Operating System Fundamentals

Module 8:
TCP/IP-based networks

Agenda

- Fundamentals of Network Communication
- Packets and Frames
- Clients and Servers
- Network Models
- Hubs, Switches and Routers
- TCP/IP
- IP Addressing

Fundamentals of Network Communication

- A computer network consists of two or more computers connected by some kind of transmission medium, such as a cable or air waves.
- In order to access the Internet, a computer has to be able to connect to a network
- The next few slides will cover what is required to turn a standalone computer into a networked computer

Network Components

- Hardware components
  - *Network interface card*—A NIC is an add-on card that's plugged into a motherboard expansion slot and provides a connection between the computer and the network.
  - *Network medium*—A cable that plugs into the NIC and makes the connection between a computer and the rest of the network. Network media can also be the air waves, as in wireless networks.
  - *Interconnecting*—Interconnecting devices allow two or more computers to communicate on the network without having to be connected directly to one another.

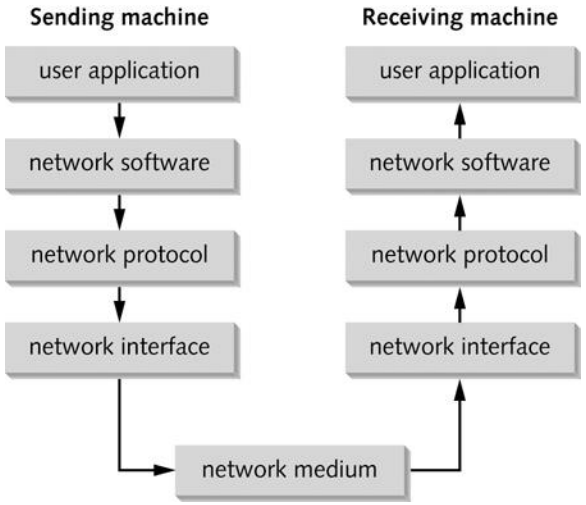
A typical network

Network Components

- Software Components
  - *Network clients and servers*—**Network client software** requests information that's stored on another network computer or device. **Network server software** allows a computer to share its resources by fielding resource requests generated by network clients.
  - *Protocols*—**Network protocols** define the rules and formats a computer must use when sending information across the network. Think of it as a language that all devices on a network understand.
  - *NIC drivers*—NIC drivers receive data from protocols and then forward this data to the physical NIC, which transmits data onto the medium.

Layers of the Network Communication Process

- Each step required for a client to access network resources is referred to as a “layer”
- Each layer has a task and all layers work together



How Two Computers Communicate

- TCP/IP is the most common protocol (language) used on networks
- TCP/IP uses 2 addresses to identify devices on a network
  - Logical address (called IP address)
  - Physical address (called MAC address)
- Just as a mail carrier needs an address to deliver mail, TCP/IP needs an address in order to deliver data to the correct device on a network
- Think of the Logical address as a zip code and the Physical address as a street address

Communication Between Two Computers

1. A user at Comp A types ping 10.1.1.2 at a command prompt
2. The network software creates a ping message
3. The network protocol packages the message by adding IP address of sending and destination computers and acquires the destination computer's MAC address
4. The network interface software adds MAC addresses of sending and destination computers and sends the message
5. Comp B receives message, verifies that the addresses are correct and then sends a reply to Comp A using Steps 2 – 4

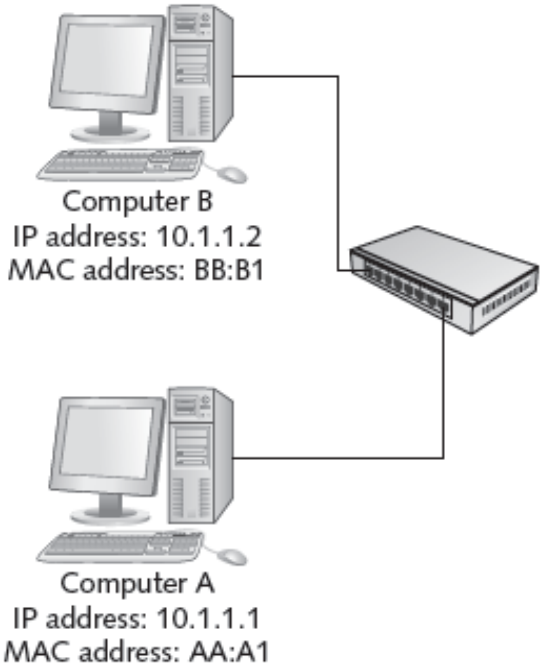


Figure 1-7 Communication between two computers

LANs, Internetworks, WANs, and MANs

- Local area network (LAN) – small network, limited to a single collection of machines and connected by one or more interconnecting devices in a small geographic area

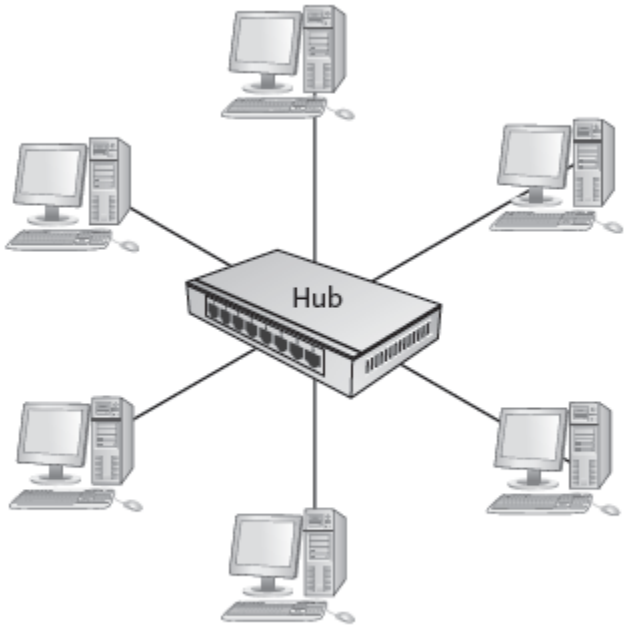


Figure 1-13 A LAN with computers interconnected by a hub

LANs, Internetworks, WANs, and MANs

- An internetwork is a networked collection of LANs tied together by devices such as routers
- Reasons for creation:
  - Two or more groups of users and their computers need to be logically separated but still need to communicate
  - Number of computers in a single LAN has grown and is no longer efficient
  - The distance between two groups of computers exceeds the capabilities of most LAN devices

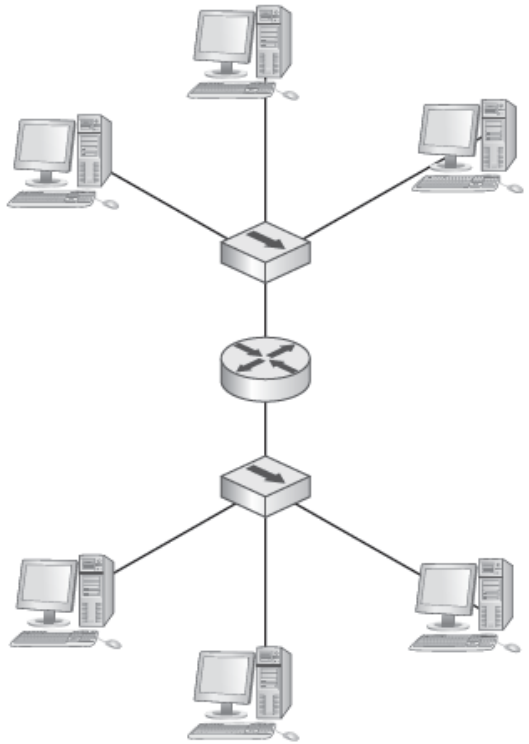


Figure 1-16 An internetwork with two LANS connected by a router

LANs, Internetworks, WANs, and MANs

- Wide area networks (WANs) use the services of third-party communication providers to carry network traffic from one location to another
- Metropolitan area networks (MANs) use WAN technologies to interconnect LANs in a specific geographic region, such as a county or city

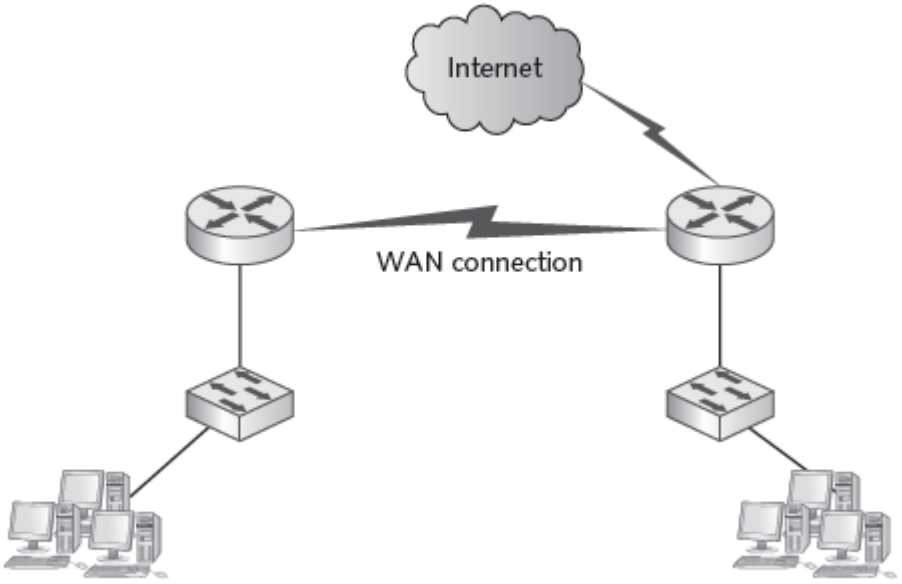


Figure 1-17 A WAN with a connection to the Internet

Packets and Frames

- Computers transfer information across networks in shorts bursts of about 1500 bytes of data
- Data is transferred in this way for a number of reasons:
  - The pause between bursts might be necessary to allow other computers to transfer data during pauses
  - The pause allows the receiving computer to process received data, such as writing it to disk
  - The pause allows the receiving computer to receive data from other computers at the same time
  - The pause gives the sending computer an opportunity to receive data from other computers and to perform other processing tasks
  - If an error occurs during transmission of a large file, only the chunks of data involved in the error have to be sent again, not the entire file

Packets

- Chunks of data sent across the network are usually called packets or frames, with packets being the more well-known term
- **Packet** is a chunk of data with source and destination IP address added to it
- Using the U.S. mail analogy, you can look at a packet as an envelope that has had the zip code added to the address but not the street address

Frames

- A **frame** is a packet with the source and destination MAC addresses added to it
- The packet is “framed” by the MAC addresses on one end and an error-checking code on the other
- A frame is like a letter that has been addressed and stamped and is ready to go
- The process of adding IP addresses and MAC addresses to chunks of data is called **encapsulation**
- Information added to the front of the data is called a **header** and information added to the end is called a **trailer**

Clients and Servers

- A **client** can be a workstation running a client OS or it can also refer to the network software on a computer that requests network resources from a server
- The word “client” is usually used in these three contexts:
  - Client operating system: The OS installed on a computer
  - Client computer: Primary role is to run user applications and access network resources
  - Client software: The software that requests network resources from server software running on another computer

