The Internet and TCP/IP

Habib Youssef, Ph.D.

youssef@ccse.kfupm.edu.sa

Department of Computer Engineering

King Fahd University of Petroleum & Minerals

Dhahran, Saudi Arabia

TCP/IP and the Internet

- TCP and IP are two of the suite of data communication protocols used on the Internet.
- IP: Internet Protocol.
- TCP: Transmission Control Protocol.
- All hosts connected to the network must speak TCP/IP.

TCP/IP Features

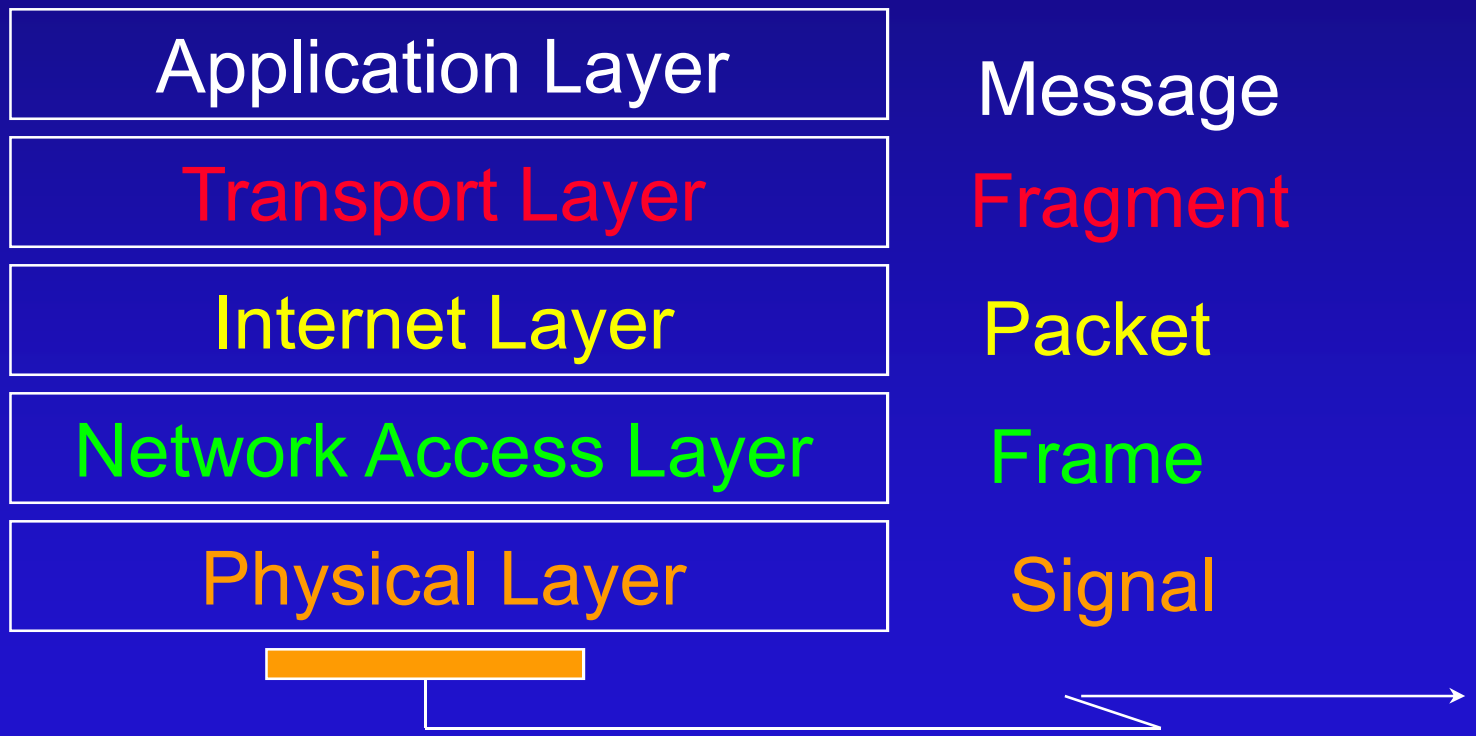
- Popularity of TCP/IP
  - simpler than OSI-ISO standard
  - provides an elegant solution to world wide data communication.
- Open Protocol Standards, freely available, and independent from any hardware platform.

