode. Each switch in the network gathers information about other switches in the same network through an exchange of data messages called bridge protocol data units, or BPDUs.

### Bridge Protocol Data Units (BPDUs)

Bridge protocol data units, or BPDUs, are data messages exchanged between the switches and bridges within an extended LAN using the STP. BPDU frames contain information regarding the originating switch port, Media Access Control (MAC) address, switch port priority, and the switch port cost. The cost of a switch port is based on the number of network segments the frame crosses before reaching its destination.

BPDU messages are also exchanged across bridges and switches to detect loops in the network topology. Any loops found are removed by shutting down the selected bridge and switch interfaces and placing the redundant switch ports in a backup, or blocked, state.

The topology of a switched LAN is determined by the following:

- The unique switch identifier or bridge ID associated with each switch - The bridge ID is made up of the MAC address and the bridge priority. As your home telephone number is unique so that you can be reached by anyone looking for you, so must the identifier of each switch be unique so that it can be found in the network. note The MAC address is the 48-byte hardware address of the network interface.
- The port cost associated with each switch port - The port cost is for communication between the switch port and the root port. This is true whether it is a financial cost, as in your long-distance telephone calls, or a logical cost, as in how fast (maximum bandwidth) each network segment is that the frame must cross on its way from source to destination.

One BPDUs is superior to another if it has a lower

- Root bridge ID
- Path cost to the root
- Sending bridge ID
- Sending port ID

Each switch originates, but does not forward, configuration BPDUs that are used to compute the spanning-tree topology. The BPDU frame is sent across the LAN, and all connected bridges and switches receive this BPDU. The receiving switch uses the information in the BPDU to determine changes in the network topology. If there is a change, the receiving switch sends a new BPDU across all attached network segments.

BPDUs contain information about the sending switch and its ports, including the following:

- Switch and port MAC address - This is the MAC address of each switch and bridge port that is part of the tree.
- Switch and port priority - When switches and bridges are running the STP, each has a bridge or switch port priority associated with it. By default, all STP switches are configured with a bridge priority value of 32,768. After the exchange of BPDUs, the switch with the lowest priority value becomes the root bridge.
- Port cost - Cost is determined according to the speeds that the ports support; the faster the port, the lower the port cost. Switches use port costs in determining the root port for each and every switch.

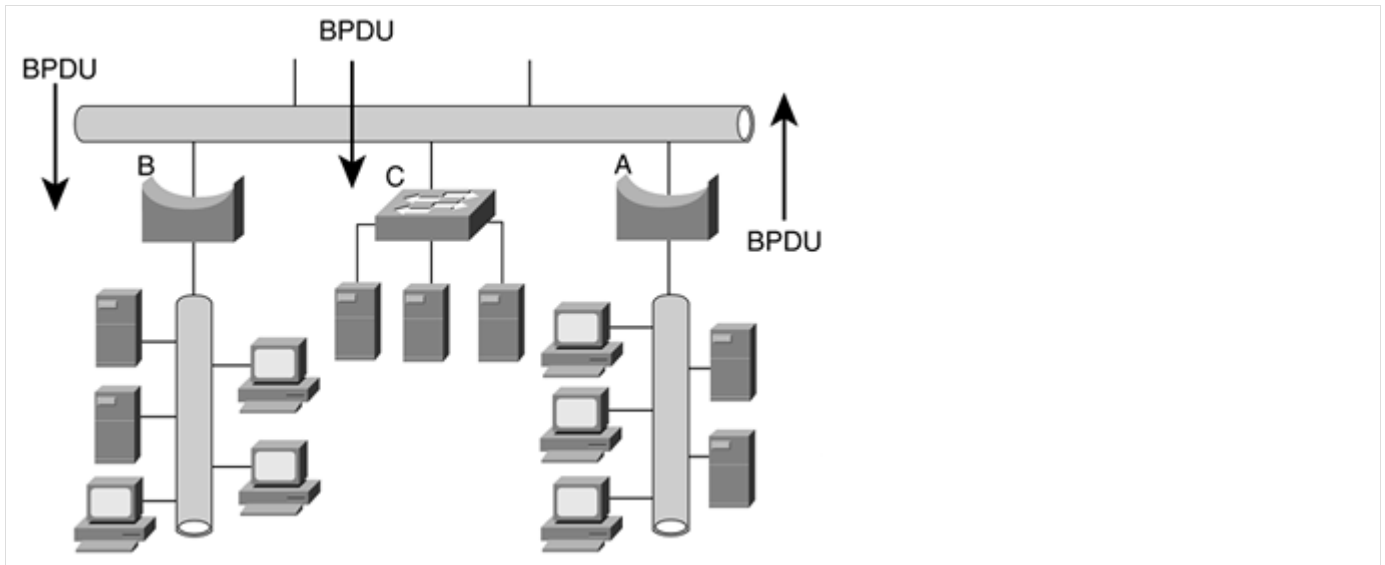
The exchange of BPDUs results in the following:

- One bridge or switch port is elected as the root bridge/switch port. This election is similar to a bunch of switches going to a voting booth and choosing their favorite switch. The BPDUs are used as a voter information guide, or ballot, to select the correct candidate. The purpose of this election is to determine which switch has the lowest identifier.
- The shortest distance to the root switch is calculated for each switch. Recall that the shortest distance between two points is a straight line, and the exchange of these BPDUs determines the direction of the straight line between bridge/switch ports.
- A designated switch is selected. This designated switch is the closest switch to the root switch through which frames will be forwarded to the root. There is only one designated switch per segment or VLAN.
- A designated port for each switch is selected, providing the best path to the root switch. Every LAN segment needs to know which switch is its entry/exit point to the rest of the network; otherwise, frames would wander aimlessly around the same network segment, never getting anywhere.
- Ports included in the STP are selected. Because all ports might not be part of the spanning tree, the exchange of BPDUs determines which ports have an invitation to the spanning tree (forwarding) and which ports do not. (Those that don't are disabled.) If STP is not running on some ports or switches, loops can occur on those non-STP ports, which then circumvent the STP blocks. note All bridge/switch ports are included in the STP BPDU message. The ports not turned on are not included as part of the spanning tree.
- Loops in the switched network are removed. Loops are detriments to networks because traffic on a network containing loops goes around in circles - stuck on the proverbial hamster wheel - and can shut down the network. These network loops are prevented by each switch placing redundant switch ports in a backup state as directed by the STP.

The following figure illustrates how BPDUs enable a spanning-tree topology based on the STP. Bridge A sends out a BPDU across the network that is received by Bridge B and Switch C.

Figure 7-4. BPDU Exchange Establishing the Spanning-Tree Topology





## Spanning Tree Protocol Configuration

Before you can understand how a network topology is built and managed using the STP, you need to understand the five states of the spanning tree. If geography isn't your strong suit, that's okay; there are no maps involved when discussing these states.

### Spanning-Tree Port States

Because of network delay caused by large LAN segments, topology changes can take place at different times and at different places in the switched network. When a switch port transitions directly from nonparticipation to an active, or forwarding, state, temporary data loops can be created. Ports must wait for new topology information to spread throughout the LAN before frames can be forwarded. Switches must also allow the frame lifetime to expire for frames that have been forwarded using the old topology.

Each port on a switch using STP is in one of the following five states:

- Blocking
- Listening
- Learning
- Forwarding
- Disabled

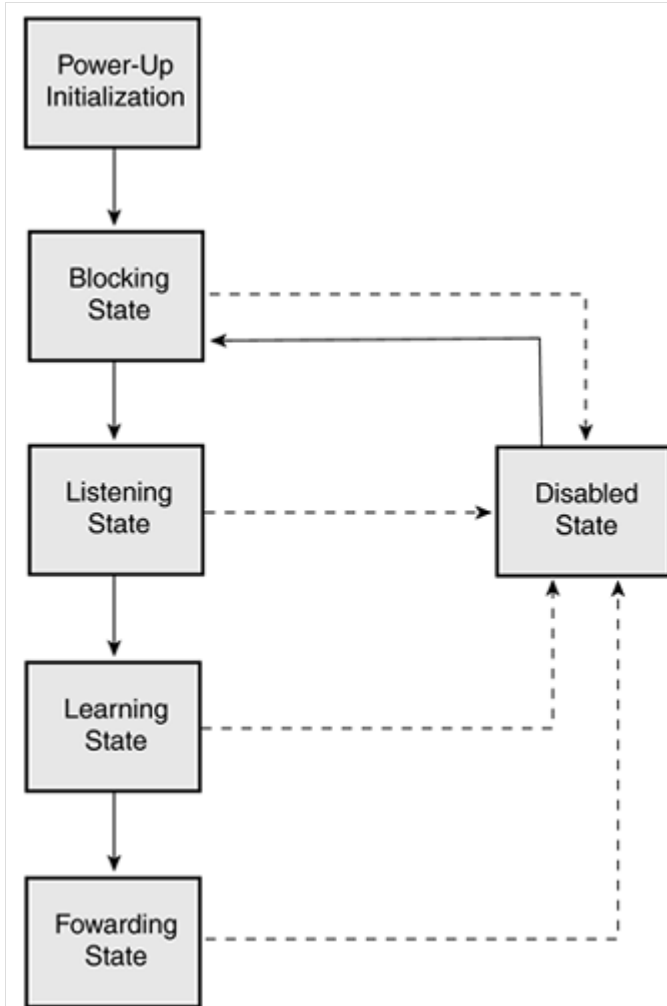
Each of these states is discussed in more detail in the following sections. A switch does not enter any of these states immediately, except the blocking state, which is entered on power up. Spanning-tree switch ports move through these five states in the timeframe described as follows:

- Initialization to blocking (0 seconds)
- Blocking to listening (20 seconds)
- Listening to learning (15 seconds)
- Learning to forwarding (15 seconds)
- Disabled

note The network administrator can disable a switch port at any time.

Figure 7-5 illustrates a bridge or switch port moving through the five STP states.

Figure 7-5. Spanning Tree Protocol States



When the STP is enabled, every bridge and switch in the network starts in the blocking state and transitions to the listening and learning states. If properly configured, the ports then stabilize to the forwarding or blocking state until a change in the network is made.

When the spanning-tree algorithm determines that a port is to be in the forwarding state, the following happens:

- The port is put into the listening state while waiting for protocol information suggesting it should go to the blocking state.
- The port waits for the expiration of a protocol, or forward delay, timer that moves the port to the learning state.
- In the learning state, the port continues to block frame forwarding as it learns network host location information for the forwarding database.
- The expiration of a protocol (forward delay) timer moves the port to the forwarding state. Both learning and forwarding are enabled while the port is in the forwarding state.

## Blocking

A port in the blocking state does not participate in frame forwarding, and after initialization, a BPDU is sent to each port in the switch. A switch assumes it is the root until it exchanges BPDUs with other switches in the network. This BPDU exchange establishes which switch in the network is the root switch. If only one switch resides in the network, no exchange occurs, and after the forward delay timer expires, the ports move to the listening state.

note A switch always enters the blocking state following switch initialization.

A port in the blocking state

- Discards frames received from the attached network segment.
- Discards frames switched from another port for forwarding.
- Does not incorporate a host location into its address database; because there is no learning at this point, there is no address database to update.
- Receives BPDUs from the network segment and directs them to the switch system module for processing.
- Unlike ports in the listening, learning, and forwarding state, a port in the blocking state does not process BPDUs received from the switch system module.

- Receives and responds to network management messages, such as a network administrator disabling the port.

After 20 seconds, the switch port moves from the blocking state to the listening state.

### Listening

The listening state is the first transitional state for a port after the blocking state. The listening state is where the STP determines that the port should participate in frame forwarding. The switch does not perform any learning or forwarding functions while in the listening state, and it therefore does not incorporate station locations into its address database as it would if the switch were in a blocking state, because there is no address table to update (while in a blocking state). In the listening state, a switch performs the following functions:

- Discards frames received from the attached network segment.
- Discards frames switched from another port for forwarding.
- Receives BPDUs from the network segment and directs them to the switch system module for processing.
- Processes BPDUs received from the switch system module.
- Receives and responds to network management messages, such as a network administrator disabling the port.

After 15 seconds, the switch port moves from the listening state to the learning state.

### Learning

In the learning state, the switch port prepares to participate in the network by forwarding frames. Learning is the second transitional state through which a port moves toward the end goal: frame forwarding. It is the STP that moves the port from the listening to the learning state.

A port in the learning state

- Discards frames received from the attached network segment.
- Discards frames switched from another port for forwarding.
- Incorporates LAN host location information into its address database.
- Receives BPDUs from the network segment and directs them to the switch system module for processing.
- Receives, processes, and transmits BPDUs received from the system module.
- Receives and responds to network management messages, such as a network administrator disabling the port.

After 15 seconds, the switch port moves from the learning state to the forwarding state.

### Forwarding

A port in the forwarding state forwards frames across the attached network segment. The forwarding state is the last state a port enters during the creation of the network topology.

A port in the forwarding state

- Forwards frames received from the attached network segment.
- Forwards frames switched from another port for forwarding.
- Incorporates LAN host location information into its address database.
- Receives BPDUs from the network segment and directs them to the switch system module for processing.
- Processes BPDUs received from the switch system module.
- Receives and responds to network management messages, such as a network administrator disabling the port.

A port stays in the forwarding state until a change occurs in the network topology, such as the addition of a new bridge or switch, a new bridge or switch port, or the failure of a bridge, switch, or port. When a change in the topology is detected, all switches recompute the network topology; this process is called [convergence](#).

### Disabled

A port in the disabled state does not participate in frame forwarding or the operation of STP because a port in the disabled state is considered nonoperational.

A disabled port

- Discards frames received from the attached network segment.
- Discards frames switched from another port for forwarding.
- Does not incorporate LAN host location information into its address database.
- Receives BPDUs, but does not direct them to the switch system module.
- Does not receive BPDUs for transmission from the switch system module.
- Receives and responds to network management messages, such as notification of a network administrator enabling a port.

## Spanning-Tree Operation

Just as a spanning-tree switch has a value, so do the individual ports on the switch, called the **port cost**. As discussed earlier, the port cost is determined based on the network bandwidth, or speeds that the port supports; the faster the port, the lower its cost.

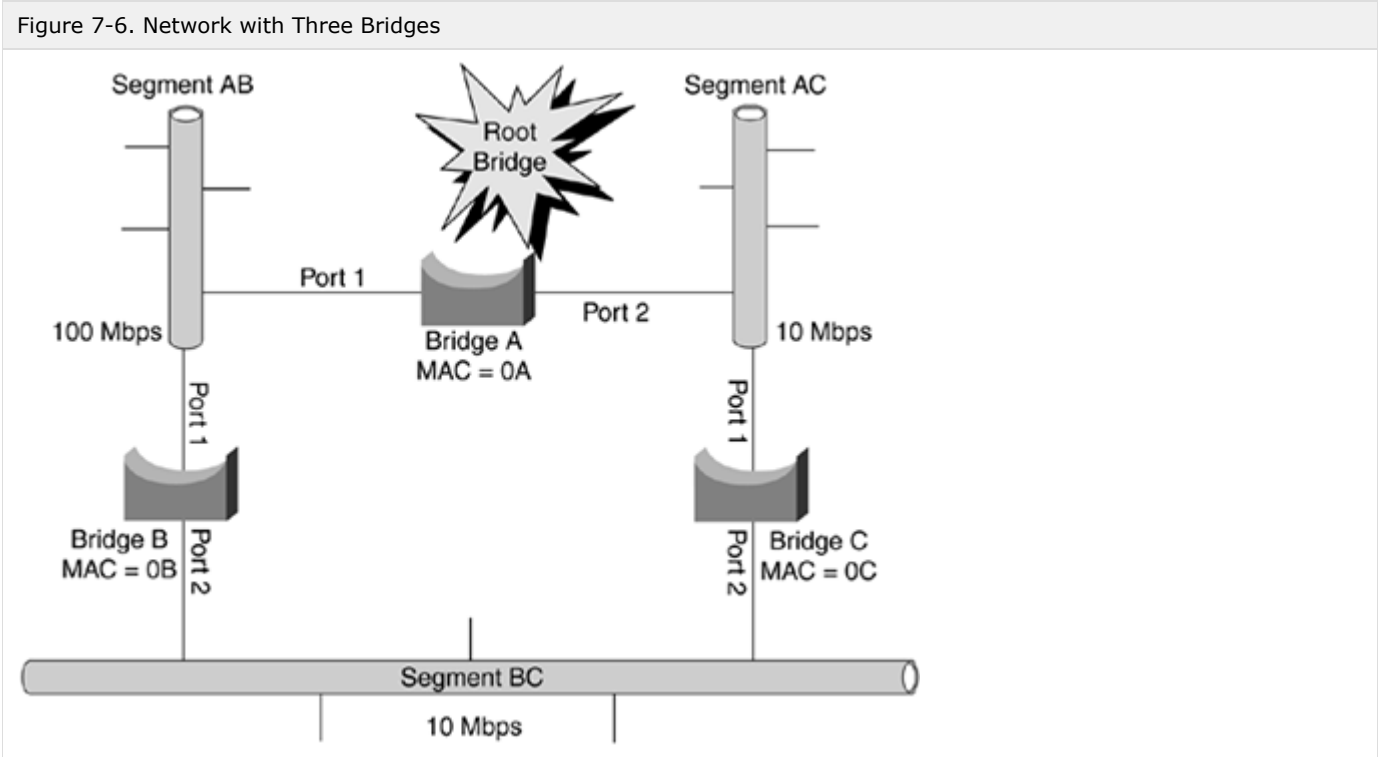
Table 7-1 lists the default IEEE costs associated with common port speeds.