Clients and Servers

- A computer becomes a **server** when software is installed on it that provides a network service to client computers
- The term “server” is also used in three contexts:
  - Server operating system: When the OS installed on a computer is designed mainly to share network resources and provide other network services
  - Server computer: When a computer’s primary role in the network is to give client computers access to network resources and services
  - Server software: Responds to requests for network resources from client software running on another computer

Network Models

- A **network model** defines how and where resources are shared and how access to these resources is regulated
- Fall into two major types
  - **Peer-to-peer network**: Most computers function as clients or servers (no centralized control over who has access to network resources)
  - **Server-based network**: Certain computers take on specialized roles and function mainly as servers, and ordinary users' machines tend to function mainly as clients

Peer-to-Peer/Workgroup Model

- Computers on a peer-to-peer network can take both a client and a server role
- Any user can share resources on his/her computer with any other user's computer
- Every user must act as the administrator of his/her computer
  - Can give everyone else unlimited access to their resources or grant restricted access to other users
  - Usernames and passwords (credentials) are used to control that access

Peer-to-Peer/Workgroup Model

- Problems with Peer-to-peer networks:
  - Must remember multiple sets of credentials to access resources spread out over several computers
  - Desktop PCs and the OSs installed on them aren't made to provide network services as efficiently as dedicated network servers
  - Data organization: If every machine can be a server, how can users keep track of what information is stored on which machine?
- Peer-to-peer networks are well suited for small organizations that have small networks and small operating budgets

Server/Domain-Based Model

- Server-based networks provide centralized control over network resources
- Users log on to the network with a single set of credentials maintained by one or more servers running a server OS
- In most cases, servers are dedicated to running network services and should not be used to run user applications

Server/Domain-Based Model

- A **domain** is a collection of users and computers whose accounts are managed by Windows servers called **domain controllers**
- Users and computers in a domain are subject to network access and security policies defined by a network administrator
  - The software that manages this security is referred to as a **directory service**
  - On Windows servers, the directory service software is **Active Directory**

Server/Domain-Based Model (cont.)

- Other network services usually found on network servers:
  - Naming services: Translate computer names to their address
  - E-mail services: Manage incoming and outgoing email
  - Application services: Grant client computers access to complex applications that run on the server
  - Communication services: Give remote users access to a network
  - Web services: Provide comprehensive Web-based application services

Server/Domain-Based Model (cont.)

- Server-based networks are easier to expand than peer-to-peer
  - Peer-to-peer should be limited to 10 or fewer users, but server-based networks can handle up to thousands of users
- Multiple servers can be configured to work together, which can be used to run a more efficient network or can provide fault tolerance
- Peer-to-peer and server-based networks both have advantages so using a combination of the two models isn't uncommon

Strengths and Weaknesses of the Two Network Models

Network attribute	Peer-to-peer network	Server-based network
Resource access	Distributed among many desktop/client computers; Makes access to resources more complex	Centralized on one or more servers; streamlines access to resources
Security	Users control their own shared resources and might have several sets of credentials to access resources; not ideal when tight security is essential	Security is managed centrally, and users have a single set of credentials for all shared resources; best when a secure environment is necessary
Performance	Desktop OS not tuned for resource sharing; access to shared resources can be hindered by users running applications	Server OS tuned for resource sharing; servers are usually dedicated to providing network services
Cost	No dedicated hardware or server OS required, making initial costs lower; lost productivity caused by increasing complexity can raise costs in the long run	Higher upfront costs because of dedicated hardware and server OSs; additional ongoing costs for administrative support

Network Servers

- A server is at the heart of any network that is too large for a peer-to-peer configuration
- A single server can be configured to fill a single role or several roles at once
- Most common server roles found on networks:
  - Domain controller/directory servers
  - File and print servers
  - Application servers
  - Communication servers
  - E-mail/fax servers
  - Web servers

Network Servers (cont.)

- Domain Controller/Directory Servers
  - Directory services make it possible for users to locate, store, and secure information about a network and its resources.
  - Windows servers permit combining computers, users, groups, and resources into domains. The server handling the computers and users in a domain is called a domain controller.
- File and Print Servers
  - Provide secure centralized file storage and sharing and access to networked printers.
  - Any Windows or Linux computer can act as a file and print server, however the Server version of Windows provides advanced sharing features.

Network Servers (cont.)

- Application Servers
  - Supply the server side of client/server applications to network clients
  - Differ from basic file and print servers by providing processing services as well as handling requests for file or print services
- Communication Servers
  - Provide a mechanism for users to access a network's resources remotely
  - Enable users who are traveling or working at home to dial in to the network via a modem or their existing Internet connection
- E-mail/Fax Servers

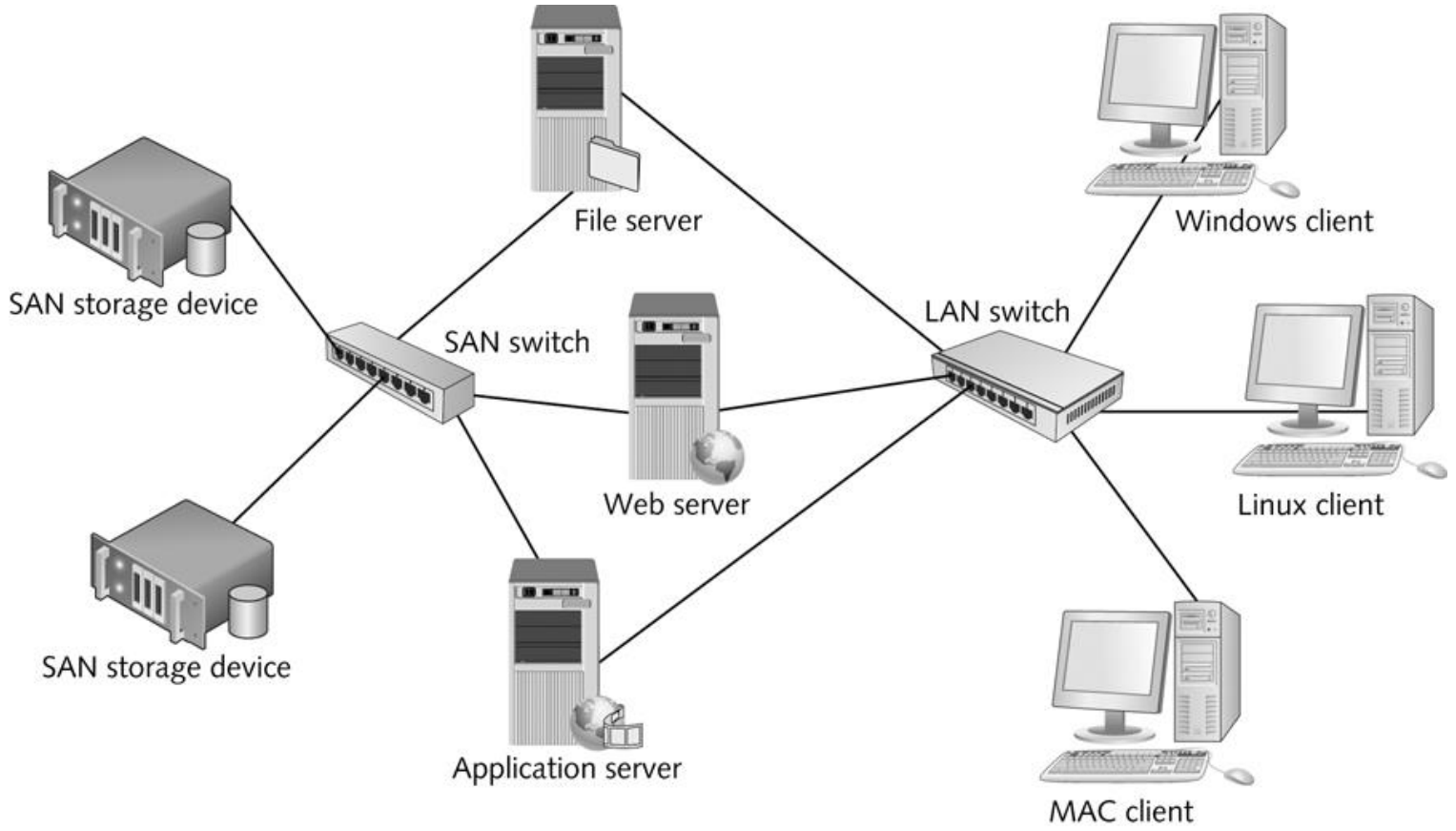
Network Servers (cont.)

- Web Servers
  - Windows Server includes a complete Web server called Internet Information Services (IIS) as well as File Transfer Protocol (FTP)
  - Apache Web Server is available as a part of most Linux distributions and remains the most widely used Web server in the world
- Other Network Services
  - Most networks require additional support services to function efficiently. The most common are Domain Name System (DNS) and Dynamic Host Configuration Protocol (DHCP)
  - DNS allows users to access both local and Internet servers by name rather than by address
  - DHCP provides automatic addressing for network clients so that network administrators do not have to assign addresses manually
  - You will learn more about these services in Chapter 5

Specialized Networks

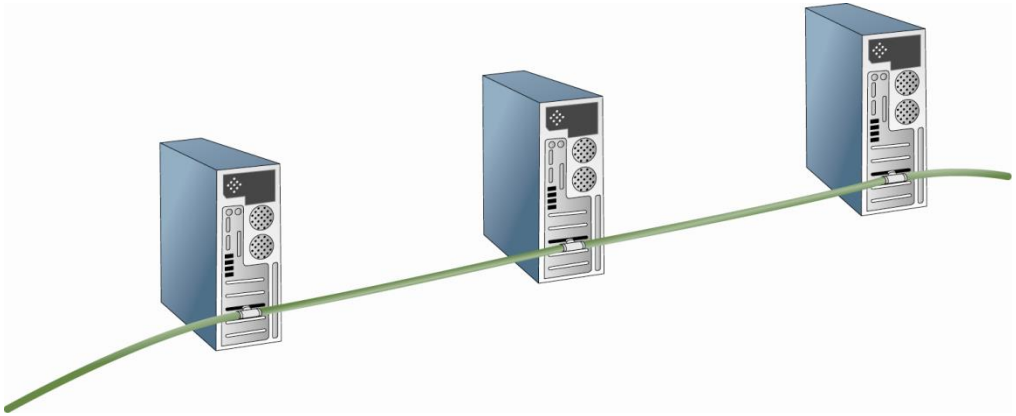
- **Storage area network (SAN)** – uses high-speed networking technologies to provide servers with fast access to large amounts of disk storage
- **Wireless personal area network (WPAN)** – short-range networking technology designed to connect personal devices to exchange information
  - These devices include cell phones, pagers, personal digital assistants (PDAs), global positioning system (GPS) devices, MP3 players, and even watches