TCP/IP Features (contd.)

- Independence from specific network hardware
  - Allows TCP/IP to integrate many types of networks (Ethernet, Token Ring, X.25)
  - TCP/IP is used in both LANs/ and WANs
  - Supports dial-up connectivity
- Common addressing scheme
  - every TCP/IP host has a unique address
- Standardized high-level protocols for world wide available network services

TCP/IP Protocol Architecture

- **Layered architecture**



Application Layer

- Includes all software programs that use the Transport Layer protocols to deliver data messages
- Examples of protocols:
  - Telnet: Network Terminal Protocol
  - FTP: File Transfer Protocol
  - SMTP: Simple Mail Transfer Protocol
  - DNS: Domain Name Service
  - WWW: World Wide Web

Transport Layer

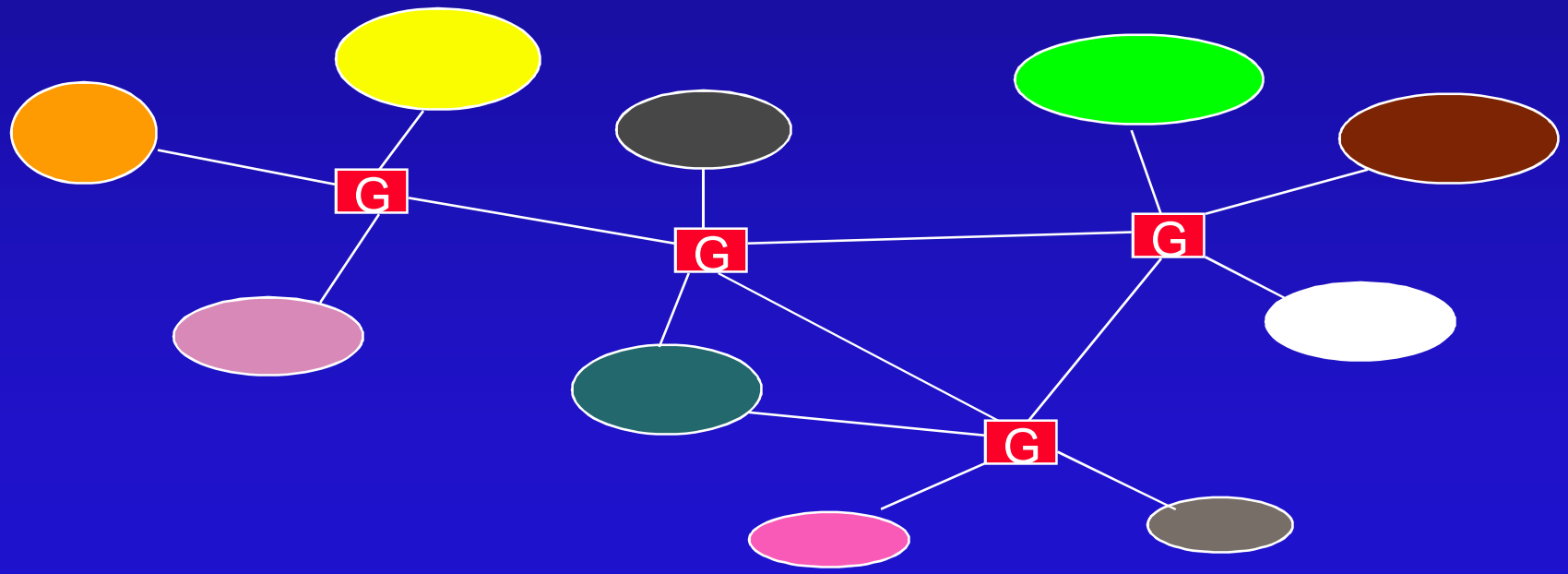
- Interface between the Application and Internet layers
- Two main protocols
  - Transmission Control Protocol (TCP)
    - Provides reliable end-to-end data delivery service
  - User Datagram Protocol (UDP)
    - Provides low overhead connection-less datagram delivery service

Internet Layer

- Heart of TCP/IP
  - Provides basic packet delivery service on which TCP/IP networks are built
- Main functions
  - Defines datagram, basic unit of transmission in the Internet
  - Provides Internet addressing
  - Routing of datagrams
  - No error control

Internetworking (cont.)