Table 7-1. Default Port Cost	
Link Speed	Port Cost
Gigabit Ethernet	4
Fast Ethernet (100 megabits per second Mbps)	10
Ethernet (10 Mbps)	100

ch07fig06ch07fig06ch07fig06ch07fig06ch07fig06A switch uses the port cost to determine the root port for each switch in the network. All nonroot bridges have one root port that is used as the link over which data traffic is forwarded across the network.

note The root port represents a switch's lowest-cost path to the root bridge, and, by default, all ports on the root bridge are also root ports and have a cost of 0. Because root ports are directly connected to the root bridge, their cost to reach the root bridge is 0.

Figure 7-6 shows a network with three bridges. Bridge A has been made root bridge 7 because it has the lowest MAC address; because all bridge priorities are equal, the bridge with the lowest MAC address is elected the root.



The following three items characterize the network topology shown in Figure 7-6:

- Bridge B is connected to Bridge A via a 100-Mbps link, and Bridge C is connected to Bridge A via a 10-Mbps link.
- Bridges B and C are connected to one another at 10 Mbps via Segment BC.
- Segment BC creates a loop in this network.

Because this network has a loop, the STP determines which links remain in a forwarding mode and which enter a blocking mode.

Bridge A is elected as the root bridge because it has the lowest MAC address based on the STP information exchanged by the BPDUs between bridges in this network. In this case, the root bridge sends out BPDUs with a port cost of 0; and because it is the root bridge, there is no cost for its own ports to reach it. Therefore, the port cost is 0. These BPDUs will be received on port 1 on Bridge B and Bridge C.

When these BPDUs are received by Bridge B, it (Bridge B) adds its own port cost to the cost provided by the root bridge; because the cost associated with a 100-Mbps port is 19, Bridge B port 1 determines that it can reach the root bridge with a

total cost of 19. Port 1 of Bridge C, connected at 10 Mbps, determines that it can reach the root bridge with a total cost of 100 (100 + 0).

note By default, BPDUs are sent across the network every two seconds.

Remember Bridge B and Bridge C are connected to Network 1 and also send out BPDUs on their interface connected to this network - port 2 for both bridges. Bridge B sends a BPDU to Bridge C over this network segment (Segment BC). In this BPDU, Bridge B announces to Bridge C that it can reach the root bridge with a cost of 19. When this message reaches Bridge C, it adds its port 2 cost to this value, calculating that it can reach the root bridge with a total cost of 119 (100 + 19) via port 2.

Bridge C now knows that it can reach the root bridge through port 1 with a cost of 100, or through port 2 with a cost of 119. Based on these two paths, Bridge C determines that port 1 should be its root port because of its lower cost to the root.

Bridge C also sends BPDUs to Bridge B across Segment BC. In these BPDU messages, Bridge C announces a cost to the root bridge of 100. When these BPDUs are received by Bridge B, Bridge B adds this cost to the cost of its port 2 interface. Bridge B now also knows that it can reach the root bridge, via Bridge C, with a total cost of 200. Based on the two possible paths, Bridge B determines that port 1 should be its root port because of its lower-cost path to the root.

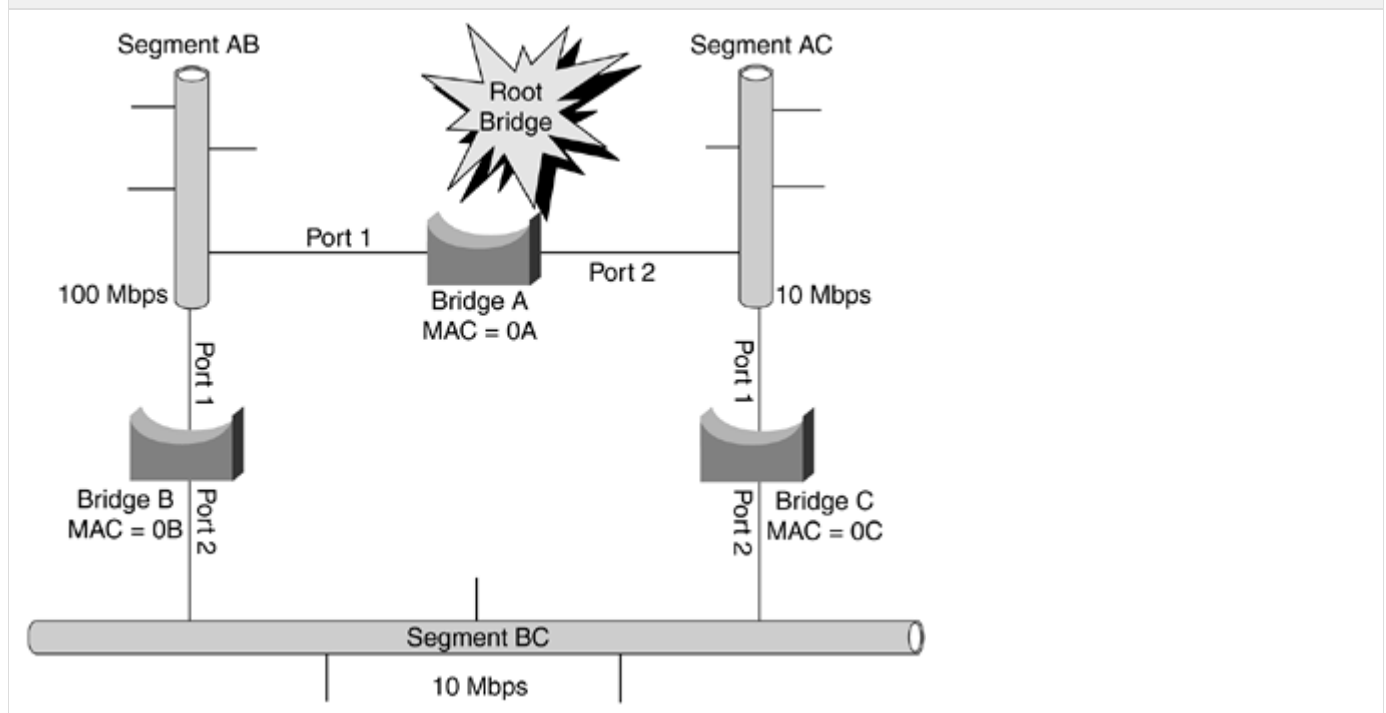
Remember, the shortest distance between two points is a straight line, or in the case of STP, the lower cost.

### Designated Ports

In the small network described previously, you have determined which port(s) should be the root ports on network bridges; however, which ports will be in a blocking or forwarding mode must still be determined.

For example, Segment BC has two possible paths to the root bridge: one via port 2 on Bridge B and the other via port 2 on Bridge C. To eliminate this loop, one of these two ports must be placed in a blocking mode, as illustrated in [Figure 7-7](#).

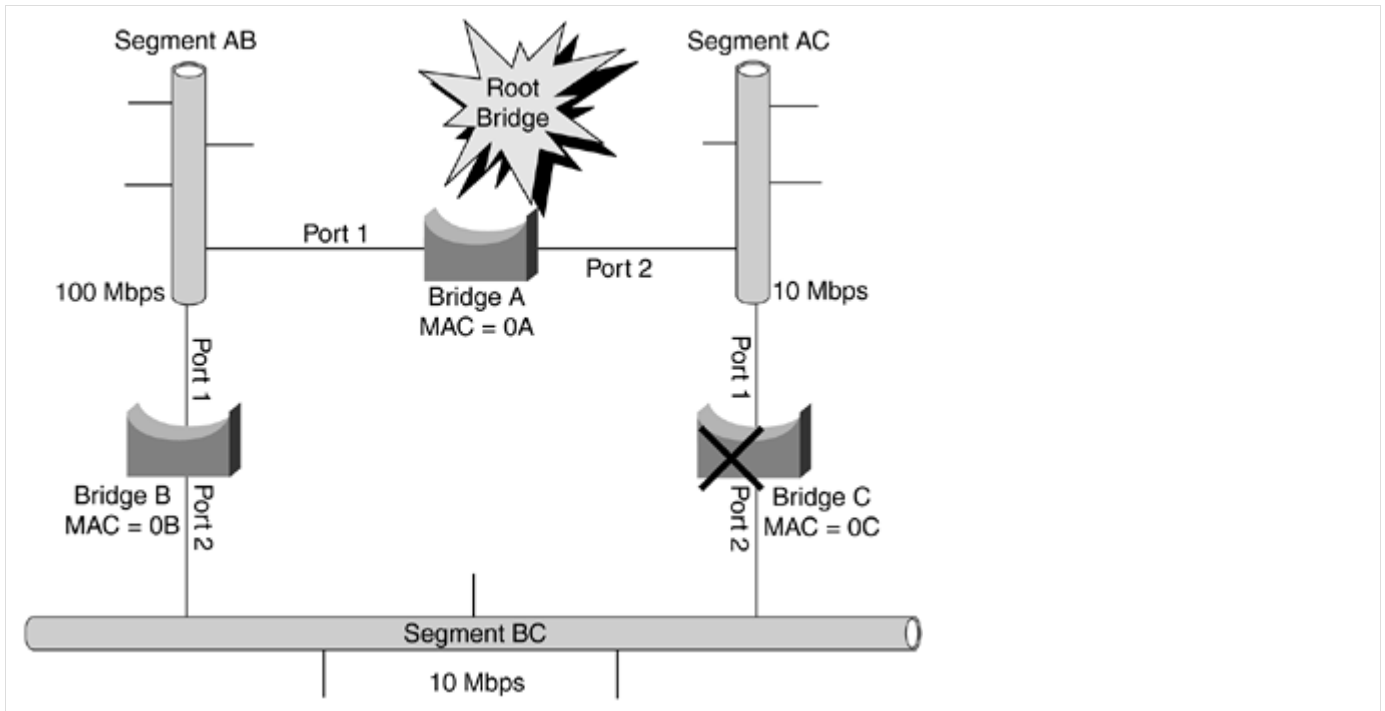
Figure 7-7. Bridge C with a Blocked Port



On a spanning-tree network, each network segment has one port identified as the designated port. The designated port is the port that is the single interface to forward traffic to the root bridge, and is determined via another election using BPDUs.

The network illustrated in [Figure 7-7](#) contains three segments: Segment AC, Segment AB, and Segment BC. On each segment, one of the connected bridge ports needs to be elected as the designated port. This is always the switch port on the segment with the lower port cost. For example, on Segment BC, two paths via port 2 on Bridge B and Bridge C are available to the root bridge, forming a loop. In this case, port 2 on Bridge B and Bridge C has a port cost of 100 on Segment BC, as illustrated in [Figure 7-8](#).

Figure 7-8. Traffic Path from Segment BC to Segment AC

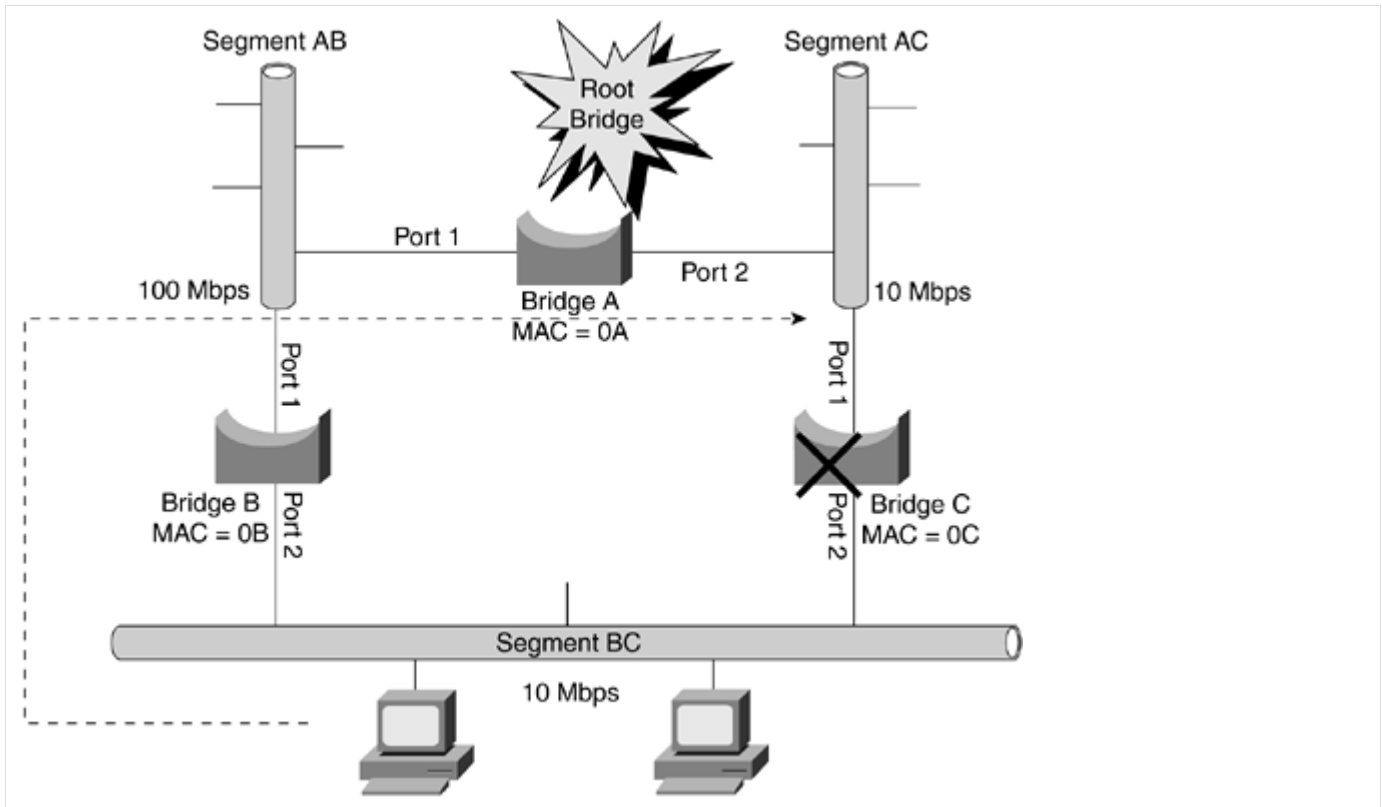


Because both bridges, Bridge B and Bridge C, have equal port costs to each other, MAC addresses are used to determine the **designated port**, making Bridge B the designated port on Segment BC because it has the lower MAC address. Therefore, port 2 on Bridge B will be placed in forwarding mode, and port 2 on Bridge C in blocking mode. When these forwarding and blocking modes are established, all traffic from Segment BC will exit the segment via Bridge B.