SAN



Network Repeaters and Hubs

- Early networks didn’t use interconnecting devices

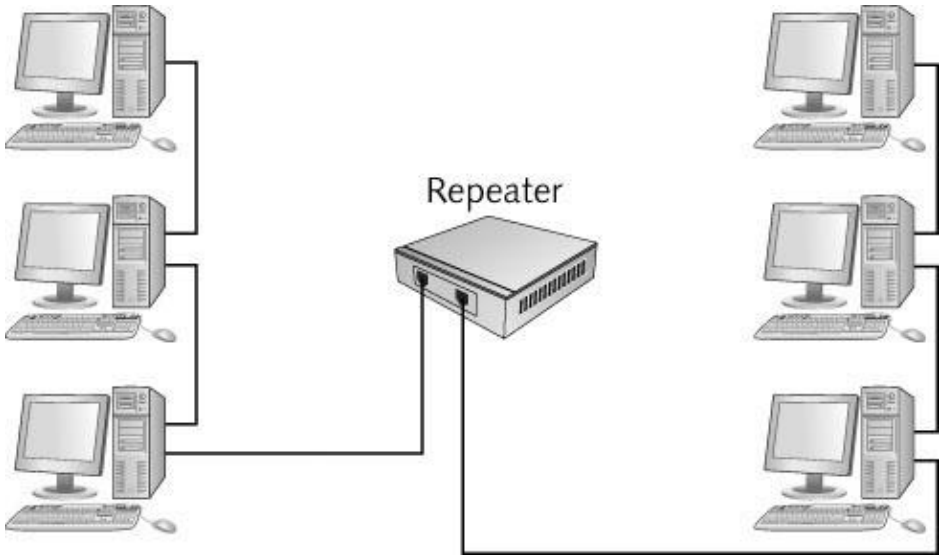


- Severely limited the total cable length and number of computers

Network Repeaters and Hubs

- Some problems were resolved with a device called a repeater
  - A repeater receives bit signals generated by NICs and other devices, strengthens them, and then “repeats” them to other parts of the network
- A repeater enables you to connect computers whose distance from one another would make communication impossible
- A traditional repeater has two ports or connections that you can use to extend your network

Network Repeaters and Hubs

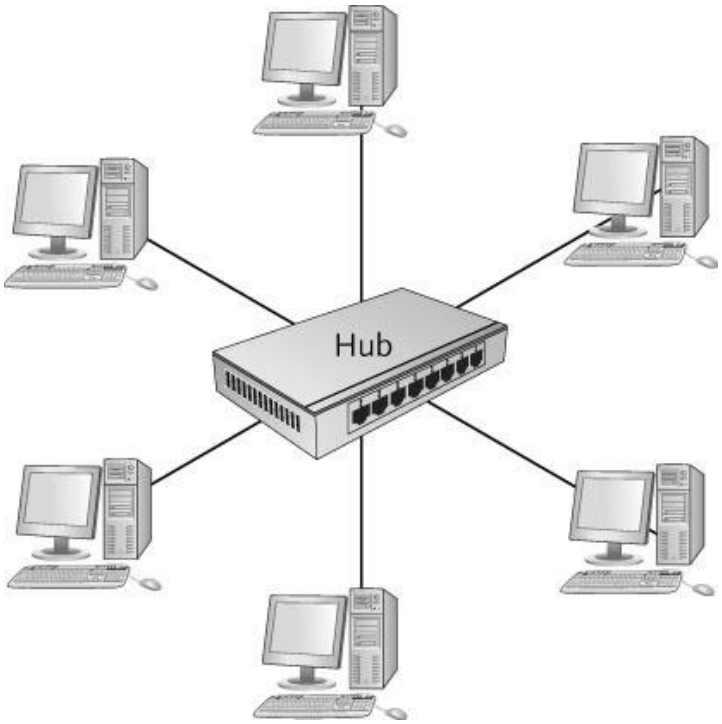


A repeater extends the length of a network

Multiport Repeaters and Hubs

- A multiport repeater is just a repeater with several ports to which you can connect cabling
- Also referred to as a **hub**
- Receives bit signals generated from a connected computer on one of its ports
- Cleans the signal by filtering out electrical noise
- Regenerates the signal to full strength
- Transmits the regenerated signal to all other ports a computer (or other network device) is connected to

Multiport Repeaters and Hubs



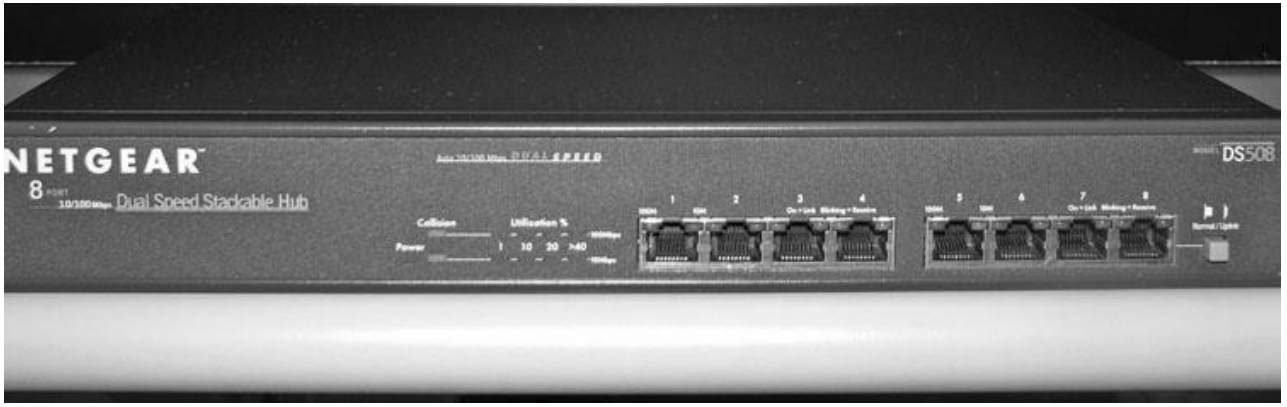
A multiport repeater or hub

Hubs and Network Bandwidth

- Amount of data that can be transferred in an interval is network bandwidth
  - Network bandwidth is usually measured in bits per second (bps) and networks operate at speeds from 10 million bps up to 10 gigabit per second (Gbps)
- Hubs share bandwidth with all other connected computers
  - Only one computer can successfully transmit data at a time
  - Creates a problem on today's networks due to an increased dependency on file sharing and the Internet
- **Bandwidth sharing** – when all computers connected to the hub must share the amount of bandwidth the hub provides

Hub Indicator Lights

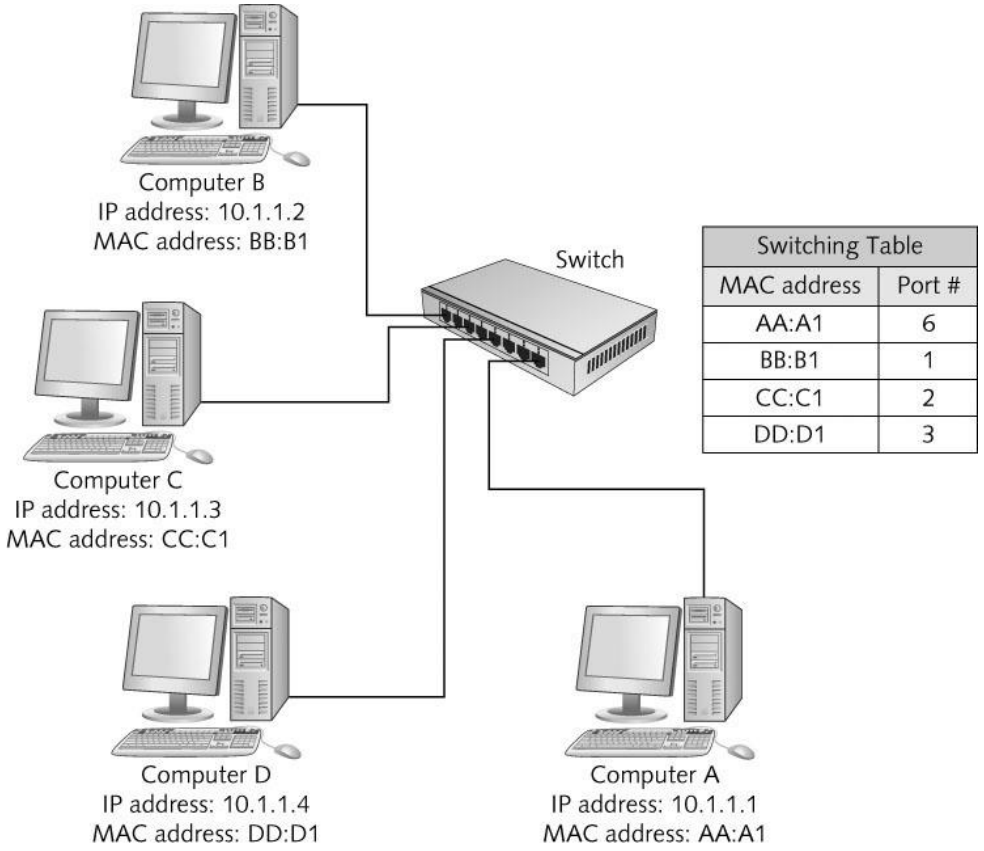
- Power, link status, network activity, collisions
- Uplink port – port used to connect two hubs together or hub to a switch



Network Switches

- Looks just like a hub
- Instead of simply regenerating incoming bit signals a switch, actually reads data in the message, determines which port the destination device is connected to, and forwards the message to only that port
- Basic Switch Operation
  - Data is sent onto the medium one frame at a time
  - The beginning of each frame has the destination MAC address
  - Switch reads the addresses:
    - Keeps a record of which computer is on which port (**switching table**)
    - Forwards the frame to the port where the destination MAC can be found

Network Switches



Switching Table	
MAC address	Port #
AA:A1	6
BB:B1	1
CC:C1	2
DD:D1	3

Switches maintain a switching table

Network Switches

Steps of switch operation:

1. The switch receives a frame.
2. The switch reads the source and destination MAC addresses.
3. The switch looks up the destination MAC address in its switching table.
4. The switch forwards the frame to the port where the computer owning the MAC address is found.
5. The switching table is updated with the source MAC address and port information.

Switches and Network Bandwidth

- Each port gets **dedicated bandwidth**
  - Instead of having to share bandwidth with all ports
- Many conversations can occur simultaneously
- Can operate in **full-duplex mode**
  - Can send and receive data simultaneously
- Hubs can only operate in **half-duplex mode**
  - Can send or receive (but not both) at one time
- Switches are the preferred device because of these advantages

Switch Indicator Lights

- Like hubs, switches have indicator lights
- Switches have link status indicators and activity indicators
- May also have indicators for whether the switch is operating in full-duplex or half-duplex mode
- Switches can be connected to one another so that your LAN can grow beyond the limitations of ports on a single switch
  - Some switches have a dedicated port for uplinking to another switch

Wireless Access Points

- The heart of a wireless network is the wireless access point (AP)
- APs operate similarly to a hub without wires
- All communication passes through the AP
- Most small business and home networks use a device typically called a wireless router that combines the functions of an AP, a switch, and a router
- Wireless LANs usually attached to wired networks

Wireless Access Points

A wireless router

Basic AP Operation

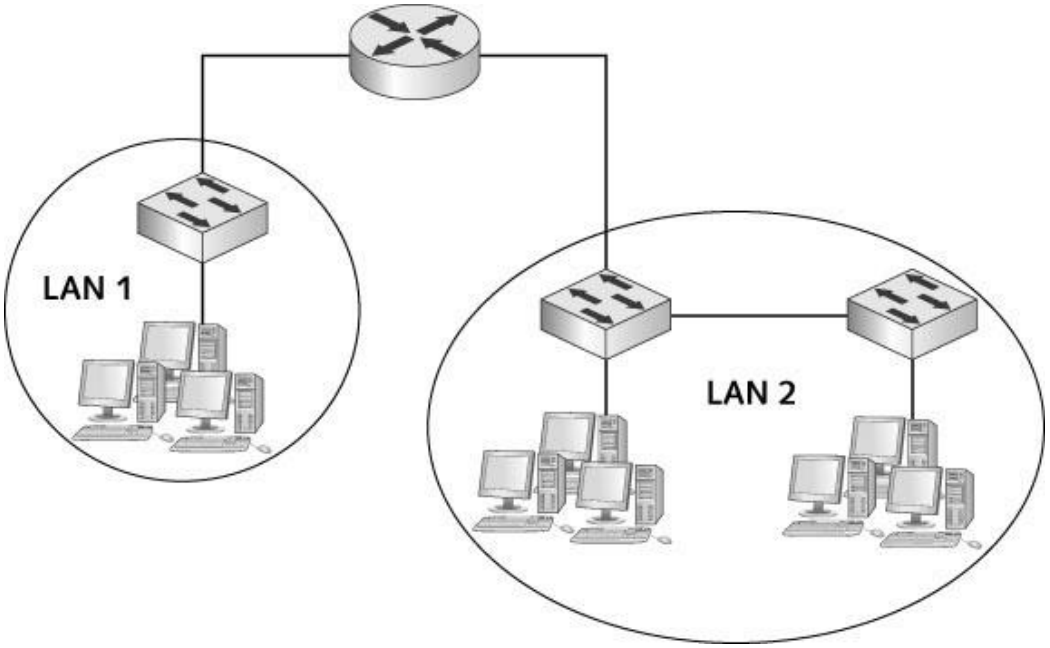
- Similar to a wired hub
- All stations hear all data being transmitted
- Extra step is required: receiving device sends an acknowledgment back to the sending device to indicate successful reception
- Some configurations require additional handshaking:
  - sending station must send a request to send (RTS) message and receive a clear to send (CTS) message before transmitting
  - The RTS/CTS “handshake” lets all other stations know that another station is about to transmit

Wireless APs and Network Bandwidth

- All the extra chatter required to send data in a wireless network slows communication
- The effective bandwidth is about half of effective bandwidth found on physical networks
- Most APs operate from 11Mbps to several hundred Mbps

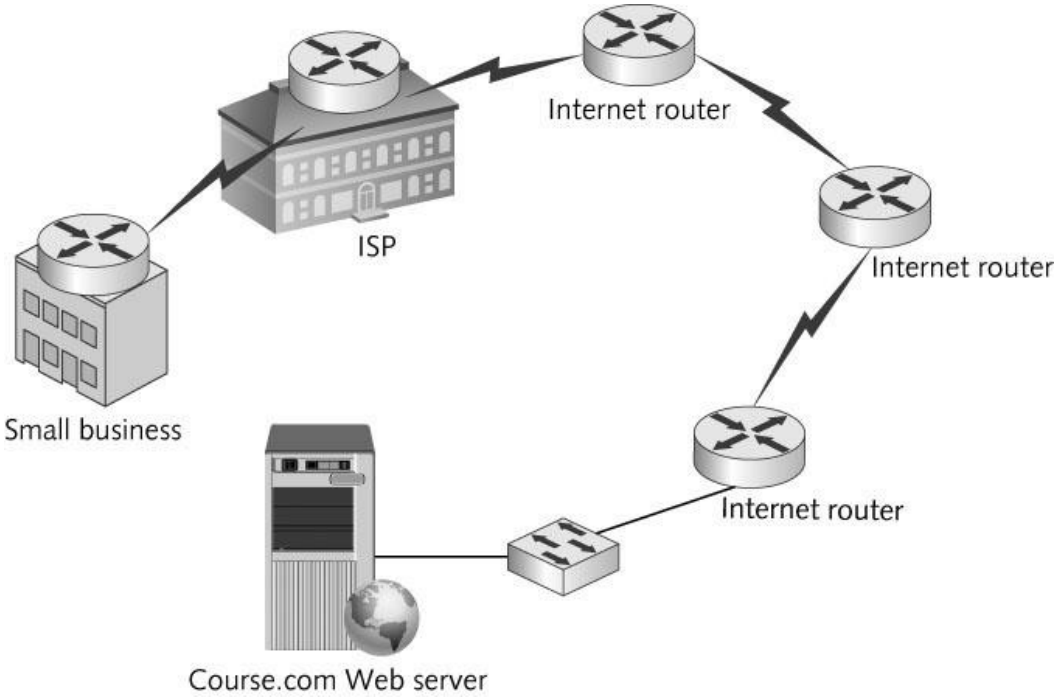
Routers

- Most complex device
- Connect LANs together to create an internetwork



Routers

- **Routers** are devices that enable multiple LANs to communicate with one another by forwarding packets from one LAN to another



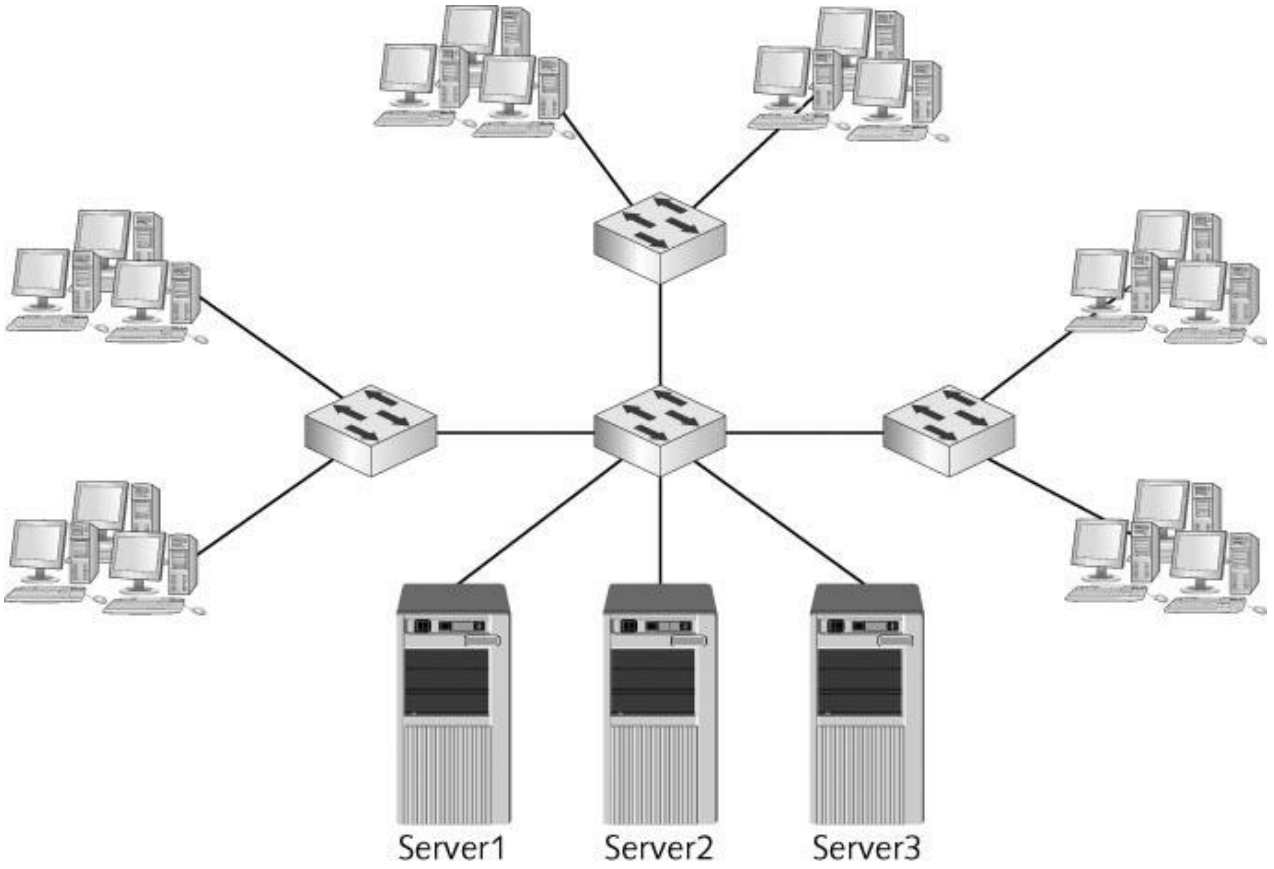
Routers

- The following are the differences between routers and switches
  - Routers connect LANs; switches connect computers
  - Routers work with logical (IP) addresses' switches work with physical (MAC) addresses
  - Routers work with packets; switches with frames
  - Routers don't forward broadcasts; switches do
  - Routers use routing tables; switches use switching tables

Routers Connect LANs

- As computers are added to a LAN, effective communication can suffer
  - Broadcast traffic is forwarded to all members of a LAN and can cause a network to become congested
- The picture on the next slide shows 3 different groups of users and 3 different servers all connected by switches
  - Since they are connected by switches, they are all part of the same LAN and all broadcast traffic will be heard by all devices