#### Convergence

After the transfer of BPDUs between systems has determined the root bridge and the root port of each bridge and switch, the network is loop free. The next topic is how the STP functions when something goes wrong in the network, such as a link failure. After the STP topology of a network has been calculated, each bridge and switch forwards BPDUs every two seconds. These BPDU messages inform the bridges and switches of which links are still active in the network, and which bridges and switches are not. For example, Bridge B in the network example illustrated in the [Figure 7-9](#) could have failed or been powered down.

Figure 7-9. Bridge B Failure



In this case, Bridge C fails to receive BPDUs from Bridge B on Bridge C's port 2 interface. Even though Bridge C port 2 is in blocking mode, it continues receiving and analyzing BPDU messages. After 20 seconds have passed without Bridge C receiving a BPDU on port 2 from Bridge B, Bridge C assumes that Bridge B is not available and transitions into the listening state. The listening state lasts for 15 seconds and is the time when Bridge B will be listening to and inspecting BPDUs from all other bridges. The bridge port still does not forward traffic during the listening stage.

After the 15 seconds of the listening state expire, the Bridge C port transitions into a learning state for another 15 seconds. During this time, Bridge C port 2 learns the MAC addresses of all connected hosts on the network segment. As it is with the listening state, Bridge C port 2 does not forward traffic during this learning state.

When the learning state is completed, Bridge C port 2 transitions into forwarding mode, in which it forwards traffic as the active path to the root bridge; at this point, the network is considered to be converged.

note During the 50 seconds the network is converging on the change, no traffic is forwarded to or from any of the network bridges and switches. In today's network environment, 50 seconds can seem like an eternity. The Rapid Spanning Tree Protocol (RSTP, IEEE 802.1w) is available to address this issue (the length of time required to transition from the blocking to forwarding state); RSTP enables designated ports to change from the blocking to forwarding state in a few seconds. The exact amount of time depends on the interval between hello timers in your network. Because RSTP does not use timed intervals, as STP does, it is difficult to discuss the precise amount of time it will take an RSTP network to converge. It is because of this lack of precise timing that convergence in an RSTP network can best be measured in "a few seconds."

## Virtual LANs

- [VLAN Overview](#)
- [VLAN Topology](#)
- [VLAN Operation](#)
  - [VLAN Membership](#)
    - [Port-Based VLAN](#)
    - [Address-Based VLAN](#)
    - [Layer 3-Based VLAN](#)
  - [Inter-VLAN Communication](#)
  - [Extending VLANs](#)
  - [VLAN Tagging](#)
- [VLAN Trunking Protocol \(VTP\)](#)
  - [VTP Modes](#)
  - [VTP Pruning](#)
  - [IEEE 802.1q](#)

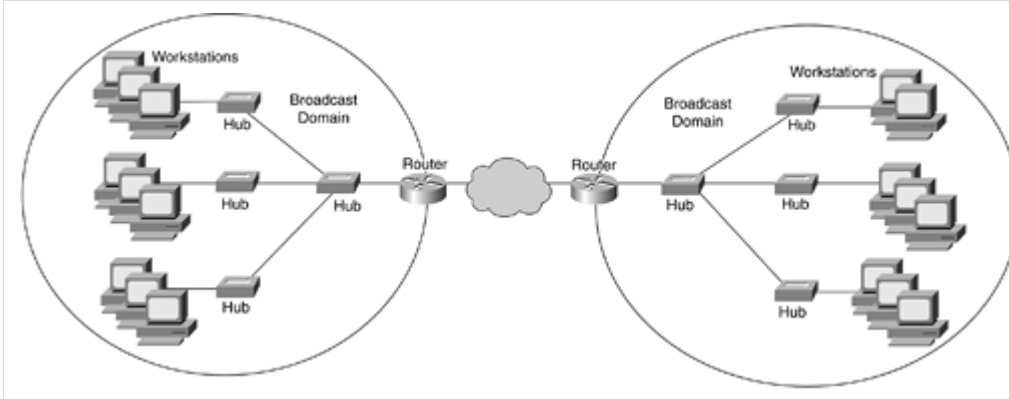
## VLAN Overview

A virtual LAN, or VLAN, is a group of computers, network printers, network servers, and other network devices that behave as if they were connected to a single network.

In its basic form, a VLAN is a broadcast domain. The difference between a traditional broadcast domain and one defined by a VLAN is that a broadcast domain is seen as a distinct physical entity with a router on its boundary. VLANs are similar to broadcast domains because their boundaries are also defined by a router. However, a VLAN is a logical topology, meaning that the VLAN hosts are not grouped within the physical confines of a traditional broadcast domain, such as an Ethernet LAN.

If a network is created using hubs, a single large broadcast domain results, as illustrated in [Figure 8-2](#).

Figure 8-2. Two Broadcast Domains Connected Across a WAN

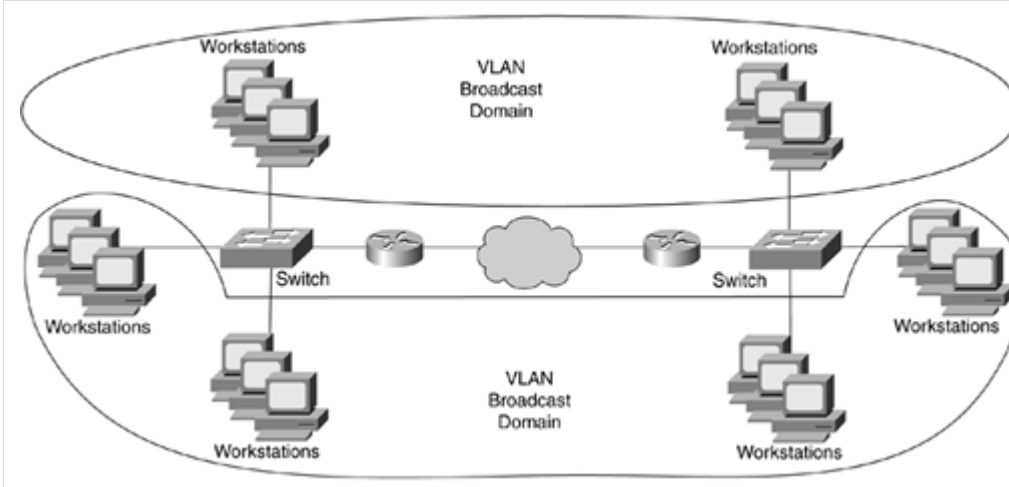


[View full size image](#)

Because all devices within the broadcast domain see traffic from all other devices within the domain, the network can become congested. Broadcasts are stopped only at the router, at the edge of the broadcast domain, before traffic is sent across the wide-area network (WAN) cloud.

If the network hubs are replaced with switches, you can create VLANs within the existing physical network, as illustrated in [Figure 8-3](#).

Figure 8-3. Two VLANs Connected Across a WAN



[View full size image](#)

When a VLAN is implemented, its logical topology is independent of the physical topology, such as the LAN wiring. Each host on the LAN can be assigned a VLAN identification number (ID), and hosts with the same VLAN ID behave and work as though they are on the same physical network. This means the VLAN traffic is isolated from other traffic, and therefore all communications remain within the VLAN. The VLAN ID assignment made by the switches can be managed remotely with the right network management software.

Depending on the type of switching technology used, VLAN switches can function in different ways; VLANs can be switched at the data link (Open System Interconnection OSI model Layer 2) or the network layer (OSI model Layer 3). The main advantage of using a VLAN is that users can be grouped together according to their network communications requirements, regardless of their physical locations, although some limitations apply to the number of nodes per VLAN (500 nodes). This



segmentation and isolation of network traffic helps reduce unnecessary traffic, resulting in better network performance because the network is not flooded. Don't take this advantage lightly, because VLAN configuration takes considerable planning and work to implement; however, almost any network manager will tell you it is worth the time and energy.

note An end node can be assigned to a VLAN by inspecting its Layer 3 address, but a broadcast domain is a Layer 2 function. If a VLAN is switched based on Layer 3 addressing, it is in essence routed. There are two basic differences between routing and switching: First, the decision of forwarding is performed by the application-specific integrated circuit (ASIC) at the port level for switching versus the reduced instruction set circuit (RISC) or main processor for routing; second, the information used to make the decision is located at a different part of the data transfer (packet versus frame).

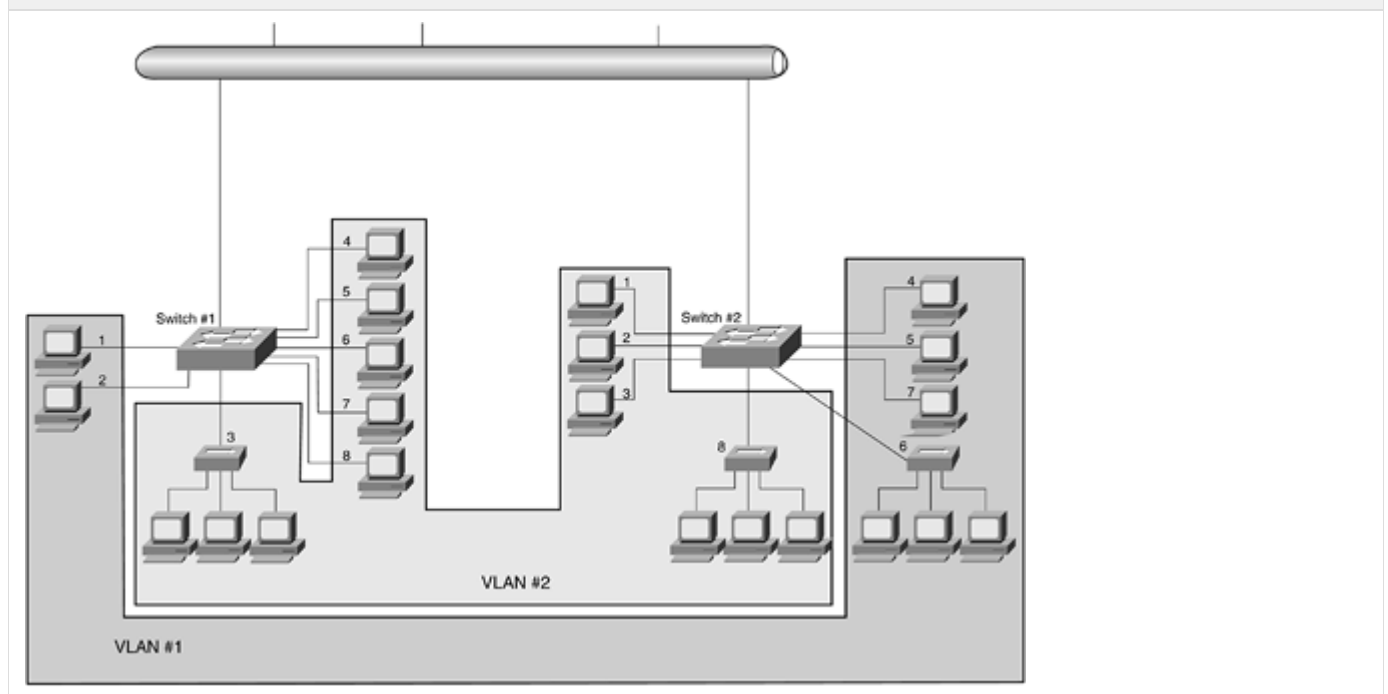
## VLAN Topology

VLANs can best be defined as a group of devices on either the same or different physical LAN segments, interacting with each other if they are on the physical LAN segment.

Suppose, for instance, that you work in a two-floor office building and each floor has a LAN switch providing network connectivity to every computer on that floor. The first floor is supported by Switch 1, and the second floor is supported by Switch 2. On each floor of this building, there is also a marketing staff and an engineering staff. Because of office real estate, people are sitting wherever an open desk can be found.

It is safe to say that the marketing and engineering departments have different jobs and therefore different network requirements. However, the fact that these two departments have different network requirements does not mean they cannot share the same network. [Figure 8-4](#) illustrates how using VLANs provides virtual dedicated network resources to the marketing (VLAN 1) and engineering (VLAN 2) departments, while using the same physical network infrastructure.

Figure 8-4. VLAN 1 and VLAN 2



[View full size image](#)

If we assign all the marketing staff on the first floor (Switch 1, ports 1 and 2) and all the marketing staff on the second floor (Switch 2, ports 4, 5, 6, and 7) to a single VLAN (VLAN 1), they can share resources and bandwidth as if they were connected to the same physical network segment. Similarly, if we assign all the first-floor engineering staff (Switch 1, ports 3, 4, 5, 6, 7, and 8) and the engineering staff on the second floor (Switch 2, ports 1, 2, 3, and 8), we create VLAN 2 for the engineering staff, providing the same illusion of physical connectivity provided to the marketing staff by VLAN 1.

It is important to remember that members of one VLAN cannot share the resources of any other VLAN without some sort of routing mechanism, such as a router or Layer 3 switch. For a member of the marketing staff in VLAN 1 to share resources with the engineering VLAN (VLAN 2), a router or a Layer 3 switch must be in place.

note Communication between VLANs can occur only if there is a router or a Layer 3 switch in place enabling such connectivity.

Switches with VLAN capability can create the same division of the network into separate LANs or broadcast domains and is similar to color coding your switch ports. In [Figure 8-4](#), ports in the light gray area can communicate with other ports in the light gray area, and ports in the dark gray area can communicate with the other ports in the dark gray area.

## VLAN Operation

Several issues are involved in the operation of a VLAN:

- Who can participate in each VLAN
- How VLANs communicate among each other
- How devices within different VLANs can communicate with one another

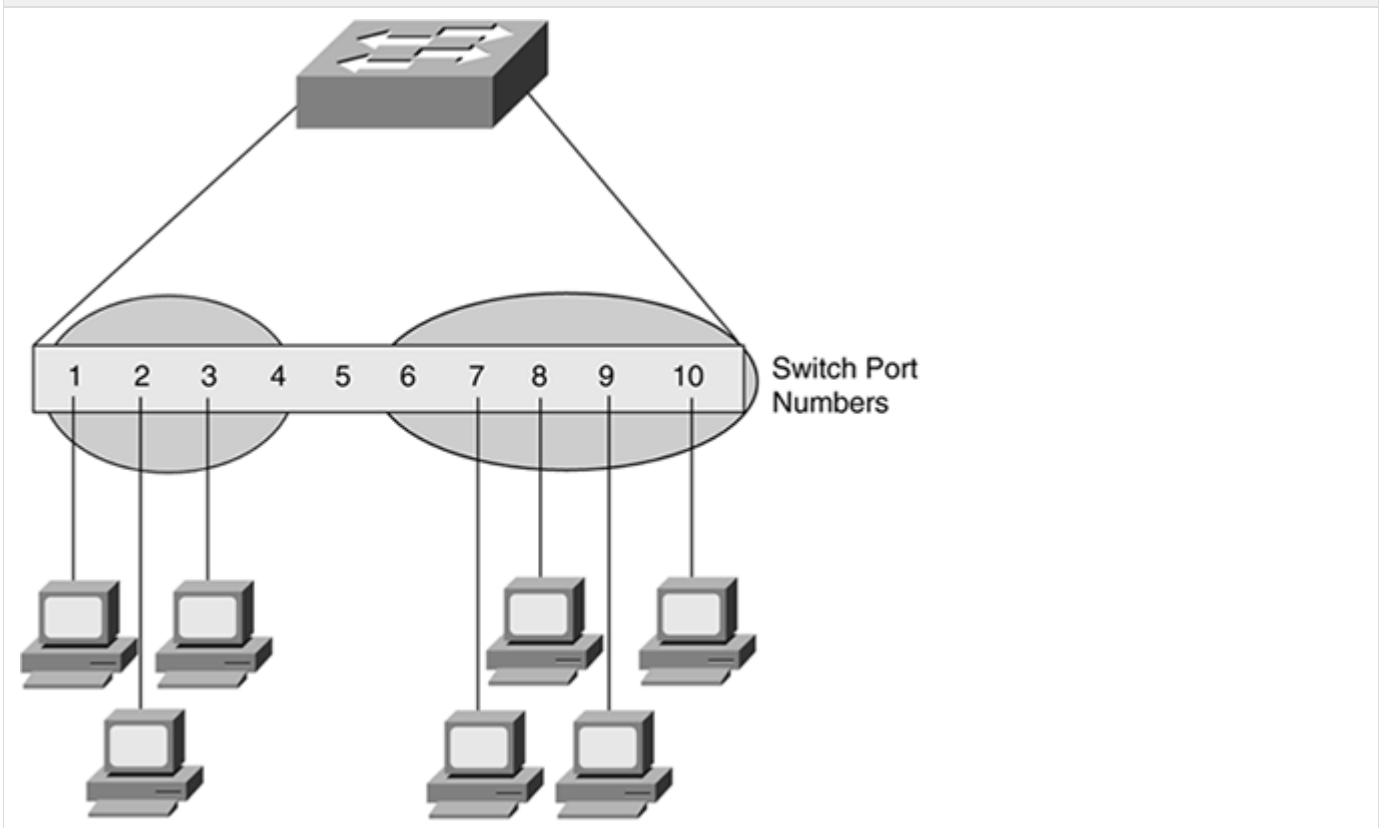
### VLAN Membership

There are three ways a network device can be assigned to a VLAN: by port, Layer 2 (MAC) address, or Layer 3 (network) address. The type of VLAN determines how a device is assigned. In a port-based VLAN, for example, you assign each switch port to a VLAN. In MAC address-based VLANs, membership is defined by the source or destination MAC address. VLANs based on Layer 3 information use the protocol type, such as the Internet Protocol (IP), and the Layer 3 (network) address in determining which VLAN the device is a member of.