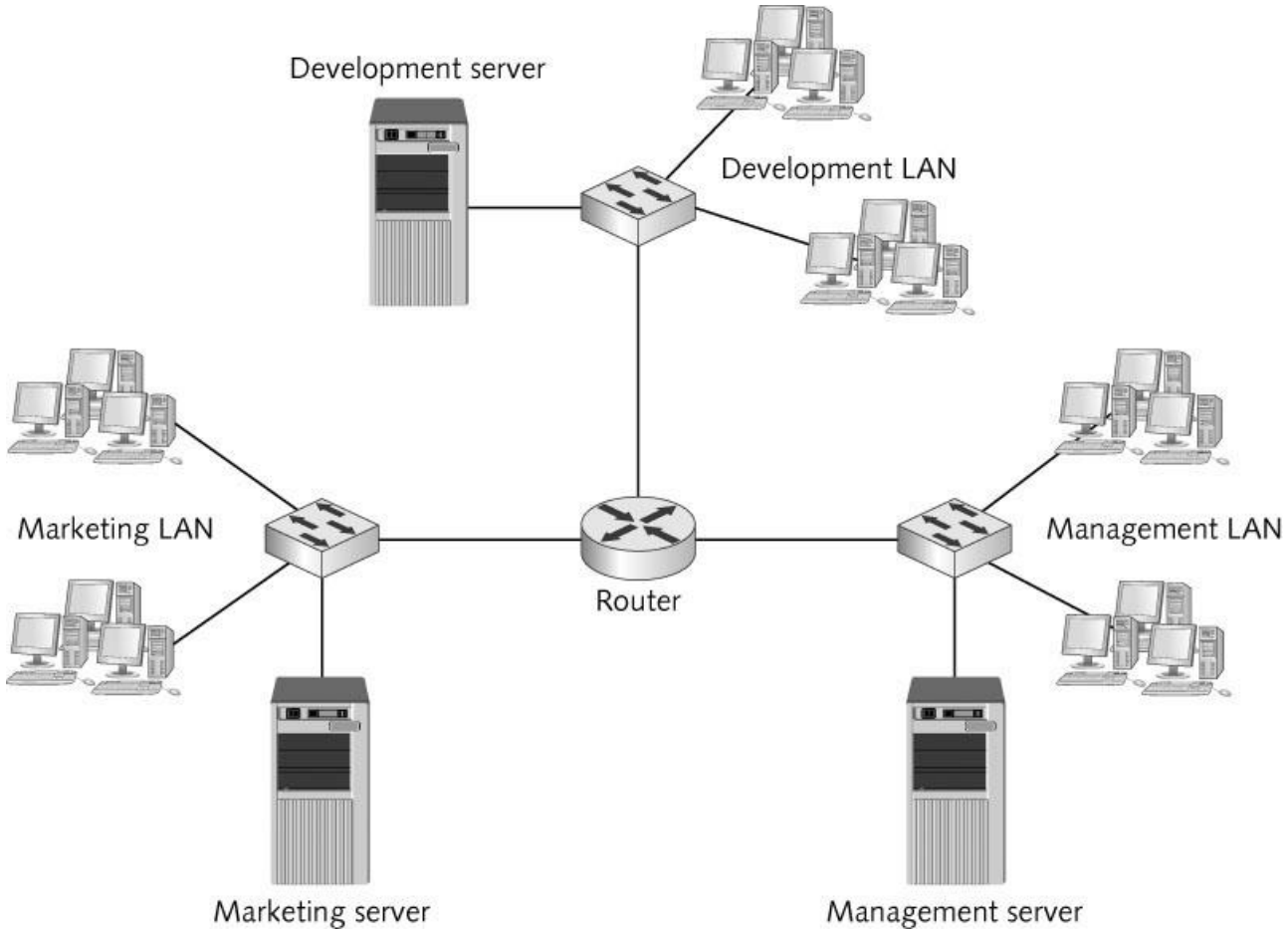
Routers Connect LANs



Routers Connect LANs

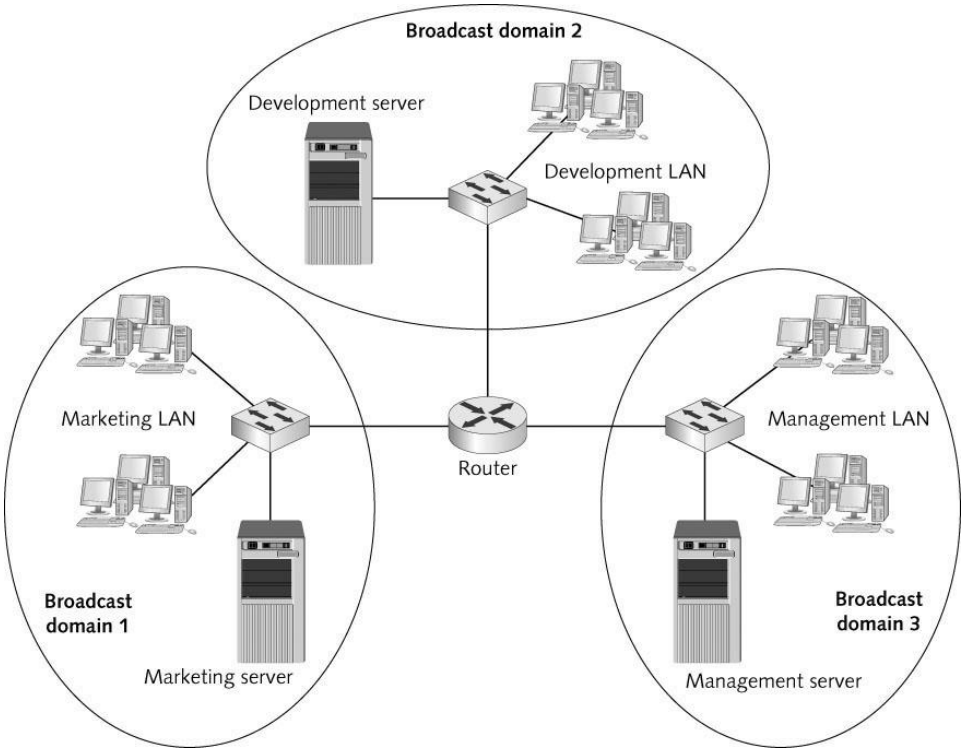
- The picture on the next slide shows a better solution for the previous network
- The administrator groups users and servers together based on their department or function
  - The router is used to create 3 separate LANs in order to contain broadcast traffic and facilitate more effective communication in each department LAN

Routers Connect LANs



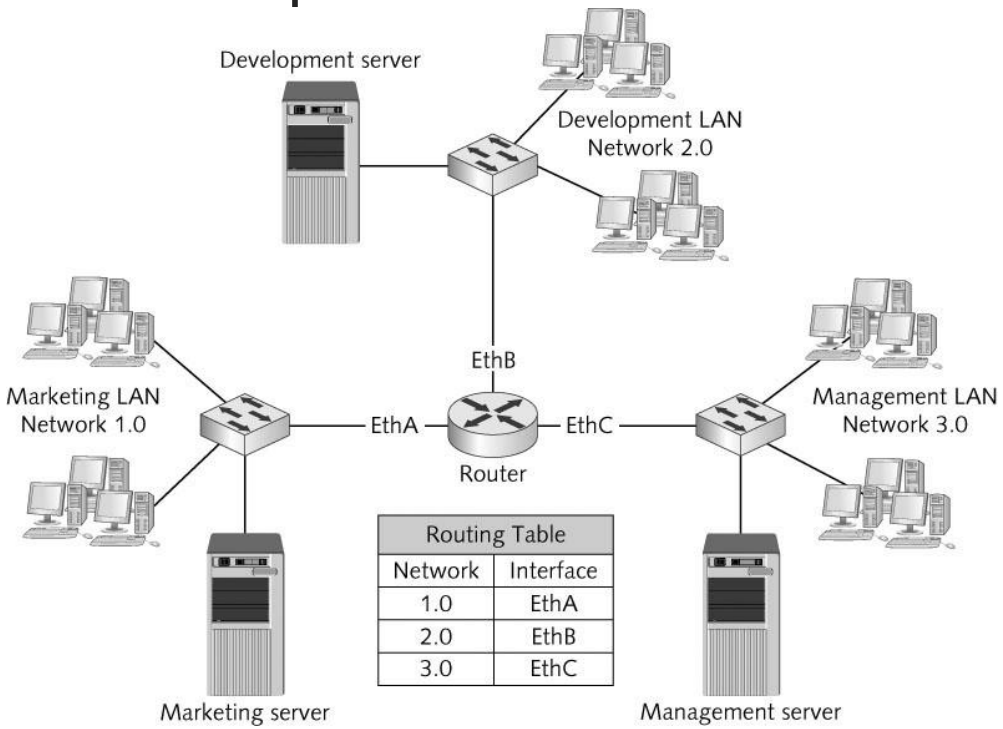
Routers Create Broadcast Domains

- The scope of devices to which broadcast frames are forwarded is called a **broadcast domain**
  - Each router interface in a network creates another broadcast domain



Routers Work with IP Addresses and Routing Tables

- Routers maintain routing tables composed of IP network addresses and interface pairs to determine where to forward packets on an internetwork



Routing Table	
Network	Interface
1.0	EthA
2.0	EthB
3.0	EthC

Routers Work with IP Addresses and Routing Tables

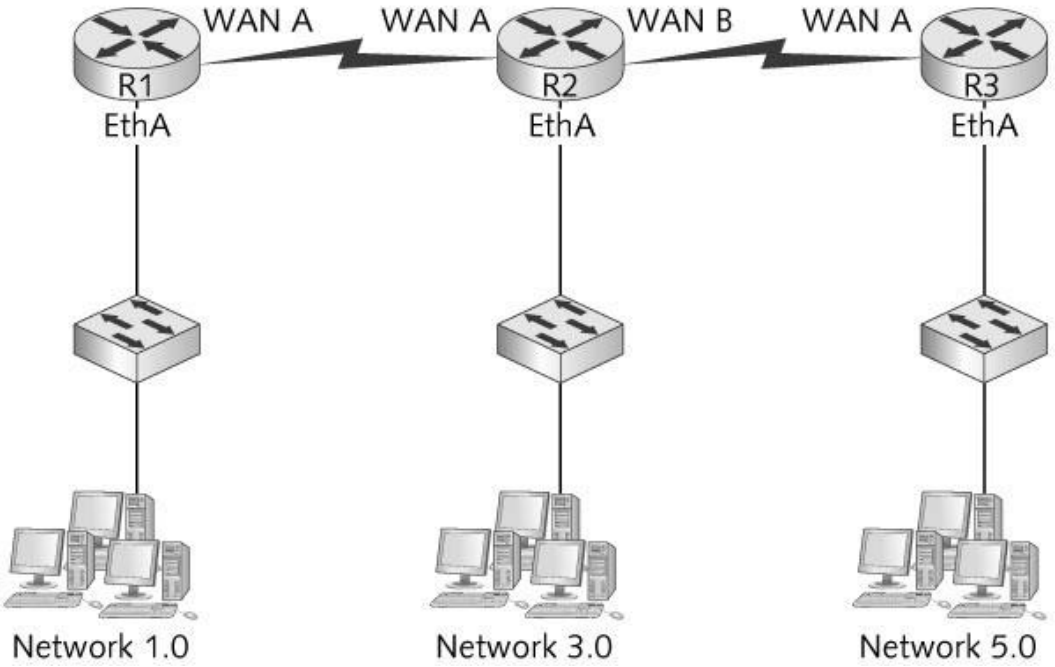
- What happens when a router isn't connected to the network the packet is addressed to?
- The picture on the following slide shows what the routing table would look like on each router between the source and destination networks

Routers Work with IP Addresses and Routing Tables

Routing Table	
Network	Interface
1.0	EthA
3.0	WAN A
5.0	WAN A

Routing Table	
Network	Interface
1.0	WAN A
3.0	EthA
5.0	WAN B

Routing Table	
Network	Interface
1.0	WAN A
3.0	WAN A
5.0	EthA



Routers Work with IP Addresses and Routing Tables

- Default route — where to send a packet when the router doesn't have an entry in its routing table
- Network unreachable — Message sent when the network can't be found and there is no default route
- Default gateway — In a computer's IP address configuration – the IP address of the computer's router

TCP/IP's Layered Architecture

- **Protocols** are rules and procedures for communication and behavior
  - Computers must "speak" the same language and agree on the rules of communication
- When a set of protocols works cooperatively it is called a **protocol stack** or **protocol suite**
- The most common protocol stack is **Transmission Control Protocol/Internet Protocol (TCP/IP)**
- TCP/IP is composed of more than a dozen protocols operating at different levels of the communication process

TCP/IP's Layered Architecture

Layer name	TCP/IP protocols			
Application	HTTP	FTP	DHCP	TFTP
	SMTP	POP3	DNS	SNMP
Transport	TCP		UDP	
Internetwork	ICMP	ARP	IPSec	
	IPv4 and IPv6			
Network access	Ethernet, token ring, FDDI, WAN technologies			

TCP/IP Layered Architecture

TCP/IP's Layered Architecture

- Example of how the layers work together:
  - You start your Web browser and your home page is *http://www.course.com*
  - The web browser formats a request for your home page by using the Application layer protocol HTTP
  - The request looks something like:

```
get the course.com home page
```

  - The unit of information the Application layer works with is simply called "data"

TCP/IP's Layered Architecture

- Example continued:
  - The Application-layer protocol HTTP passes the request down to the Transport-layer protocol (TCP)
  - TCP adds a header to the request that looks like:

TCP header	get the course.com home page

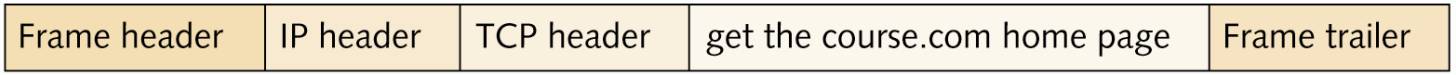
  - The unit of information the Transport layer works with is called a segment
  - TCP passes the segment to the Internetwork layer protocol (IP)
  - IP places its header on the segment:

IP header	TCP header	get the course.com home page

  - The unit of information is now called a packet

TCP/IP's Layered Architecture

- Example continued:
  - The packet is passed down to the Network access layer, where the NIC operates
  - A frame header and trailer are added



  - The frame is then delivered to the network medium as bits on its way to the *www.course.com* server
  - The web server processes it and returns a Web page to the computer that originated the request

Role of the Network Access Layer

- Provides a physical (MAC) address for the network interface
- Verifies that incoming frames have the correct destination MAC address
- Defines and follows media access rules
- Receives packets from the Internetwork layer and encapsulates them to create frames
- De-encapsulates received frames and sends the resulting packets to the Internetwork layer

Role of the Network Access Layer

- Provides frame error detection in the form of a CRC code
- Transmits and receives bit signals
- Defines the signaling needed to transmit bits, whether electrical, light pulses, or radio waves
- Defines the media and connectors needed to make a physical network connection

Role of the Internetwork Layer

- The Internetwork layer is where administrators usually do the most network configuration
- This is where the IP protocol operates and is the heart of the TCP/IP protocol suite
- Responsible for four main tasks:
  - Defines and verifies IP addresses
  - Routes packets through an internetwork
  - Resolves MAC addresses from IP addresses
  - Delivers packets efficiently

Defines and Verifies IP Addresses

- An IP address is assigned to every computer and network device using TCP/IP for communications
- IP addresses are used for two main purposes
  - To identify a network device at the Internetwork layer
  - To identify the network on which a device resides
- When a device receives an IP packet, it compares the destination IP address with its own
  - If it matches or is a broadcast, the packet is processed
  - It is does not match then it is discarded

Routes Packets Through an Internetwork

- The Internetwork layer determines the best way to get a packet from network to network until it reaches its destination
- Most large internetworks, such as the Internet, have multiple paths for getting from one network to another
- Routers work at the Internetwork layer and it is their job to select the best path to the destination
  - Routers use the network identifier portion of IP addresses along with their routing tables to determine the best path

Resolves MAC Addresses from IP Addresses

- Every frame contains source and destination physical MAC and logical IP addresses
- When a packet is ready to be sent to the Network access layer, the destination device's MAC address must be retrieved before the frame header can be constructed
- TCP/IP uses Address Resolution Protocol (ARP) to find MAC addresses
  - ARP is discussed in more detail later in the chapter

Delivers Packets Efficiently

- Internetwork-layer protocols are primarily focused on efficient delivery of packets
  - Internetwork-layer protocols don't include features such as flow control, delivery confirmation, or message assembly
  - These features require overhead to ensure reliable delivery
  - Rely on the protocols in the Transport and Application layers to provide these reliability features
  - Considered a connectionless protocol – relies on upper-layer protocols to ensure the packet's safe journey

Protocols at the Internetwork Layer

- Internet Protocol version 4
  - More commonly known as “IP”
  - Most common version in networks and the first version that was in widespread use
  - Defines a 32-bit dotted decimal IP address
    - Example: 172.31.149.10
    - 172.31 = network id and 149.10 = host id (actual device on the network)
  - Important fields in an IP packet:
    - Version = version of IP in use
    - Time to live (TTL) = prevents a packet from wandering aimlessly through an internetwork
    - Protocol = numeric code specifying the type of IP packet
    - Checksum = value that protects the IP header’s contents
    - Source and Destination address

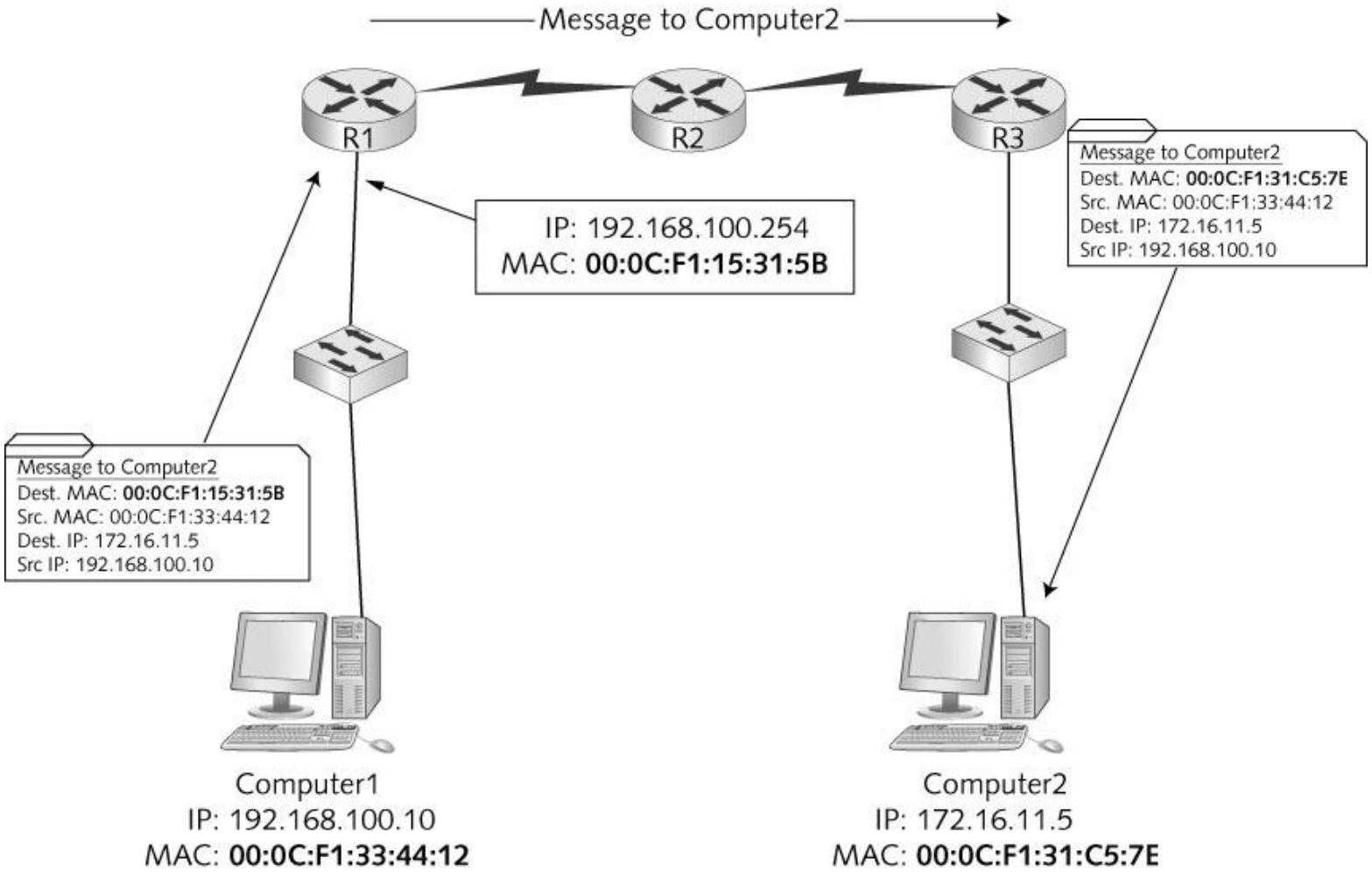
Protocols at the Internetwork Layer

- Internet Protocol version 6 (IPv6)
  - Has many of the features of IPv4
  - Uses a different format for IP addresses
  - Can run alongside IPv4 without needing to change the Transport Layer or Network Access Layer
- Address Resolution Protocol
  - Used to resolve a logical (IP) address to physical (MAC) address
  - When a source doesn't have the destination's MAC, it sends out an ARP broadcast frame requesting the MAC address corresponding to the host's IP address
  - A network configured with the specified IP address responds with an ARP reply containing its MAC address

Protocols at the Internetwork Layer

- Address Resolution Protocol (cont.)
  - To avoid sending an ARP request every time an IP packet is sent, PCs and other devices store learned IP address/MAC address pairs in an **ARP cache**, which is a temporary location in RAM
  - If the destination computer is on another network, the computer uses ARP to retrieve the MAC address of the router configured as its default gateway
    - The packet is delivered to the router and the router determines where the packet should go next to get to its destination
    - When the packet gets to the destination network, the router on the destination network uses ARP to get the destination computer's MAC address

Protocols at the Internetwork Layer



Protocols at the Internetwork Layer

- Internet Control Message Protocol (ICMP)
  - Used to send error and control messages between systems or devices
  - Specialized IP packet with its own header
  - Ping program uses ICMP Echo packets to request a response from another computer or to verify whether it is available for communication
  - An ICMP Reply indicates whether the host is reachable and how long the message's round trip from sender to receiver took

Protocols at the Internetwork Layer

- Internet Protocol Security (IPSec)
  - Works with IPv4 to ensure secure delivery of packets
  - IPSec can be used to secure sensitive network transmissions between computers needing extra security
  - Provides security by using authentication and encryption
  - Requires additional network and computer resources so it should be enabled only for highly sensitive communication and in environments where security risks are high

Role of the Transport Layer

- Transport layer provides reliability needed to handle the unpredictable nature of the Internet
- Two protocols:
  - Transmission Control Protocol (TCP):
    - Connection-oriented and designed for reliable transfer of information in complex internetworks
  - User Datagram Protocol (UDP):
    - Connectionless and designed for efficient communication of generally small amounts of data
  - Both:
    - Work with segments
    - Provide a means to identify the source and destination applications involved in a communication
    - Protect data in the segment with a checksum

Working with Segments

- Both Transport-layer protocols work with units of data called segments
- Both TCP and UDP add a header to data
- The Transport-layer protocol then passes the segment to the Internetwork protocol (IP)
- With incoming data, the Transport-layer receives the segment from the Internetwork protocol, processes it, de-encapsulates it and sends the resulting data up to the Application layer