#### Port-Based VLAN

In a port-based VLAN, such as that illustrated in [Figure 8-5](#), each computer is assigned to its VLAN based on the port to which the computer is connected.

Figure 8-5. VLAN Membership Based on Switch Port Number



For example, ports 1 through 4 can be assigned to the sales VLAN, ports 6 through 10 to the engineering VLAN, and port 5 kept open as a spare port that you can assign to either VLAN. Or you can create a third VLAN with port 5 as a member. When a computer is connected to port 4, it becomes part of the sales VLAN. When that same computer is connected to port 6, however, it becomes part of the engineering VLAN.

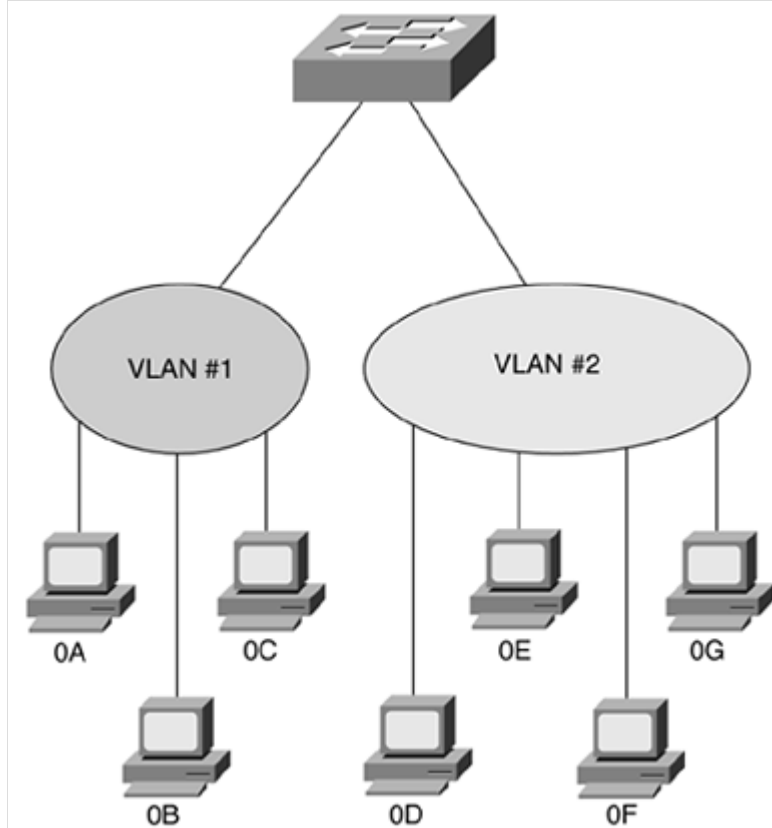
note On almost all switches today, all ports by default are part of VLAN 1.

The main drawback of port-based VLANs is that you must reconfigure VLAN membership when a user moves from one port to another. If you are in an environment in which people are moving around all the time, port-based VLANs can become quite the headache.

#### Address-Based VLAN

In an address-based VLAN, such as that illustrated in [Figure 8-6](#), each computer is assigned to its VLAN based on the Media Access Control (MAC) address of the computer.

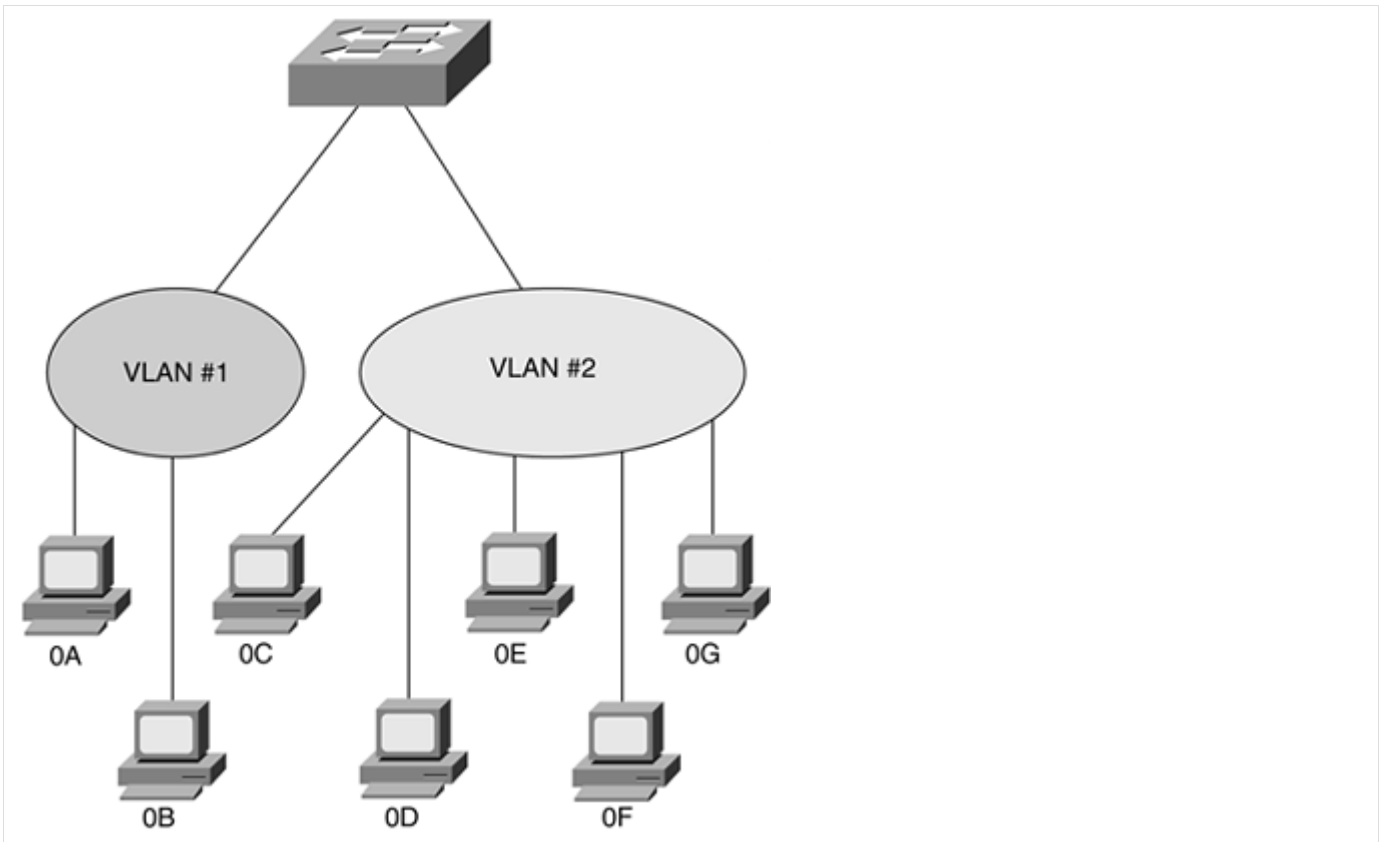
Figure 8-6. Address-Based VLAN



The computers with the MAC addresses 0A, 0B, and 0C are assigned to VLAN 1, and the computers with the MAC addresses 0D, 0E, 0F, and 0G are assigned to VLAN 2. (Note that these are not real MAC addresses.)

The main advantage of the address-based model is that the switch does not need to be reconfigured when a user moves to a different port, as illustrated in [Figure 8-7](#).

Figure 8-7. Machine 0C Moved to New VLAN



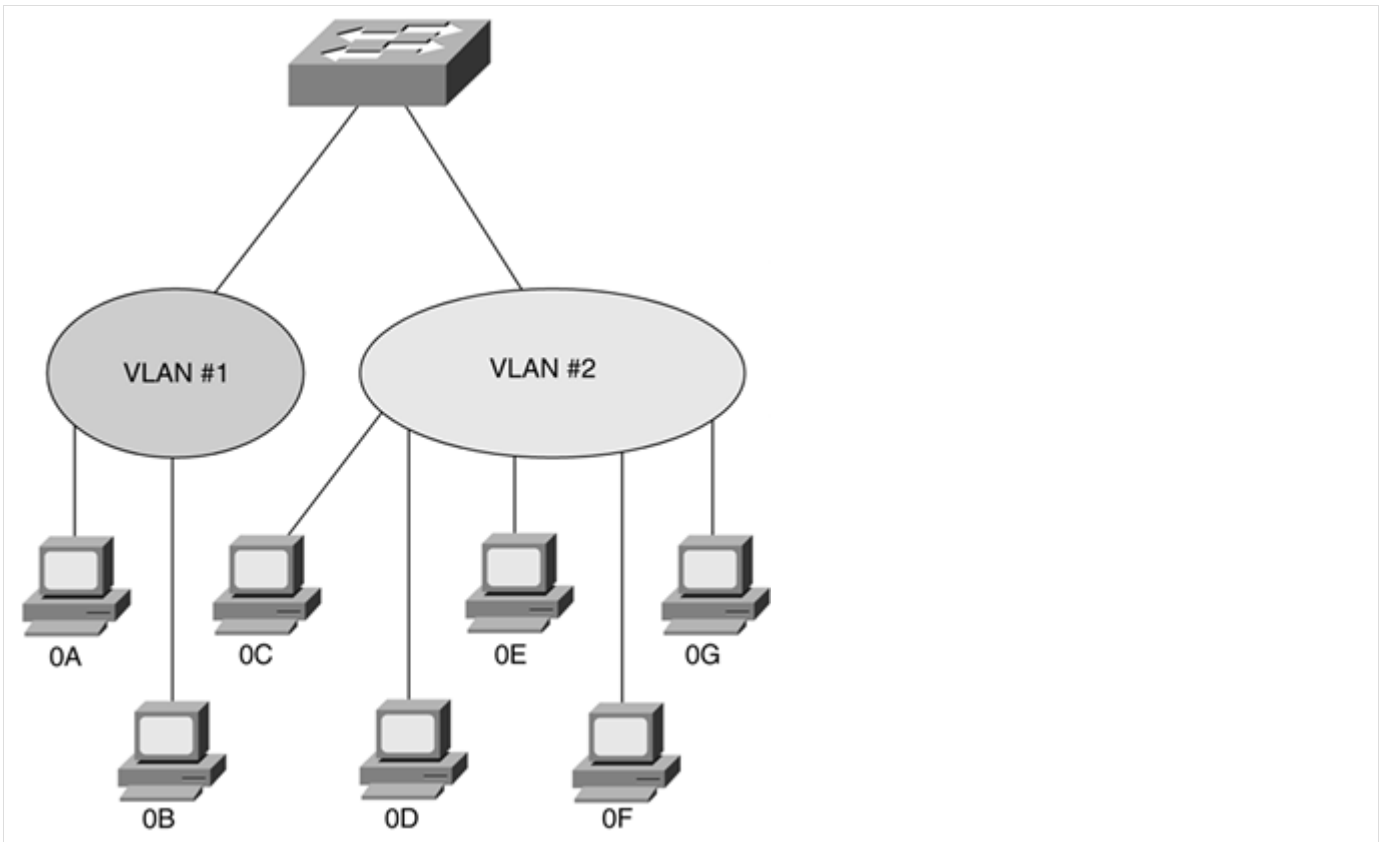
The user at machine 0C changed departments, and to support this move the network administrator removed the MAC address (0C) from VLAN 1 and assigned 0C to VLAN 2 without reconfiguring any switch ports. This type of change can happen about as quickly as you can type on a keyboard.

The primary issue with MAC address-based VLANs is that a single MAC address cannot be a member of multiple VLANs without special features available on the switch enabling the multiple VLAN membership.

### Layer 3-Based VLAN

In a Layer 3-based VLAN, such as that illustrated in [Figure 8-8](#), each computer is assigned to its VLAN based on the OSI model Layer 3, the network layer, and the address of the computer.

Figure 8-8. Layer 3-Based VLAN



The primary benefit of using a Layer 3-based VLAN is that users can physically move their workstations to any network jack without the workstation's network address being reconfigured. This might make your life as a network manager much easier because you assign a network address, or range of addresses, to a VLAN only once, instead of having to reassign a MAC address to a new VLAN. The downside of Layer 3 VLANs is the slow performance caused by additional switch processing.

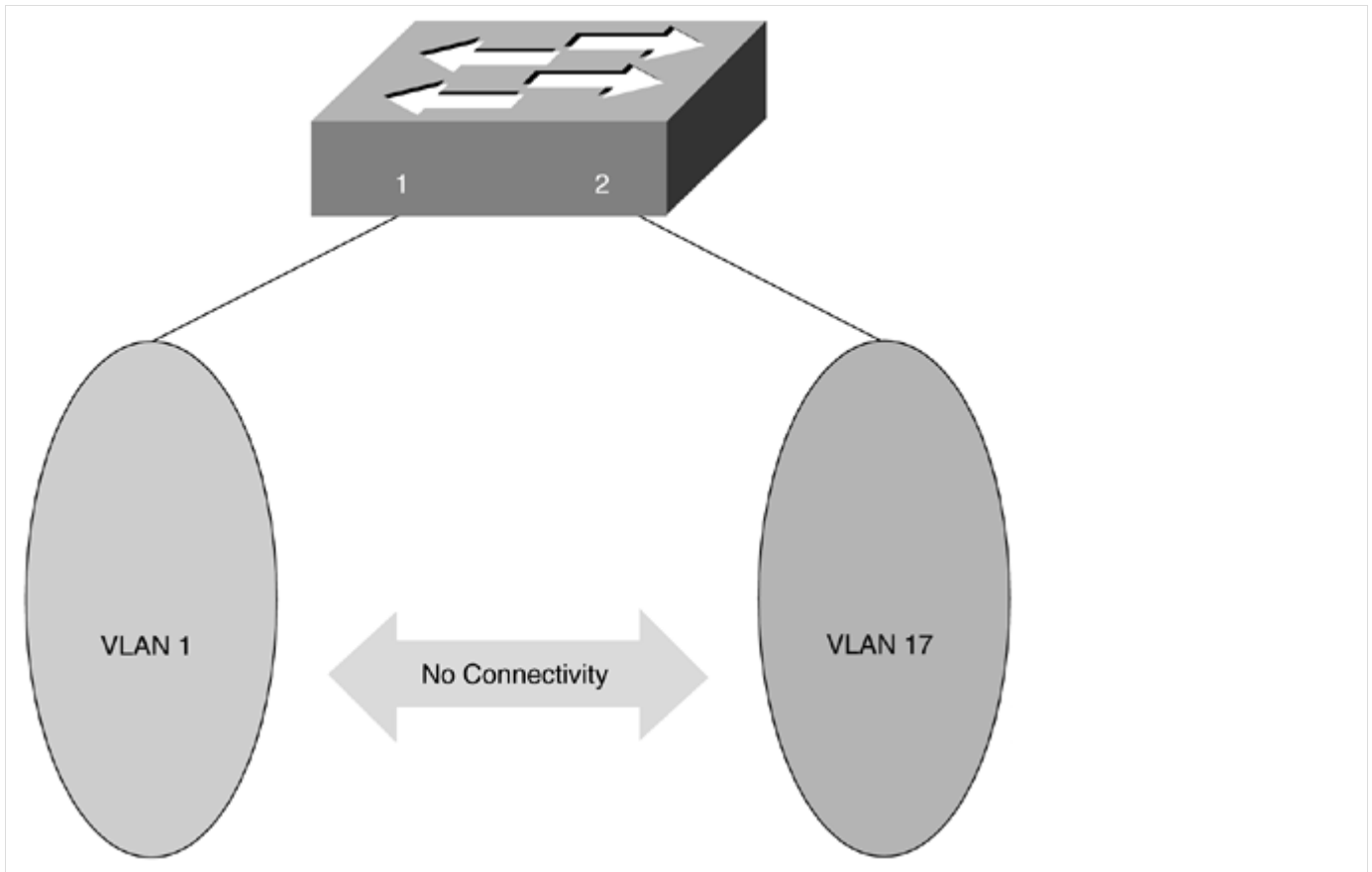
note Because switches are Layer 2 devices, not Layer 3, additional processing cycles are needed for the switch to manage Layer 3-based VLANs. Even though you are using a Layer 3 address to differentiate, the device is being assigned to a Layer 2 broadcast domain (not forwarding the packet).

### Inter-VLAN Communication

We have discussed VLANs that are basically a special type of broadcast domain, in that a VLAN is defined by a switch port rather than by traditional physical boundaries, such as wiring hubs. Recall that when a host in one broadcast domain wants to communicate with another, a router must be involved, and the same holds true for VLANs.

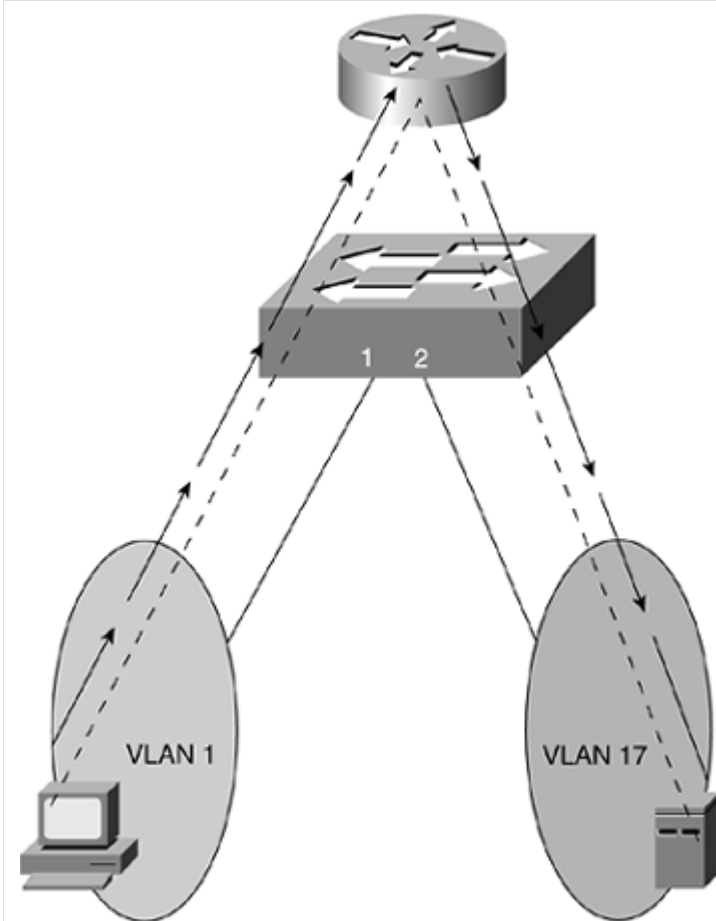
For example, suppose that port 1 on a switch is part of VLAN 1, and port 2 part of VLAN 17, as illustrated in [Figure 8-9](#).

Figure 8-9. VLAN 1 and VLAN 17



If all of the switch's ports were part of VLAN 1, the hosts connected to these ports could communicate with each other without issue. However, when the ports are made part of different VLANs, this communication is no longer possible. For a host connected to port 1 to communicate with another connected to port 2, a router must be involved, as illustrated in [Figure 8-10](#).

Figure 8-10. VLAN 1 and VLAN 17 with a Router



Traffic leaving the host in VLAN 1 passes through the switch to the router so that the traffic can be passed back through the switch to reach the host server in VLAN 17. Instead of using a router to enable this inter-VLAN communication, a Layer 3 switch might be used.

A Layer 3 switch is essentially a Layer 2 switch that can also act as a router, often through additional hardware and/or software features. If a switch is capable of Layer 3 functions, it can be configured to route traffic between VLANs defined within the switch, without the need for traffic to ever leave the switch for routing decisions. If a switch includes only Layer 2 functions, however, an external router must be configured to route traffic between the VLANs. In some cases, a packet can leave switch port 1, be forwarded to an external router, and then be routed right back to port 2 on the originating switch, as illustrated in [Figure 8-10](#). For this reason, Layer 3 switches are popular to use throughout a corporate network.

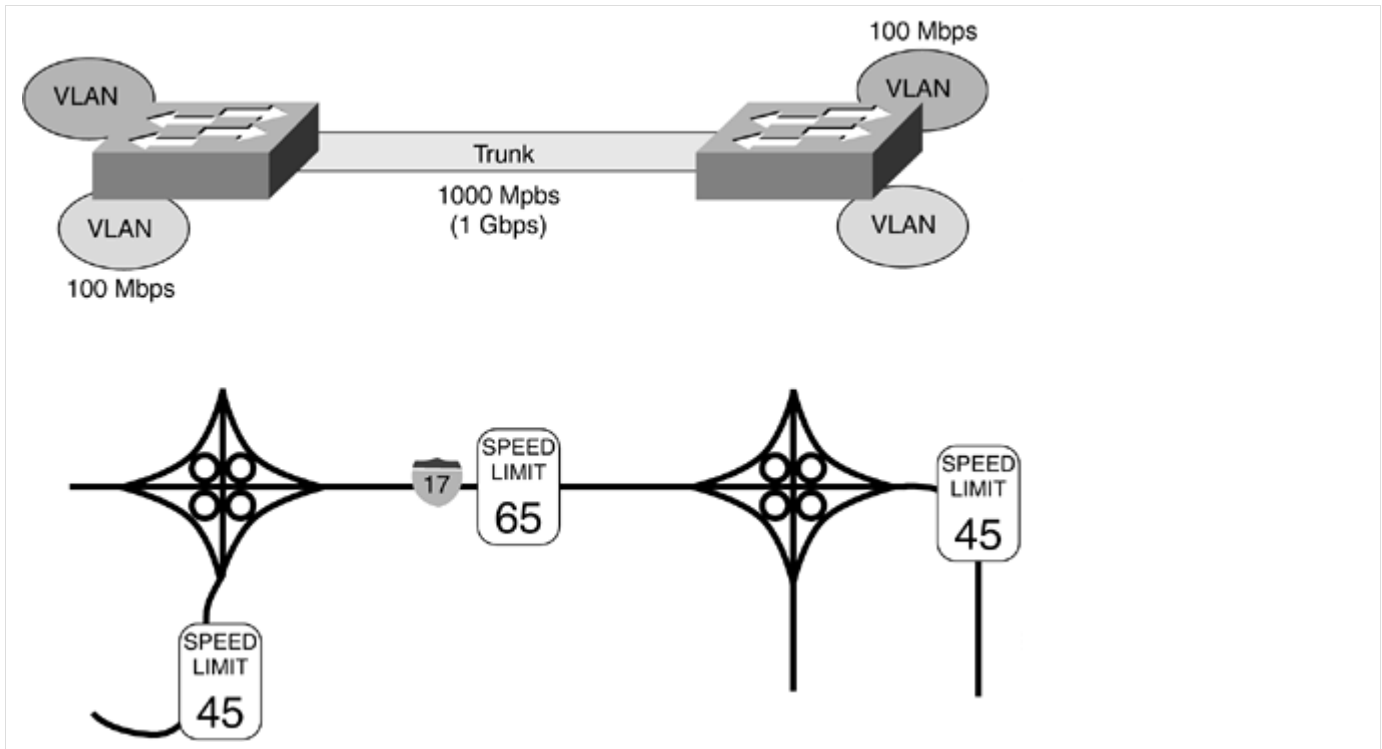
Devices that are called Layer 3 switches track the Layer 3 addresses in and out of each port and build a table similar to a MAC address table for Layer 2. If they see the same address more than once, they forward the packet without looking at the routing table or sending it up to the main processor.

note Regardless of the method chosen for inter-VLAN communication, either a router or Layer 3 switch, the most important point to remember is that when a host on one VLAN wants to communicate with a host on another, a routing (Layer 3) device must be involved.

## Extending VLANs

To extend VLANs across different switches, a trunk link must be implemented, interconnecting the switches. This trunk link is often faster than the VLANs themselves. Think of a trunk link as being similar to an interstate highway; several small roads converge to one larger, and faster, road, as illustrated in [Figure 8-11](#).

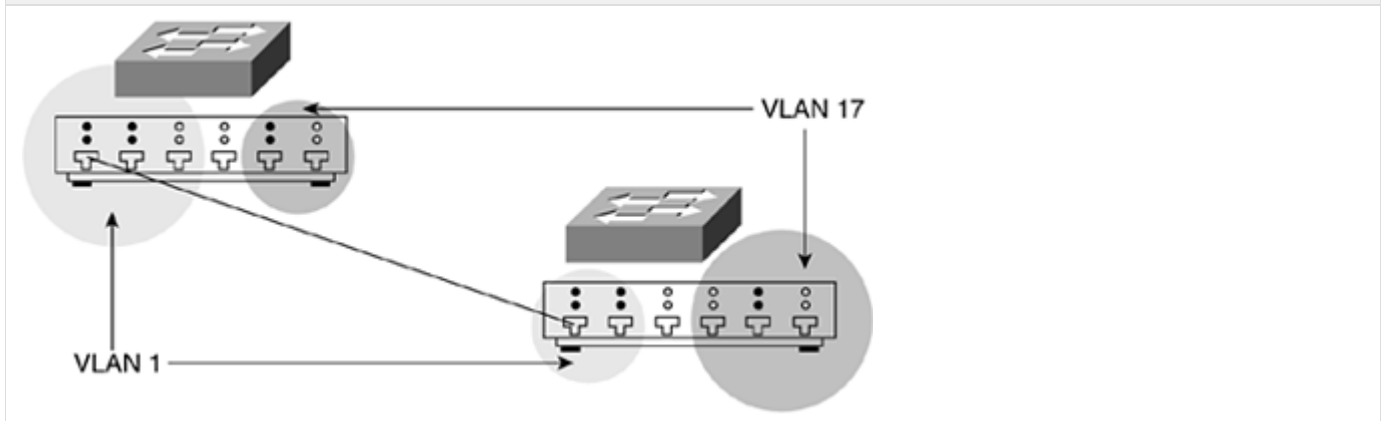
Figure 8-11. VLAN Trunks and Interstate Highways



For example, you might interconnect two Gigabit Ethernet ports on different switches enabling the communication between the 100-Mbps VLANs on each switch. It is recommended that you use the fastest port available for trunk connections between switches, because this link often carries a great deal of traffic, most often for multiple VLANs.

Assume you have connected a link between the 100-Mbps ports of two switches, as illustrated in [Figure 8-12](#).

Figure 8-12. 100-Mbps Link Between VLAN 1 Ports



Note these ports are members of VLAN 1 on each switch. By default, without additional configuration, these ports act as a trunk link between these two switches; however, these ports pass traffic only for the VLAN associated with their port connections (in this case, VLAN 1). This type of link, in which traffic for only a single VLAN is passed, is referred to as an [access link](#), as opposed to a [trunk link](#), which carries traffic for multiple VLANs.

Access links get the job done in a single VLAN environment; however, multiple access links would be required if traffic from multiple VLANs were to be passed back and forth between switches. Having multiple access links between the same pair of switches would be a waste of switch ports. When traffic for multiple VLANs needs to be transferred across a single trunk link, VLAN tagging is used.

## VLAN Tagging

When traffic from multiple VLANs travels across a link interconnecting two switches, you need to configure a VLAN tagging method on the ports that supply the link so that the receiving switch can identify the destination VLAN's traffic.

A number of tagging methods are in use for different technologies. The two discussed here are known as Inter-Switch Link (ISL) and 802.1q. ISL is a Cisco proprietary VLAN tagging method, whereas 802.1q is an open standard. This means that if you

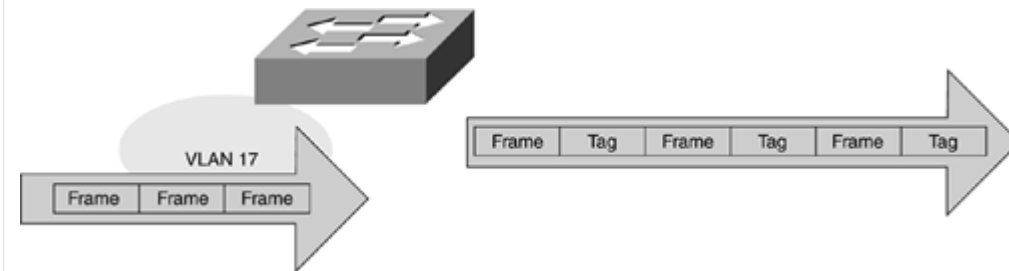


are connecting two Cisco switches, you could use ISL; if any non-Cisco switches are involved, however, 802.1q is your best option.

note ISL is a Cisco proprietary VLAN tagging method; 802.1q is an open standard although both are similar in operation.

ISL tags a frame as it leaves a switch with information about the VLAN to which the frame belongs. If a frame from VLAN 17 is leaving a switch, for example, the ISL port adds information to the frame header, designating that the frame is part of VLAN 17, as illustrated in [Figure 8-13](#).

Figure 8-13. Frames Before and After Tagging by the Switch

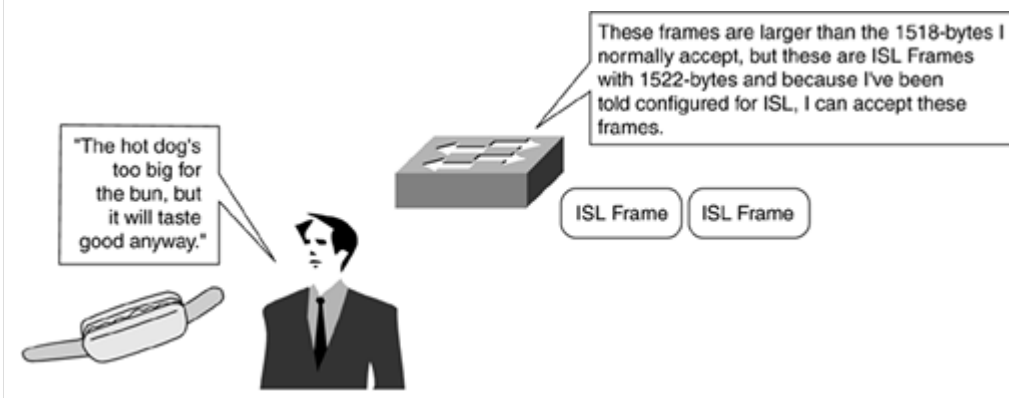


[View full size image](#)

When this ISL frame reaches the port at the other end of the switch, it looks at the ISL header, determines that the frame is meant for VLAN 17, strips off the ISL information, and forwards it into VLAN 17.