Identifying Source and Destination Applications

- How do computers keep track of incoming data when a Web browser, email application, chat and a word processing program are all running at the same time?
- TCP and UDP use port numbers to specify the source and destination Application-layer protocols
  - Port numbers are 16-bit values assigned to specific applications running on a computer or network device

Protecting Data with a Checksum

- To protect data integrity, TCP and UDP provide a checksum similar to the CRC in the Network access layer
- Intermediate devices don’t recalculate the checksum in the Transport layer so if data corruption occurs during the transmission, the final receiving station detects the checksum error and discards the data

TCP: The Reliable Transport Layer

- If an application requires reliable data transfer, it uses TCP as the Transport-layer protocol
- TCP provides reliability by using these features:
  - Establishing a connection
  - Segmenting large chunks of data
  - Ensuring flow control with acknowledgments
- TCP is a connection-oriented protocol
  - It establishes a connection with the destination, data is transferred, and the connection is broken

Establishing a Connection: The TCP Handshake

- A client sends a TCP synchronization (SYN) segment to the destination device, usually a server
  - A destination port is specified and a source port is assigned dynamically
- When the server receives the SYN segment, it usually responds by sending either an acknowledgment-synchronization (ACK-SYN) segment or a reset connection (RST) segment
  - RST is sent when the server refuses the request to open the session
  - If an ACK-SYN is returned, the client completes the **three-way handshake** by sending an ACK segment back to the server

Segmenting Data

- When TCP receives data from the Application layer, the size might be too large to send in one piece
- TCP breaks the data into smaller segments (max frame sent by Ethernet is 1518 bytes)
- Each segment is labeled with a sequence number so that if segments arrive out of order they can be reassembled in the correct order

Flow Control with Acknowledgments

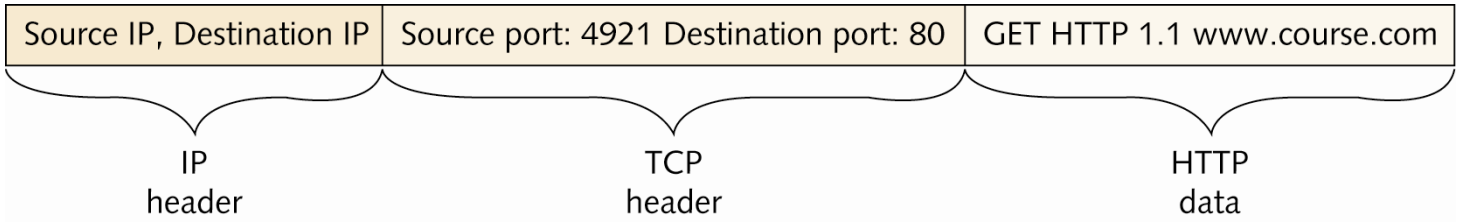
- Flow control prevents a destination from becoming overwhelmed by data, resulting in dropped packets
- TCP establishes a maximum number of bytes, called the window size, that can be sent before the destination must acknowledge the receipt of data
- If no acknowledgment is received within a specified period of time, the sending station will retransmit from the point at which an acknowledgment was last received

Role of the Application Layer

- The Application layer provides network services to user applications that access network resources
- With most Application layer protocols, both a client and a server version exist
- The Application layer provides these functions:
  - Access by applications to network services
  - Client/server data access
  - Name resolution
  - Dynamic address assignment
  - Authentication/user logon
  - Data formatting and translation

HTTP:Protocol of the World Wide Web

- Originally, its main purpose was to transfer static web pages written in HTML
- Now, it is also used for general file transfer and downloading/displaying multimedia files
- Uses TCP as its Transport-layer protocol
- Default TCP port number is 80



POP3, IMAP, and SMTP: E-mail Protocols

- Post Office Protocol version 3 (POP3) is used to download incoming messages from e-mail servers to local desktops (uses TCP port 110)
- Internet Message Access Protocol (IMAP) is used to manage email messages locally, yet stores them on a server (uses TCP port 143)
- Simple Mail Transfer Protocol (SMTP) is the standard protocol for sending email over the Internet (uses TCP port 25)

Dynamic Host Configuration Protocol

- A drawback of using TCP/IP in a large network is keeping track of assigned addresses and to which machine they are assigned
- DHCP is used to automatically assign IP addresses as needed
  - When a computer is turned on, it requests an address from a server that is configured as a DHCP server
  - The server assigned an address for a specific amount of time (called a lease)
  - If the computer is still on and the lease is 87.5% expired, a broadcast DHCP renewal request is sent
- Uses UDP Transport layer because DHCP messages are short in length

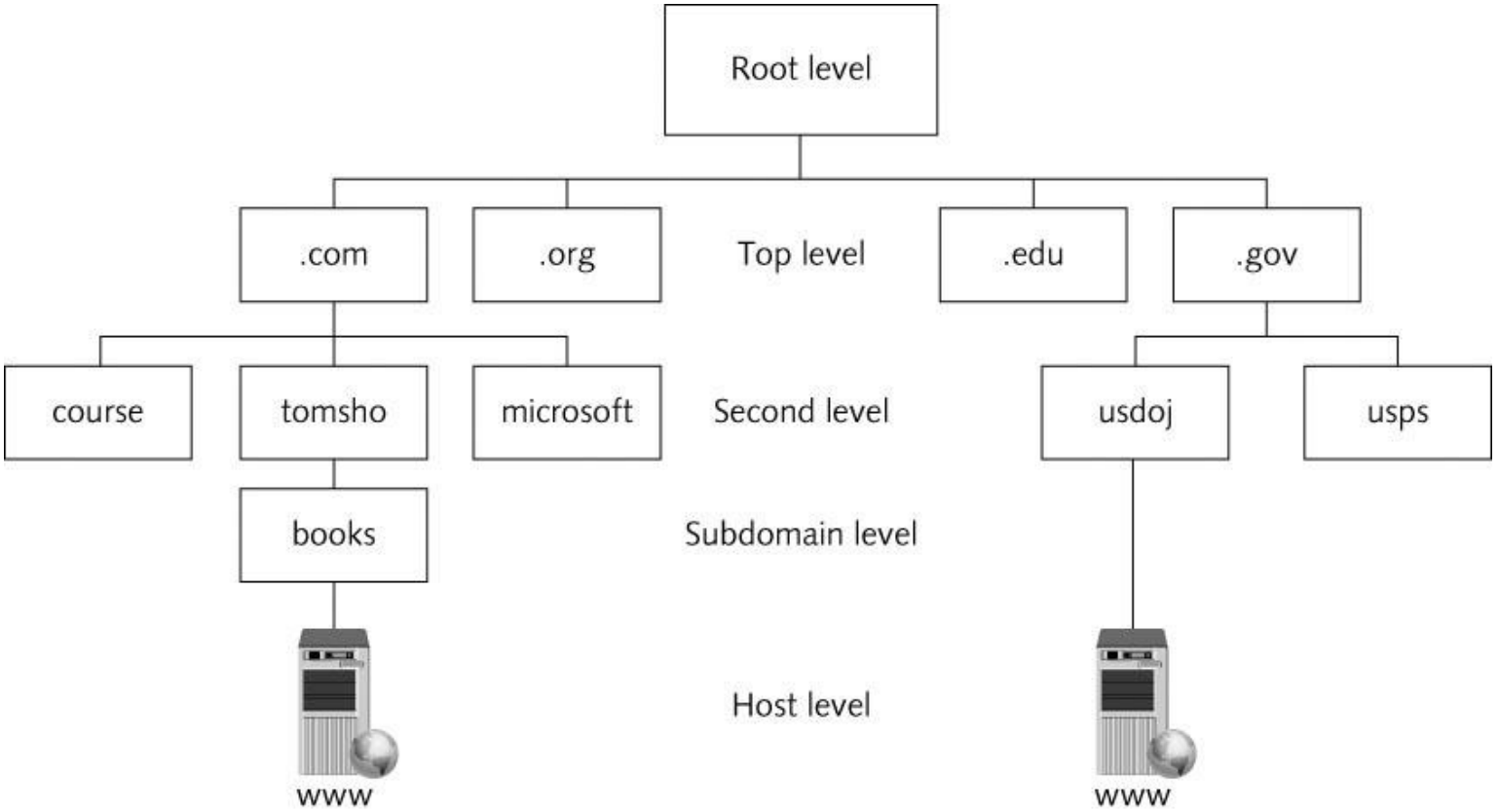
Domain Name System

- DNS is a name-to-address resolution protocol that keeps a list of computer names and their IP addresses
- With a DNS, a user can use a computer's name instead of its IP address
- Example:
  - When you enter www.course.com in your Web browser, the browser contacts the DNS server specified in your OS's IP configuration and requests that the url be resolved to an IP address
  - Once the IP address for the website is returned, your computer can contact Web server to request a Web page
- DNS uses UDP because DNS messages usually consist of a single packet of data

Domain Name System

- DNS is organized as a treelike hierarchy
- Organized into domain levels
  - Top-level domains are organized into categories such as commercial (.com), nonprofit organizations (.org), government (.gov) or country of origin indicated by a two-letter code
  - Second-level domains are usually the name of a company or institution
  - Subdomain is optional and can consist of names separated by a period
  - Host level represents individual computers hosting network services
- Example: www.books.tomsho.com
  - com is the top-level domain name, tomsho is the second-level domain, books is the subdomain, and www is the hostname

Domain Name System



IP Addressing

- IP is responsible for addressing and routing in the TCP/IP environment
- An IP address is 32-bits in length
  - Grouped into four 8-bit octets and each octet is represented by a decimal number from 0-255
  - Four decimal numbers are separated by periods in a format called dotted decimal notation
  - Example: 172.24.208.192
  - Divided into two parts: network ID and host ID
  - In the above address, host ID 208.192 resides on network 172.24

IP Address Classes

- IP addresses are categorized in Classes A-E
  - Only IP addresses in the A, B, and C classes are available for host assignment
- Class A
  - Value of the first octet is between 1 and 127
  - Addresses beginning with 127 are reserved for loopback
  - IP registry assigns the first octet, leaving the last three octets to be assigned to hosts
  - Intended for large corporations and government
- Class B
  - Value of the first octet is between 128 and 191
  - IP registry assigns the first two octets, leaving the third and fourth octets to be assigned to hosts
  - Intended for use in medium to large networks

IP Address Classes

- Class C
  - Value of the first octet is between 192 and 223
  - IP address registry assigns the first three octets
  - These networks are limited to 254 hosts per network
  - Intended for small networks
- Class D
  - Value of the first octet is between 224 and 239
  - Reserved for multicasting
- Class E
  - Value of the first octet is between 240 and 255
  - Reserved for experimental use and can’t be used for address assignment