One of the issues with VLAN tagging is that by adding information to an Ethernet frame, the size of the frame can move beyond the Ethernet maximum of 1518 bytes to 1522 bytes. Because of this, all non-ISL ports see frames larger than 1518 bytes as giants, and therefore invalid. As shown in [Figure 8-14](#), this is similar to putting a jumbo-sized hot dog in a regular-sized hot dog bun. Just because the hot dog is oversized doesn't make it a bad hot dog. ISL works in much the same way, although without the mustard and relish.

Figure 8-14. ISL Frames and Hot Dogs

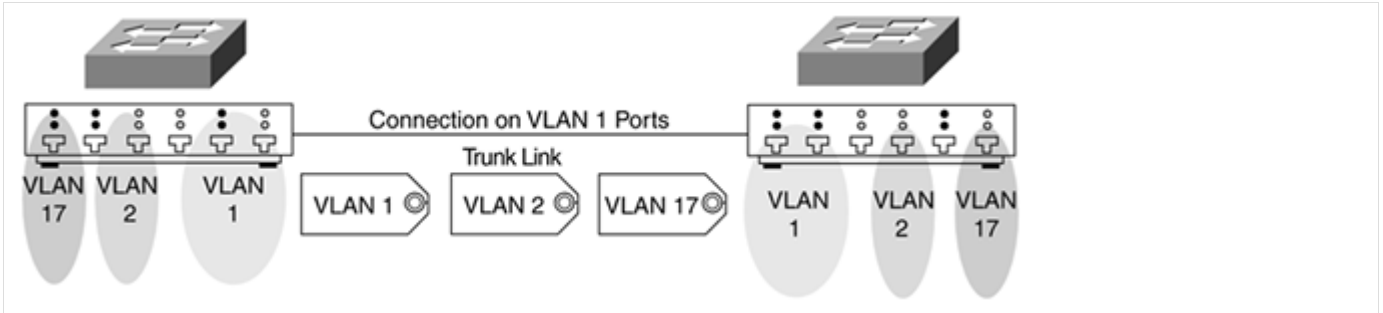


[View full size image](#)

Because the port might see the ISL frame as a giant, the port needs to be configured for ISL so that it can understand the different frame format.

After VLAN tagging has been configured on the ports associated with the link connecting switches, the link is known as a trunk link, as illustrated in [Figure 8-15](#).

Figure 8-15. VLAN Tagging on a Trunk Link



[View full size image](#)

A trunk link transfers frames from many different VLANs by using Cisco ISL or the standard IEEE 802.1q.

## VLAN Trunking Protocol (VTP)

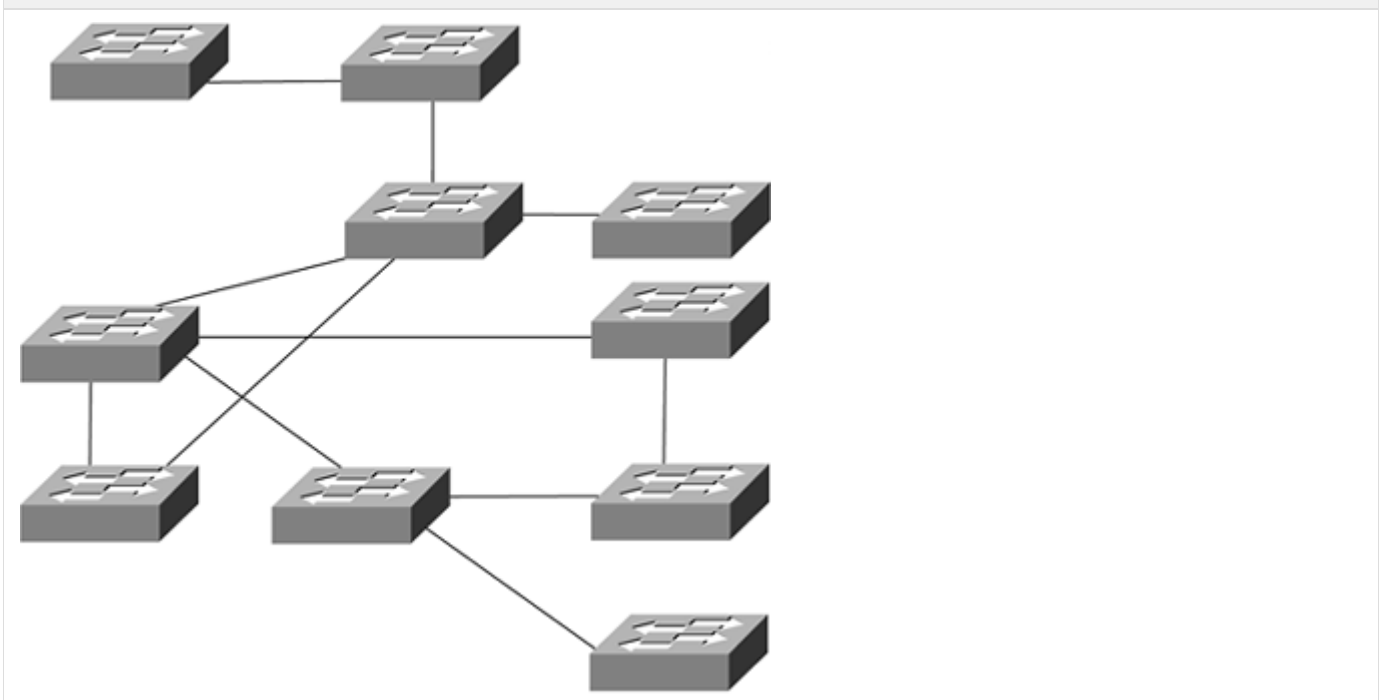
Recall that the purpose of configuring VLAN tagging is to enable traffic from multiple VLANs to cross a trunk link interconnecting switches. However, VLAN tagging does not help ease the burden of configuring individual VLANs on multiple switches; this is where the Cisco VLAN Trunking Protocol (VTP) can help.

note The VTP is a Cisco-proprietary protocol and is useful in large Cisco switch-based environments that include multiple VLANs.

The purpose of VTP is to provide a way to manage Cisco switches as a single group for VLAN configuration purposes. For example, if VTP is enabled on Cisco switches, the creation of a new VLAN on one switch makes that VLAN available to all switches within the same [VTP management domain](#). A switch can be part of only one VTP management domain at a time, and is part of no VTP management domain by default.

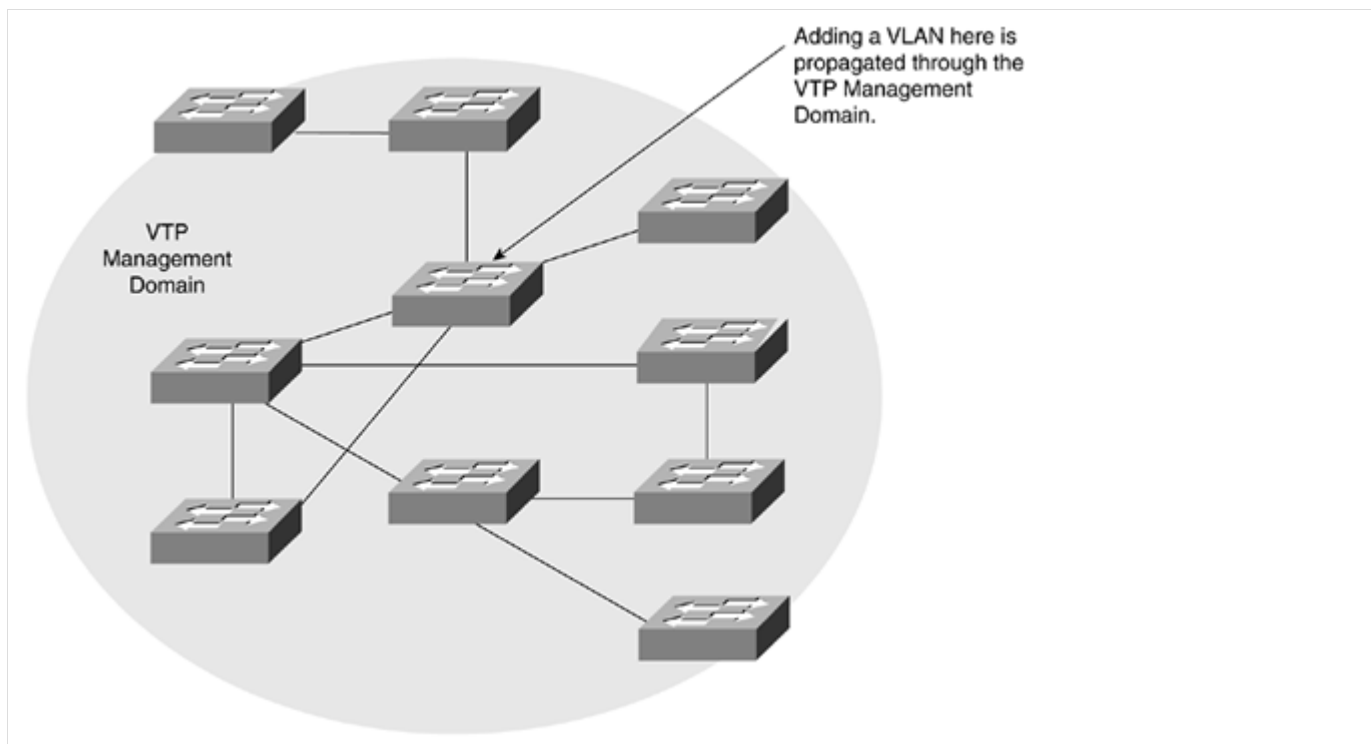
Envision an environment in which you must manage 10 switches, as illustrated in [Figure 8-16](#).

Figure 8-16. 10-Switch Network



Without VTP, the creation of a new VLAN would require you to define that new VLAN individually on all necessary switches, a process that is subject to error and that is time-consuming to say the least. Instead, with VTP, you define the VLAN once and have VTP spread the information to all other switches in the same domain automatically, as illustrated in [Figure 8-17](#).

Figure 8-17. 10 Switches in 1 VTP Management Domain



[View full size image](#)

The primary benefit of VTP is that in large environments it facilitates adding and deleting VLANs, as well as making changes to VLAN configurations. Without VTP you would have to add a VLAN manually to each switch; with VTP you can add a VLAN to one switch and let the switches propagate the changes throughout the VTP management domain, and all before lunch!

When a VTP management domain name is defined on each switch, the switches exchange VTP information automatically and require no further configuration or day-to-day management.

## VTP Modes

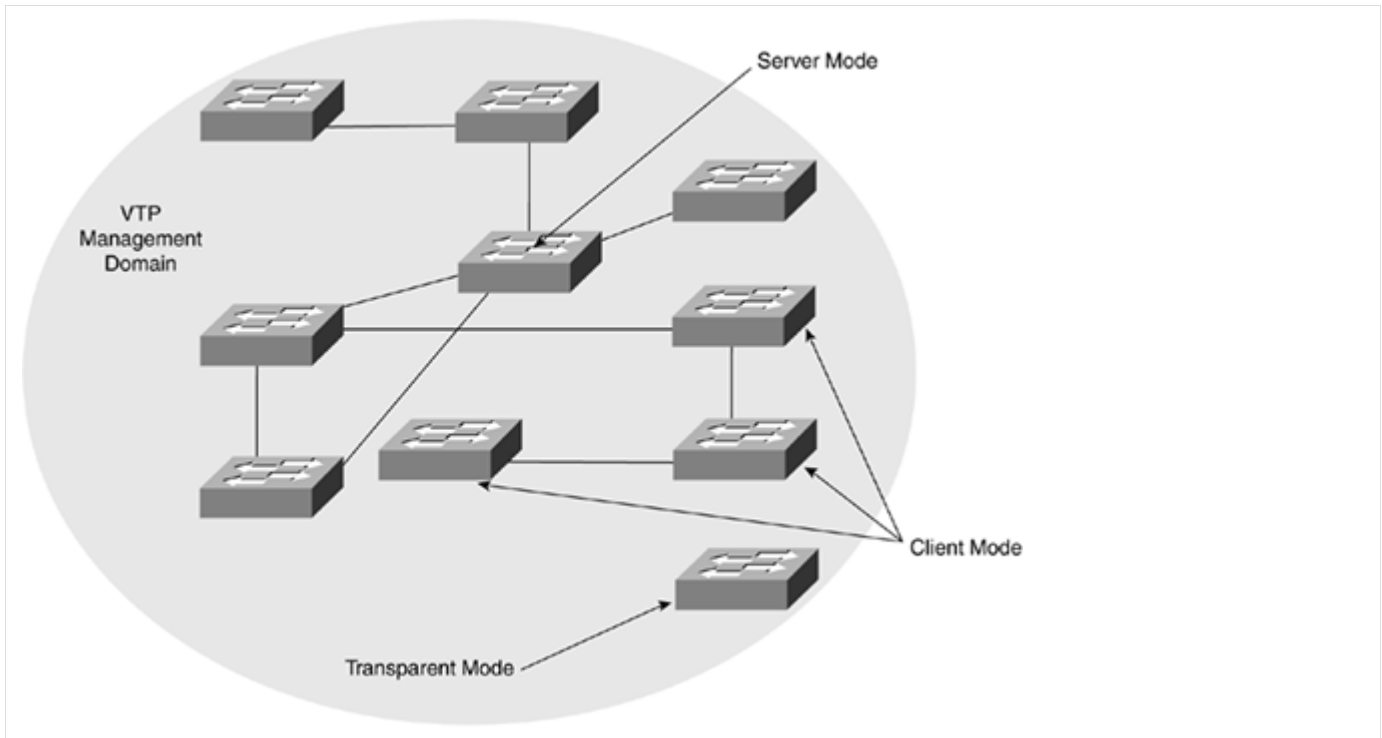
If you intend to make a switch part of a VTP management domain, each switch must be configured in one of four possible VTP modes: server, client, transparent, and off. The VTP mode assigned to a switch determines how the switch interacts with other VTP switches within the VTP management domain.

The following list details each of these four VTP modes:

- **Server mode** - A switch configured in server mode can be used to add, delete, and change VLANs within the VTP management domain. Server is the default mode used after a VTP has been configured on a Cisco switch. Within any VTP management domain, at least one switch must be in server mode. When in server mode, changes are passed to all other switches within the VTP management domain.
- **Client mode** - A switch configured in client mode is the recipient of any changes within the VTP management domain, such as the addition, deletion, or modification of VLANs by a server mode switch. A switch in VTP client mode cannot make any changes to VLAN information.
- **Transparent mode** - A switch configured in transparent mode passes VTP updates received by switches in server mode to other switches in the VTP management domain, but does not process the contents of these messages. When individual VLANs are added, deleted, or changed on a switch running in transparent mode, the changes are local to that particular switch only, and are not passed to other switches within the VTP management domain.
- **Off** - With the introduction of IOS version 7.1.1, the option now exists to disable VTP completely on a switch.

Figure 8-18 illustrates the use of each VTP mode.

Figure 8-18. VTP Modes in Action



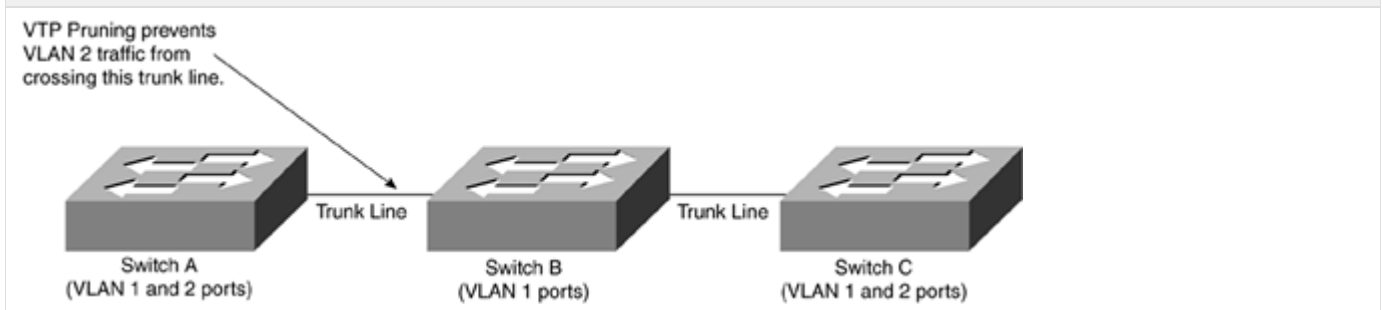
[View full size image](#)

For example, think of the 10-switch network described earlier in this chapter. You could configure each switch to be in the same VTP management domain. Although each could be left in the default server mode, it might be easier to leave only one switch in server mode and configure all remaining switches for VTP client mode. When you need to add, delete, or change a VLAN, the change can be carried out on the VTP server-mode switch and passed to all client-mode switches automatically. When you need a switch to act in a relatively standalone manner, or don't want it to propagate information about its configured VLANs, transparent mode should be used.

## VTP Pruning

Although the configuration of trunk links by using protocols such as ISL enables traffic from multiple VLANs to travel across a single link, this is not always the optimal choice. For example, suppose three switches are connected by two trunk links, as illustrated in [Figure 8-19](#).

Figure 8-19. VTP Pruning



[View full size image](#)

In this example, all three switches include ports that are part of VLAN 1, but only Switches A and C include ports in VLAN 2. Without VTP pruning, traffic for VLAN 2 will be passed to Switch B, even though it does not have any ports configured for VLAN 2.

When VTP pruning is implemented within a VTP management domain, traffic for a given VLAN is passed only to a switch across a trunk link if necessary. In [Figure 8-19](#), for example, implementing VTP pruning in the management domain would ensure that traffic for VLAN 2 is never passed to Switch B until Switch B has VLAN 2 ports configured.

VTP advertisements are sent every 5 minutes or when a change occurs. Switches overwrite only information with a higher revision number. If a switch receives an update with VTP revision 14 but the switch is running on VTP revision 16, for example, it ignores the older revision, much as you ignore yesterday's newspaper when today's arrives on your doorstep.

## IEEE 802.1q

The Institute of Electrical and Electronics Engineers (IEEE) has defined the 802.1q standard for VLANs, ensuring the interoperability of VLAN implementations between switches and network interface cards (NICs) from different vendors. Because of the various types of VLAN definitions, each vendor has developed its own unique and proprietary VLAN solution and product, such as the Cisco VTP. Without some common ground, such as an open standard, switches from one vendor will not interoperate with VLANs from other vendors.

## Switching Security

- [Network Security Basic Rules](#)
- [Port Security](#)
  - [Port Security Configuration Guidelines](#)
- [Virtual LANs](#)
  - [VLAN 1 Precautions](#)
  - [Trusted and Untrusted Ports](#)
- [VLAN-Based Network Attacks](#)
  - [MACFlooding Attack](#)
  - [ARP Attacks](#)
  - [Private VLAN Attack](#)
  - [Multicast Brute-Force Attack](#)
  - [Spanning-Tree Attack](#)
  - [Random Frame-Stress Attack](#)
- [Chapter Summary](#)

## Network Security Basic Rules

You need to keep in mind several basic rules when setting up secure Layer 2 switched VLANs:

- VLANs should be set up in such a way that the VLAN clearly separates the network's various logical components from each other, in turn segregating logical workgroups. This is the first step toward segregating those portions of your network that need more security from portions that need less.
- If some switch ports are not being used, it is best practice to disable these ports and assign them to a special VLAN that collects these unused ports. This special VLAN should have no Layer 3 connectivity, such as to a router or other Layer 3 device capable of switching.

Although devices in a particular VLAN cannot access devices in another VLAN unless a trunking or routing mechanism is available, VLANs should not be used as the single mechanism for providing network security. VLAN protocols are not designed with network security as the primary goal, and because of this VLAN protocols can be compromised rather easily. Unfortunately, VLANs enable loopholes into the network. Because VLAN protocols are not security conscious, you should use other mechanisms, such as those discussed in the next sections, to secure the network.

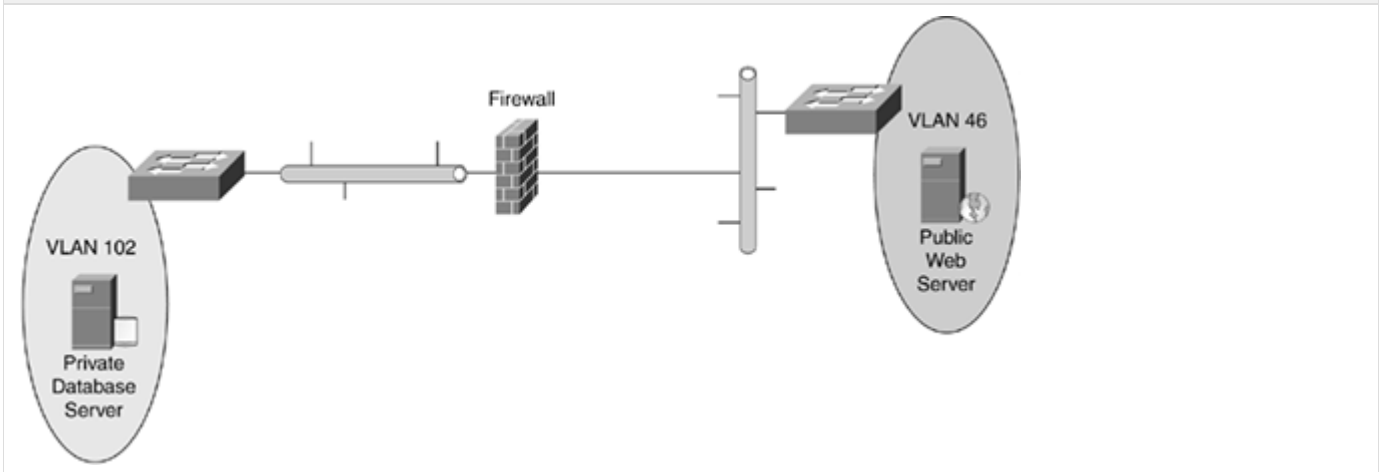
Because VLANs lack security, devices at different security levels should be isolated on physically separate Layer 2 devices. For example, having the same switch chassis on both the inside and outside of a firewall is not recommended, as illustrated in [Figure 9-2](#).

Figure 9-2. Public and Private VLAN Behind the Same Firewall



Putting both the public (VLAN 46) and private (VLAN 102) VLANs on the same switch, behind the firewall, is not a good idea. The VLAN separation does not provide enough security for your private information, such as a corporate database. This is not recommended because the management of the switch is more easily compromised by having a public VLAN. In addition, this is not recommended because a simple misconfiguration or incorrect cabling could expose the management interface of the switch. [Figure 9-3](#) illustrates the solution to this type of scenario.

Figure 9-3. Public and Private VLANs Separated by Two Switches



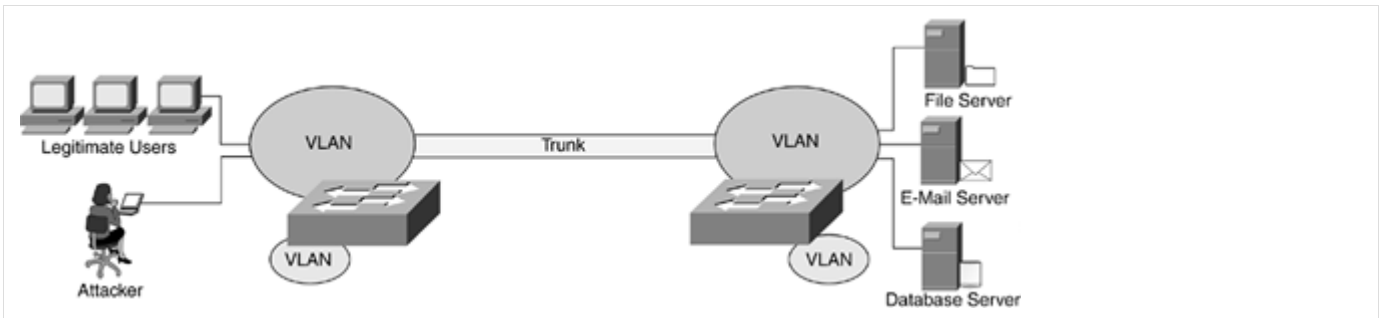
[View full size image](#)

Two separate switches should be used for the [secure](#) and [nonsecure](#) sides of the [firewall](#): one switch on the public side of the firewall and one switch on the private side of the firewall.

note High-end switches can perform firewall functions without using an external firewall device.

An important point to remember is that you need to make sure VLAN trunking in your network does not become a security risk in the network switching environment. VLAN trunks should not use switch port numbers that belong to the native VLAN. Because the native VLAN is a VLAN that is not associated explicitly to a trunk link, the native VLAN enables network packets from the trunk port to reach other ports located in the same native VLAN, as illustrated in [Figure 9-4](#).

Figure 9-4. When VLAN Trunking Goes Awry



[View full size image](#)

The VLAN trunk between the two switches in [Figure 9-4](#) is part of an active VLAN. Therefore, if an attacker gains access to that VLAN, that same person now has access to all network resources inside that VLAN, such as user workstations or servers. (Aren't network attackers annoying?)

Switch ports that do not require trunking should have trunking disabled because, as illustrated in [Figure 9-4](#), an attacker can use this trunking to hop from one VLAN to another. The attacker can do this by pretending to be another switch looking to establish a VLAN trunk with an active switch. This impersonation of a switch enables the attacker's machine to become a part of all the VLANs on the switch being attacked.

It's a good idea to use dedicated VLAN IDs for all VLAN trunks rather than using VLAN IDs that are also being used for nontrunking ports. If you don't use separate VLAN IDs, you enable an attacker to be part of a trunking VLAN pretty easily and then in turn use trunking to hop on to other VLANs as well. In other words, your attacker just bribed Patches with a steak.

note Layer 3 interfaces between switches provide additional access control.

If one of your network users does not want his workstation to be tampered with, that user must control the physical access to that workstation, such as powering off the computer at the end of the day. In addition, it is important for any network administrator or manager to use all the proven security tools available for his or her specific platforms. These security tools range from the very basic configuration of system passwords, IP permit filters, and login banners, to more advanced tools such as Remote Authentication Dial-In User Service (RADIUS), Terminal Access Controller Access Control System Plus (TACACS+), and intrusion detection systems (IDSs).

Only after the basic security components are in place is it possible to turn attention to some of the more sophisticated security details, such as the use of port security or VLANs in your network, which are discussed in the following sections.

## Port Security

When port security is enabled on a switch, any Media Access Control (MAC) address not specified for that port is denied access to the switch, and to any networks to which the switch is connected. Port security can be used to block input to an Ethernet, Fast Ethernet, or Gigabit Ethernet switch port.

The total supply, or global resource, of MAC addresses for the switch is 1024 MAC addresses. However, not all Cisco switches have 1024 MAC addresses; some have only 64 MAC addresses. In addition to this total supply, there is space for one default MAC address per port to be secured. The total number of MAC addresses that can be specified per port is limited to the global resource of 1024 MAC addresses plus one default MAC address (per port).

note The total number of MAC addresses on any port cannot exceed 1025. Bear in mind that the switch limit is 1024 MAC addresses total for use.

The maximum number of MAC addresses for each port depends on your network configuration. The following combinations are some examples of valid allocation of MAC addresses:

- 1025 (1 + 1024) addresses on 1 port and 1 address each on the rest of the ports
- 513 (1 + 512) each on 2 ports in a system and 1 address each on the rest of the ports
- 901 (1 + 900) on 1 port, 101 (1 + 100) on another port, 25 (1 + 24) on the third port, and 1 address each on the rest of the ports

Each of these examples is listed in [Table 9-1](#), grouped together by shades of gray. Note that the total number of allocated MAC addresses does not exceed 1024.

Table 9-1. MAC Address Allocation Examples

Number of Ports	x	Number of MAC Addresses	=	Total
1		1024		1024

2		512		1024
1		900		900
1		100		100
1		24		24
				1024

After you have allocated the maximum number of MAC addresses on a switch port, you can do one of two things:

- Manually specify the secure MAC address for the port
- Have the port dynamically configure the MAC address of the connected devices

From an allocated number of maximum MAC addresses on a port, you can manually configure all, allow all to be autoconfigured, or configure some manually and allow the rest to be autoconfigured. After the port addresses have been configured, manually or automatically, they are stored in nonvolatile rapid-access memory (NVRAM).

After you allocate a maximum number of MAC addresses on a port, you specify a period of time, called the age time, during which the addresses on the specified port remain secure. After this age time expires, the MAC addresses on the port become insecure and are no longer trusted.

note All addresses on a port are permanently secured by default.

If a security violation occurs, you can configure the port to go into shutdown mode or restrictive mode. Shutdown mode gives you the option of specifying whether the port is permanently disabled or disabled for a specified amount of time. The default action during a security violation is for the port to permanently shut down. Restrictive mode allows port configuring to remain enabled during the security violation, only stopping packets coming in from insecure hosts.

When a secure port receives a frame, the frame's source MAC address is compared to the list of secure source addresses that were configured (manually or learned via autoconfiguration) on the port. If the MAC address of a device attached to the port is not on the secure address list, the port is shut down, either permanently or for a period of time you've configured.

## Port Security Configuration Guidelines

When configuring port security, consider the following guidelines:

- You cannot configure port security on a [trunk port](#).
- Port security cannot be enabled on a Switched Port Analyzer (SPAN) port.
- You cannot configure dynamic, static, or permanent content-addressable memory (CAM) entries on a secure port.
- When you enable port security on a port, any static or dynamic CAM entries associated with the port are cleared; any currently permanent CAM entries that are configured by an administrator are treated as secure.

## Virtual LANs

Recall from [Chapter 8](#) that a virtual LAN, or VLAN, is a group of computers, network printers, network servers, and other network devices behaving as if they were connected to a single, network segment.

Network attackers or malicious users often seek to gain access to the management console of a networking device, because if they are successful, they can easily alter the network configuration to their advantage.

In a VLAN switch, in addition to having a direct connection to an [out-of-band](#) management port (a port not used for user traffic), the network management station can use one or more VLANs for [in-band](#) management. The network management station can also use one or more VLANs to exchange protocol traffic with other networking devices.

Basic physical security guidelines require networking equipment to be in a controlled or locked space, such as a telephone room or communications closet. VLAN-based security's primary rule is confining in-band management and protocol traffic to a logically controlled environment, by implementing the following tools and best practices:

- Using traffic and protocol access control lists (ACLs) or filters preventing untrusted traffic from being filtered, or passed, through the switch
- Disabling Layer 2 protocols on untrusted ports, such as disabling the Cisco [Dynamic Trunking Protocol \(DTP\)](#) on switch access ports
- Configuring in-band management switch ports only in dedicated and trusted VLANs
- Not using VLAN 1 to carry user or network data traffic

There is a VLAN used for special requirements within your switch network: VLAN 1.

### VLAN 1 Precautions

VLAN 1 is special because switches need to have a default VLAN to assign to their ports, including management ports, and



VLAN 1 is the default VLAN. In addition, many Layer 2 protocols need to send their information across a specific VLAN on trunk links. It was for these purposes that VLAN 1 is used, and therefore VLAN 1 should not be used for user-related traffic.

As a result of this selection, VLAN 1 can sometimes end up spanning the entire network if not appropriately configured. If the diameter of VLAN 1 is large enough, the risk of instability significantly increases. Using a universal VLAN for management purposes puts trusted network devices, such as workstations and servers, at higher risk of security attacks from untrusted network devices. These untrusted network devices might gain access by switch misconfiguration, or accidentally gain access to VLAN 1 and then try to exploit this unexpected security hole in your network.