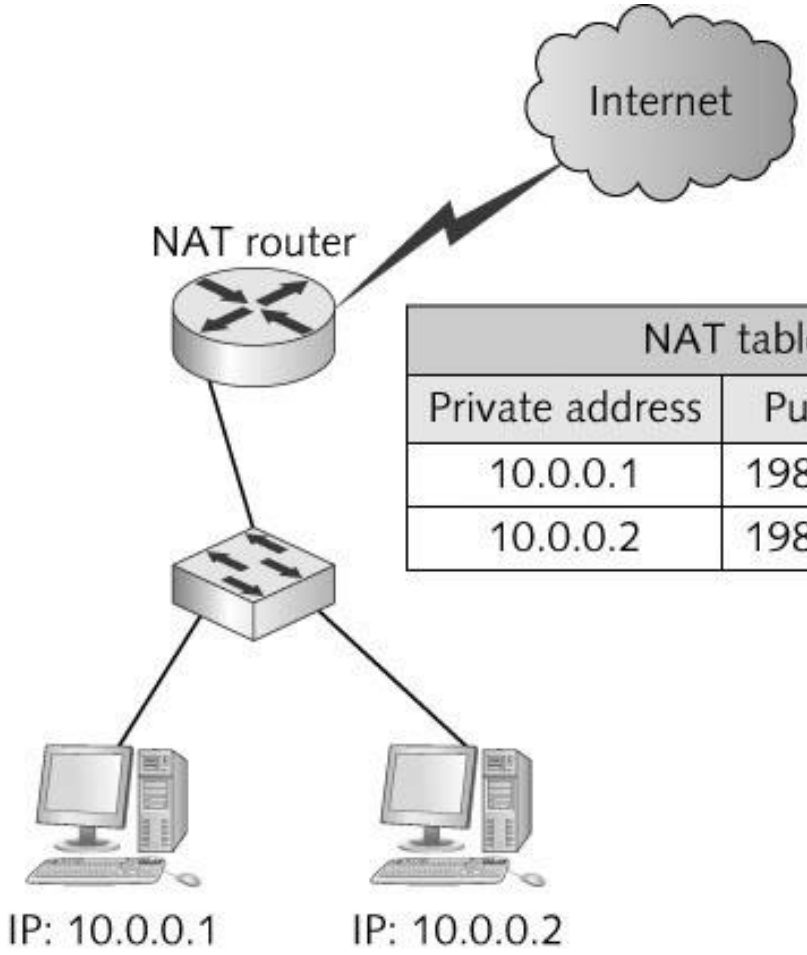
Private IP Addresses

- Due to the popularity of TCP/IP and the Internet, we are running out of unique IP addresses
- A series of addresses have been reserved for private networks (networks whose hosts can't be accessed directly through the Internet)
- Reserved addresses:
  - Class A addresses beginning with 10
  - Class B addresses from 172.16 to 172.31
  - Class C addresses from 192.168.0 to 192.168.255
- The addresses in those ranges can't be routed across the Internet

Network Address Translation

- **NAT** allows an organization to use private IP addresses while connected to the Internet
- The NAT process translates a workstation's private address (as a packet leaves the corporate network) into a valid public Internet address
  - When data returns to the workstation, the address is translated back to the original private address
  - NAT is usually handled by a network device connected to the Internet, such as a router
  - Address translation is kept track of in a NAT table

Network Address Translation



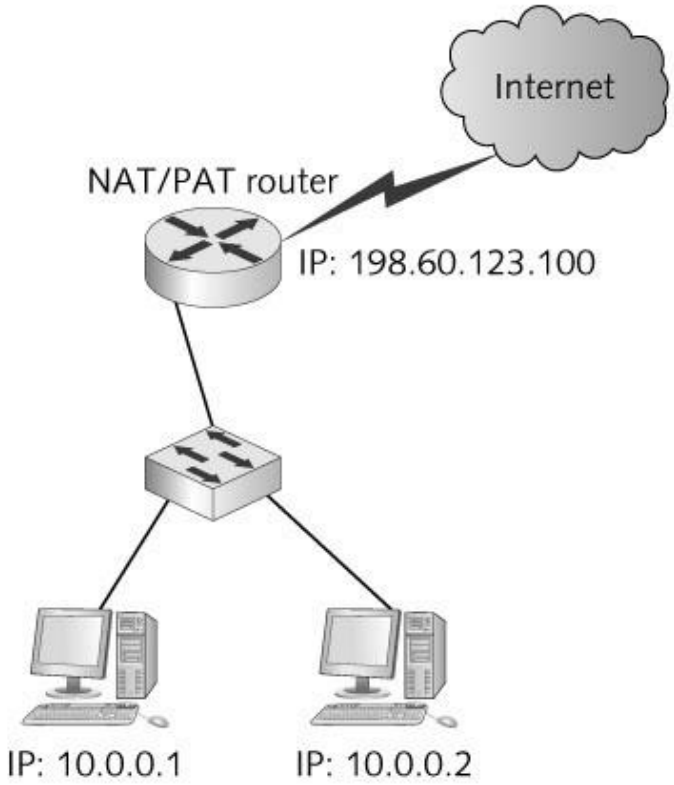
NAT table	
Private address	Public address
10.0.0.1	198.60.123.101
10.0.0.2	198.60.123.102

Network Address Translation

- **Port Address Translation (PAT)**
  - Allows several hundred workstations to access the Internet with a single public Internet address
  - Each packet contains source and destination IP addresses along with source and destination port numbers
  - A single public IP address is used for all workstations, but different source port numbers are used for each communication session
- The next slide shows an example of how PAT is used

Network Address Translation

NAT/PAT table	
Private address: Port	Public address: Port
10.0.0.1:2562	198.60.123.100:5311
10.0.0.2:12441	198.60.123.100:3105



Subnet Masks

- IP uses an address's **subnet mask** to determine which part of the address identifies the network portion and which part identifies the host portion
- Subnet masks are 32-bit numbers in dotted decimal format
  - Default subnet mask for Class A is 255.0.0.0
  - Default subnet mask for Class B is 255.255.0.0
  - Default subnet mask for Class C is 255.255.255.0
- Example: If a computer has the IP address 153.92.100.10 and the subnet mask is 255.255.0.0 then the network portion is 153.92 and the host portion is 100.10

Subnet Masks

- Example (continued): Using the same address of 153.91.100.10 but with a subnet mask of 255.255.255.0, the network portion is now 153.92.100 and the host portion is 10
- By altering the subnet mask, the network ID has been altered

Value of first octet	Class	Default subnet mask	Number of hosts/network
1–127	A	255.0.0.0	16,777,214
128–191	B	255.255.0.0	65,534
192–223	C	255.255.255.0	254

How is the Subnet Mask Used?

- Here's what happens when Computer1 has a packet to send to Computer2:
  - Computer1 must first know its network address. It determines this by doing a logical AND operation between its IP address and subnet mask. A logical AND is an operation between two binary values. AND operations can have the following results:

    0 AND 0 = 0

    1 AND 0 = 0

    0 AND 1 = 0

    1 AND 1 = 1

Subnet Masks

- Example (continued)
  - The logical AND operation between Computer1's IP address and subnet mask looks like this:

    10101100.00010011.00101110.10111100 (binary for 172.19.44.211)

    AND

    11111111.11111111.11111111.00000000 (binary for 255.255.255.0)

    _____

    10101100.00010011.00101110.00000000 (binary for 172.19.44.0)

    The resulting network address is 172.19.44.0.
  - The next step is to determine whether Computer2's address is on the same network by performing the same AND calculation between Computer 2's IP address and Computer1's subnet mask
  - If Computer2 is on a different network, Computer1 knows that the packet must be sent to the router, which forwards it to the correct network

Binary Arithmetic

- Review how the decimal number system works
  - 0 through 9 are used to represent any possible number
  - Each place in a decimal number can 10 possible values
  - The ones place can be expressed as a number 0 through 9, multiplied by 10 raised to the 0 power or 10^0 (any number raised to the 0 power equals 1)
  - The decimal number 249 can be expressed as either of the following:

    $2 * 10^2 + 4 * 10^1 + 9 * 10^0 = 249$

    $2* 100 + 4 * 10 + 9 * 1 = 249$

Binary Arithmetic

- With binary arithmetic, there are only 2 possible values (1 or 0)
- For example, using the same method you used to solve the decimal example, you can express the binary number 101 as either of the following. The numbers in bold are the binary digits.

  1 * 2^2 + **0** * 2^1 + **1** * 2^0 = 5

  1 * 4 + **0** * 2 + **1** * 1 = 5

Binary Arithmetic

128	64	32	16	8	4	2	1
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
0	**1**	**1**	**1**	**1**	**1**	**0**	**1**

Decimal to binary conversion table = Convert 125 to binary:

• 125 is less than 128, so you place a **0** in the column under the 128. The test number remains 125.
• 125 is greater than 64, so you place a **1** in the column under the 64 and subtract 64 from 125, leaving your new test number as 61.
• 61 is greater than 32, so you place a **1** in the column under the 32 and subtract 32 from 61, leaving your new test number as 29.
• 29 is greater than 16, so you place a **1** in the column under the 16 and subtract 16 from 29, leaving your new test number as 13.
• 13 is greater than 8, so you place a **1** in the column under the 8 and subtract 8 from 13, leaving your new test number as 5.
• 5 is greater than 4, so you place a **1** in the column under the 4 and subtract 4 from 5, leaving your new test number as 1.
• 1 is less than 2, so you place a **0** in the column under the 2.
• 1 is equal to 1, so you place a **1** in the column under the 1 and subtract 1 from 1, leaving your new test number as 0. When your test number is 0, you're done.

Binary Arithmetic

Binary	Decimal
00000000	0
10000000	128
11000000	192
11100000	224
11110000	240
11111000	248
11111100	252
11111110	254
11111111	255

- Many of the numbers you work with when subnetting have patterns
- The subnet mask always consists of a series of zero or more 1s, followed by a series of zero or more 0s, as the table above shows

Calculating a Subnet Mask

- There are usually two approaches to subnetting, and they depend on the answer to these questions:
  - Am I subnetting to provide a network with a certain number of host addresses?
  - Or am I subnetting to provide a network with a certain number of logical subnets?
  - If you're working for an ISP, the answer is usually yes to the first question, and if you’re a network administrator for a corporate network, the answer is more likely to be yes to the second question.
  - Sometimes the answer is a combination of both.

Calculating a Subnet Mask

- Subnetting Example: You have a large internetwork and need to break an IP address space into several subnets. Follow this process:
- First, decide how many subnets you need. Each router interface connection indicates a required subnet.
- Decide how many bits you need to meet or exceed the number of required subnets.
- Use the formula 2^n, with *n* representing the number of bits you must reallocate from the host ID to the network ID.
- The number of subnets you create is always a power of 2, so if you need 20 subnets, you must reallocate 5 bits ($2^5 = 32$) because reallocating 4 bits gives you only 2^4 or 16 subnets.

Calculating a Subnet Mask

- Subnetting example (continued)
  - Reallocate bits from the host ID, starting from the most significant host bit (that is, from the left side of the host ID).
  - You must also ensure that you have enough host bits available to assign to computers on each subnet. To determine the number of host addresses available, use the formula $2^n - 2$, with *n* representing the number of host (0) bits in the subnet mask.

Supernetting

- Sometimes necessary to solve certain network configuration problems and to make routing tables more streamlined
- Sometimes referred to as “route aggregation” or “route summarization”
- Supernetting reallocates bits from the network portion of an IP address to the host portion
  - Making two or more smaller subnets a larger supernet