At present VLAN 1 has a bad reputation to overcome; with a little bit of help, however, VLAN 1 can redeem itself. To redeem VLAN 1, a simple security principle should be used: As a rule, the network administrator should prune any VLAN, most notably VLAN 1, from all the ports where that VLAN is not strictly needed.

note VLAN Trunking Protocol (VTP) pruning is disabled by default.

The rule of VLAN pruning means four things to you:

- Do not use VLAN 1 for in-band or out-of-band management traffic. Instead, use a different dedicated VLAN, thereby keeping management traffic separate from user data and other necessary network protocol traffic.
- Prune VLAN 1 from all VLAN trunks and from all access ports that do not require participation in VLAN 1, including switch ports that are not connected or shut down. If a switch port is not being used for any reason, move it to a new VLAN created for this purpose. This VLAN should also be pruned.
- Do not configure the management VLAN on any trunk or access port not requiring participation in the management VLAN. This includes switch ports not connected to any network segments and ports that are shut down and not in use.
- When feasible, for near-foolproof security, use an out-of-band network management platform, separating your network management traffic from your network user, or data, traffic.
- If VLANs other than VLAN 1 or the management VLAN represent a security concern, automatic or manual VLAN pruning should be applied. When a VLAN is automatically pruned, the VLAN must be manually enabled.

## Trusted and Untrusted Ports

Apart from VLAN pruning in your network, another security principle you should put into practice is this: Connect untrusted (nonsecured) devices to untrusted ports, trusted (secured) devices to trusted ports, and disable all remaining ports.

This security principle means four things to you:

- If a switch port is connected to an unknown, or foreign, device, do not try to speak the language of this unknown device because doing so could be turned to an attacker's advantage and used against you. On the switch port in question, disable any unnecessary network management protocols, such as the DTP, because you do not want to risk potentially dangerous communication with an untrustworthy neighbor.
- To prevent undesirable protocol interactions within the network-wide VLAN configuration in your network, configure VTP domains appropriately or turn off VTP. This precaution limits or prevents the risk of human error, in the form of mistakes made by a network administrator, from spreading throughout the network. Because the switch with this error would have a newer VTP revision number, the entire domain's VLAN configuration is at risk of being reconfigured with the error. Oops.
- By default, only switch ports known to be trusted should be treated as such and all other ports should be configured as untrusted. There is an adage that fits here: We trust, but verify.
- Create a VLAN to collect unused switch ports, and disable unused switch ports and put them in this unused VLAN. By not granting connectivity to this VLAN, or by placing a device into a VLAN not in use, unauthorized network access can be stopped through physical and logical barriers. In other words, while Patches (physical barrier) is enjoying his steak, the home burglar is contained in the garage because of an alarm system on the house door (logical barrier).

## VLAN-Based Network Attacks

The majority of Layer 2 (data link layer) attacks exploit the inability of a switch to track an attacker, because the switch has no inherent mechanism to detect that an attack is occurring. This inability to detect an attacker means that this same attacker can perform malicious acts against the network path, altering the path and exploiting the change without detection.

note Some of the newer switches introduced to the market can track network attackers with the implementation of firewall and IDS modules or Cisco Network-Based Application Recognition (NBAR). Firewalls are used to prevent unauthorized access to your network, and IDS sensors are used to track network attack and intrusion attempts. Cisco NBAR adds intelligent network classification to network infrastructures by using a classification engine that recognizes a wide variety of applications, including web-based applications.

Some of the most common Layer 2 attacks are as follows:

- MAC flooding attack
- Address Resolution Protocol (ARP) attacks
- Private VLAN attack
- Multicast brute-force attack
- Spanning-tree attack
- Random frame stress attack

Each of these attacks is discussed in detail in the following sections.

## MACFlooding Attack

A MAC flooding attack is not a network attack but more a limitation of the way switches and bridges work. Switches and bridges possess a finite hardware-learning table to store the source addresses of all received packets. When this table becomes full, traffic directed to addresses that cannot be learned anymore is permanently flooded. Traffic flooding is constrained within the VLAN of origin, and therefore no VLAN hopping is permitted.

On nonintelligent switches, this flooding problem arises because a sender's Layer 2 identity is not verified, and therefore the sender can impersonate an unlimited number of network devices by counterfeiting frames.

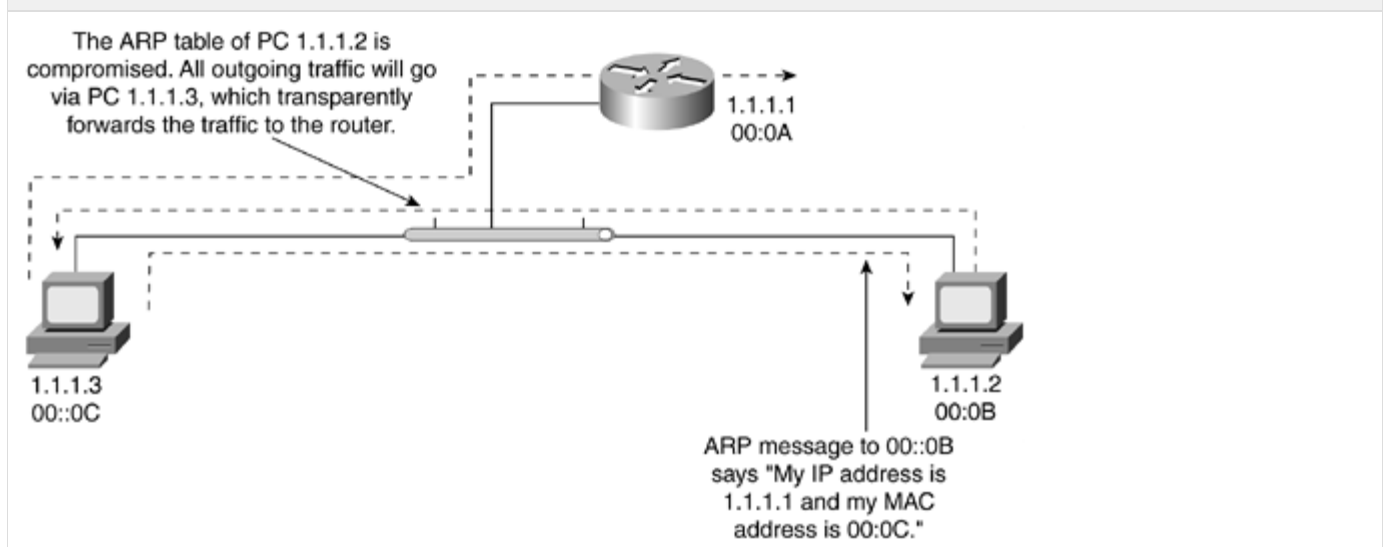
Port security, 802.1x, and dynamic VLANs are three features that you can use to limit a device's connectivity based on its user's login ID and the device's own MAC layer identification. With port security, for example, preventing MAC flooding attacks is as simple as limiting the number of MAC addresses that can be used by a single port. By using port security in this way, you tie the identification of the device's traffic to its port of origin. Dynamic VLANs enable you to dynamically assign switch ports to VLANs based on the Media Access Control (MAC) address of the device connected to the port. When you move a host from one switch port to another switch port in the network, that switch dynamically assigns the new port to the assigned VLAN for that device.

## ARP Attacks

**Address Resolution Protocol (ARP)** is an old protocol and was developed back in the time when everyone in a network was supposed to be friendly. Because ARP was designed for a friendly environment, no security was built in to the ARP function. As a consequence, anyone can claim to be the owner of any IP address he likes. In other words, an attacker can say that his MAC address is associated to any IP address in your network. These false claims result from the fact that ARP requests and replies carry information that associates the MAC address with the IP address of a device. Because there is no way to verify these identities, anyone trying to break into your network can pretend to be someone else, such as a legitimate user of your network, and gain access to resources on your network, such as a corporate database.

ARP attacks are targeted to fool a switch into forwarding packets to a device in a different VLAN by sending ARP packets containing forged identities. Within the same VLAN, ARP attacks, also known as ARP poisoning, can fool network end nodes, such as workstations or routers, into learning these false identities. These counterfeited identities enable a malicious user to pretend to the network that she is an intermediary between two endpoints and perform a **man-in-the-middle (MiM) attack**, as illustrated in [Figure 9-5](#).

Figure 9-5. ARP Spoofing Attack



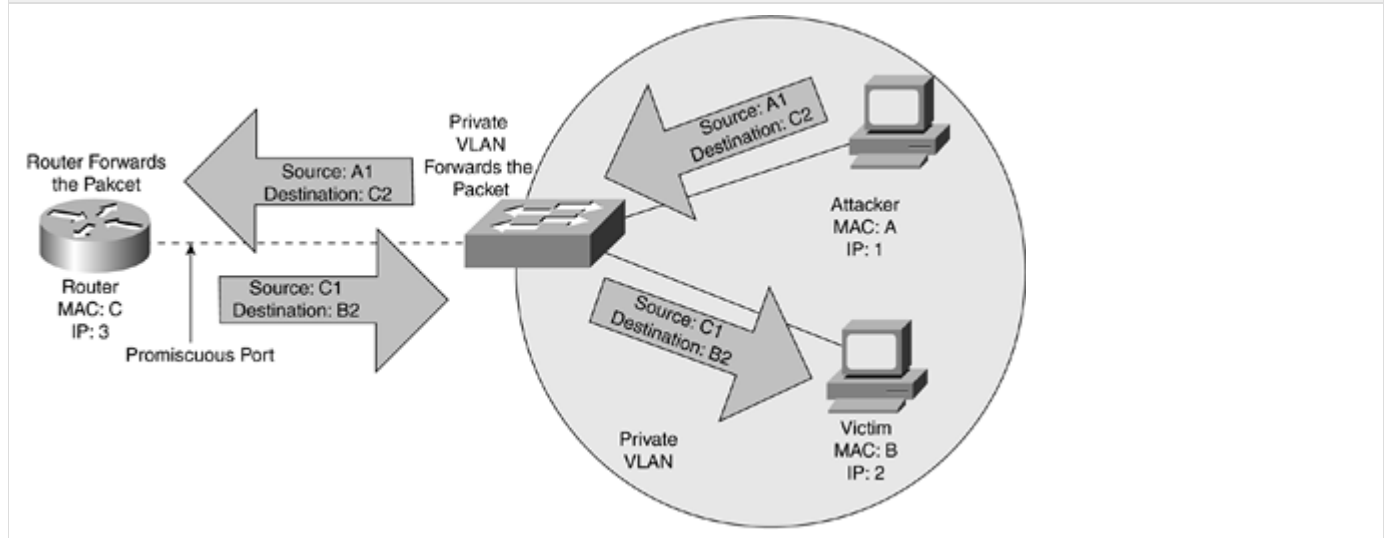
The man-in-the-middle attack occurs when one network device impersonates another network device, such as your **default gateway**. The attacker uses the ARP packets sent to the device targeted for attack because these ARP packets are not verified by the receiver. These ARP packets poison the receiver's ARP table with forged information, injecting the attacker into your network. This attack is similar to identity theft, in which someone obtains a piece of information related to your identity and uses that information to gather more information about you. Eventually, the pretender can convince everyone he is you.

Man-in-the-middle attacks can be prevented either by blocking direct Layer 2 communication between the attacker and the attacked device or by embedding intelligence into your network, such as a Layer 3 device that can check forwarded ARP packets for identity correctness.

## Private VLAN Attack

Private VLANs allow traffic to be further segmented at Layer 2, limiting the size of your broadcast domain. A private VLAN attack uses the expected behavior of a private VLAN against the VLAN itself. Private VLANs are a Layer 2 feature that is supposed to restrict traffic to Layer 2. However, recall that a router is a Layer 3 device and as such, when the router is connected to the [promiscuous port](#) of a private VLAN, the switch forwards all Layer 3 traffic received on that port to whatever destination is identified. This forwarding occurs even if the destination is in the same local network as the source, as illustrated in Figure 9-6.

Figure 9-6. Private VLAN Attack



[View full size image](#)

note Configuring access control lists (ACLs) on the router is a way to prevent private VLAN attacks.

It is normal for two hosts in an isolated VLAN to fail in communicating with each other through direct Layer 2 communication but instead succeed in talking to each other using the router as a packet relay. As it is with regular routed traffic, packets relayed through a Layer 2 proxy can be filtered, if desired, by an appropriately configured ACL on the forwarding device.

## Multicast Brute-Force Attack

Multicast brute-force attacks exploit the [potential vulnerability](#) of a switch to a storm of multicast frames. When a switch receives a significant amount of Layer 2 multicast traffic (frames) in rapid succession, the switch should limit the traffic to its original VLAN; failing to do so would leak frames to other VLANs if there is a routing mechanism in place between the VLANs.

This type of attack often proves ineffective against switches because switches should contain all the frames within their appropriate broadcast domain.

## Spanning-Tree Attack

Another attack that can leverage switch vulnerability is the spanning-tree attack. Recall from [Chapter 7](#), "Spanning Tree Protocol (STP)," that by default STP is turned on and every port on the switch both talks and listens for STP messages on the network. The spanning-tree attack consists of sniffing the network STP frames on the wire and getting the ID of the port on which STP was transmitting.

When the attacker has this port ID information, she can begin sending out STP Configuration/Topology Change Acknowledgement BPDUs (bridge protocol data units) announcing that she (the attacker) is the new root bridge with a much lower priority. This enables the attacker to listen in on all the network traffic and possibly change traffic flow.

## Random Frame-Stress Attack

Random frame-stress attacks can have many incarnations but in general this attack is a brute-force attack, randomly varying several fields of a packet and leaving only the source and destination addresses untouched.

Private VLANs can be used to better isolate hosts at Layer 2 and protect these hosts from unwanted or malicious traffic from untrustworthy devices. Communities of mutually trusting hosts can be created so that a Layer 2 network can be divided into smaller Layer 2 networks where only friendly devices are permitted to communicate with each other.

## Chapter Summary

Network security should be applied to all seven layers of the OSI model; however, this chapter discussed network security from

a Layer 2 (data link layer) perspective. Some basic rules to keep in mind when setting up a secure Layer 2 switch-based network include the following:

- VLANs should be set up so that they clearly separate logical components of your network.
- VLANs are based on the level of security each VLAN requires.
- If any switch ports are not being used, these ports should be placed in a VLAN designed to collect these unused ports.

Using port security on your switch as a security mechanism provides a level of security because port security is based on permitted and denied MAC addresses. Because a MAC address is a hardware address, it lends itself to being a type of physical separation for your network. This differs from using VLANs, which provide more of a logical security for your network. Physical security for your network can be achieved by locking your wiring closets and preventing physical access to your network equipment.

VLANs use logical separation of network components to achieve a level of security in your network. Because VLANs are organized by assigned groups, any host that is not a member of the VLAN is denied access to any of that VLAN's resources. The switch will not recognize that host as part of that VLAN because you did not configure the VLAN to recognize that host.

Port security and VLANs are each susceptible to certain types of network attacks; when used together, however, each provides a level of network security that complements the other. No matter what your comfort level concerning network security, remember that you must take whatever precautions available to protect your network, its resources, and its users from threats both inside and outside your network.

## Practice

- [Configuring Switches](#)
- [Implementing and Tuning Spanning Tree](#)
- [Troubleshooting the LAN Switching Configuration](#)

## Configuring Switches

1. Switch base configuration
  - a. CLI access
  - b. Global switch settings
  - c. Per-vlan settings
2. Access ports settings
  - a. Speed
  - b. Duplex
  - c. Vlan association
3. Trunk ports settings
  - a. Native VlanID
  - b. Vlan range
  - c. Vlan propagation method


## Implementing and Tuning Spanning Tree

1. Default configuration
2. Dedicated (manually assigned) root configuration
3. Timeouts and keepalives tuning
4. STP-RSTP, STP-MSTP migration

## Troubleshooting the LAN Switching Configuration

1. Misconfiguration of Vlan propagation
2. Misconfiguration of ports settings
3. Misconfiguration of timeouts and keepalives

## Hometasks

 Optional part of course. Essay.

1. Layer 2 protocols. Purposes and functionality
2. Functionality and purpose of MAC and LLC sub-layers
3. Differences b/w STP protocols