- Internet Gateways/Routers are used to connect networks together.
- Gateways have knowledge of internet topology
- Gateways route packets based on destination network not on destination host



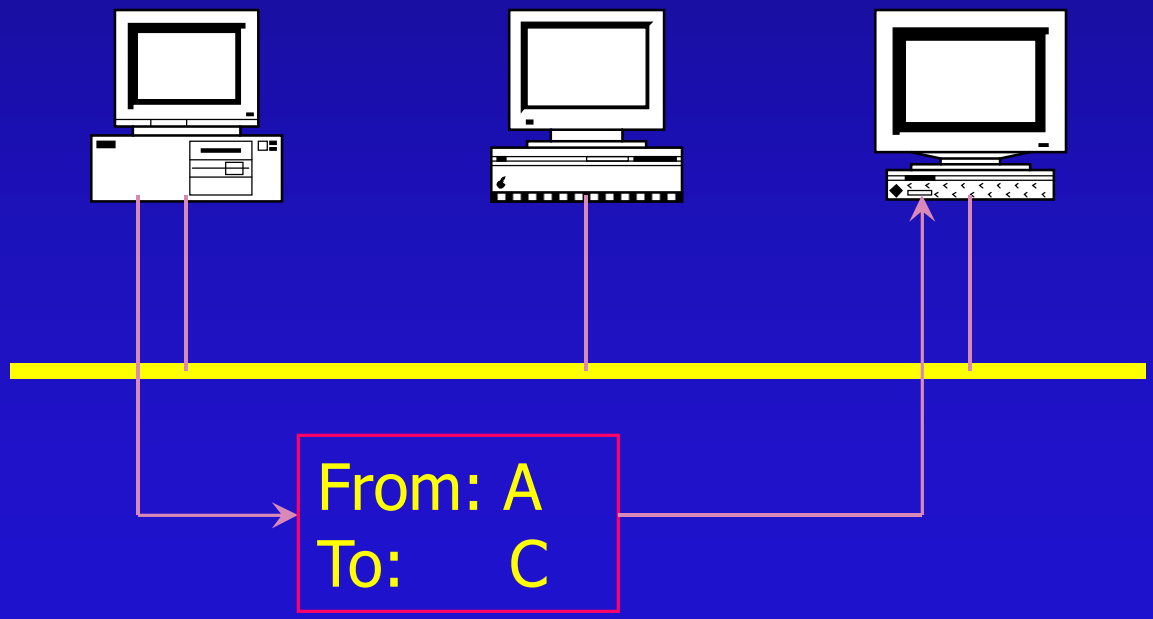
Internetwork Addressing

- Each device on a network or an internetwork is identified by a unique address, often called a **device** or **node address**.
- These addresses are frequently hard-coded into the network hardware.
- Each Ethernet and Token-Ring interface possesses a 48-bit address guaranteed to be unique throughout the world.

- A local delivery mechanism enables devices to place messages on the medium and retrieve messages that are addressed to them.
- This local delivery is performed by using the device address.
- The local delivery is handled by the physical and data link layers.

Simple Addressing

- On simple networks, delivery of messages between devices is simple.



- A mechanism is also needed to deliver messages that must cross network boundaries and travel through the internetwork.
- Internetworks can be very complex, so there must be a way to find out the best possible path from one node to another across the internetwork.
- This process of finding the best possible paths is referred to as **routing**.

TCP/IP-based Internetworks

- TCP/IP provides an excellent and simple approach with the widest acceptance.
- TCP/IP consists of the layers above and including the network layer.
- The lower layers (physical and data link) can be of many types, such as Ethernet, Token-Ring, X.25, Frame Relay, ATM, Serial Line, etc.

- TCP/IP was designed explicitly without data link and physical layer specifications because the goal was to make it adapt to most types of physical media.
- TCP/IP relies on the physical layer to deliver messages on the local network.
- For delivering messages across network boundaries, TCP/IP has its own addressing mechanism.

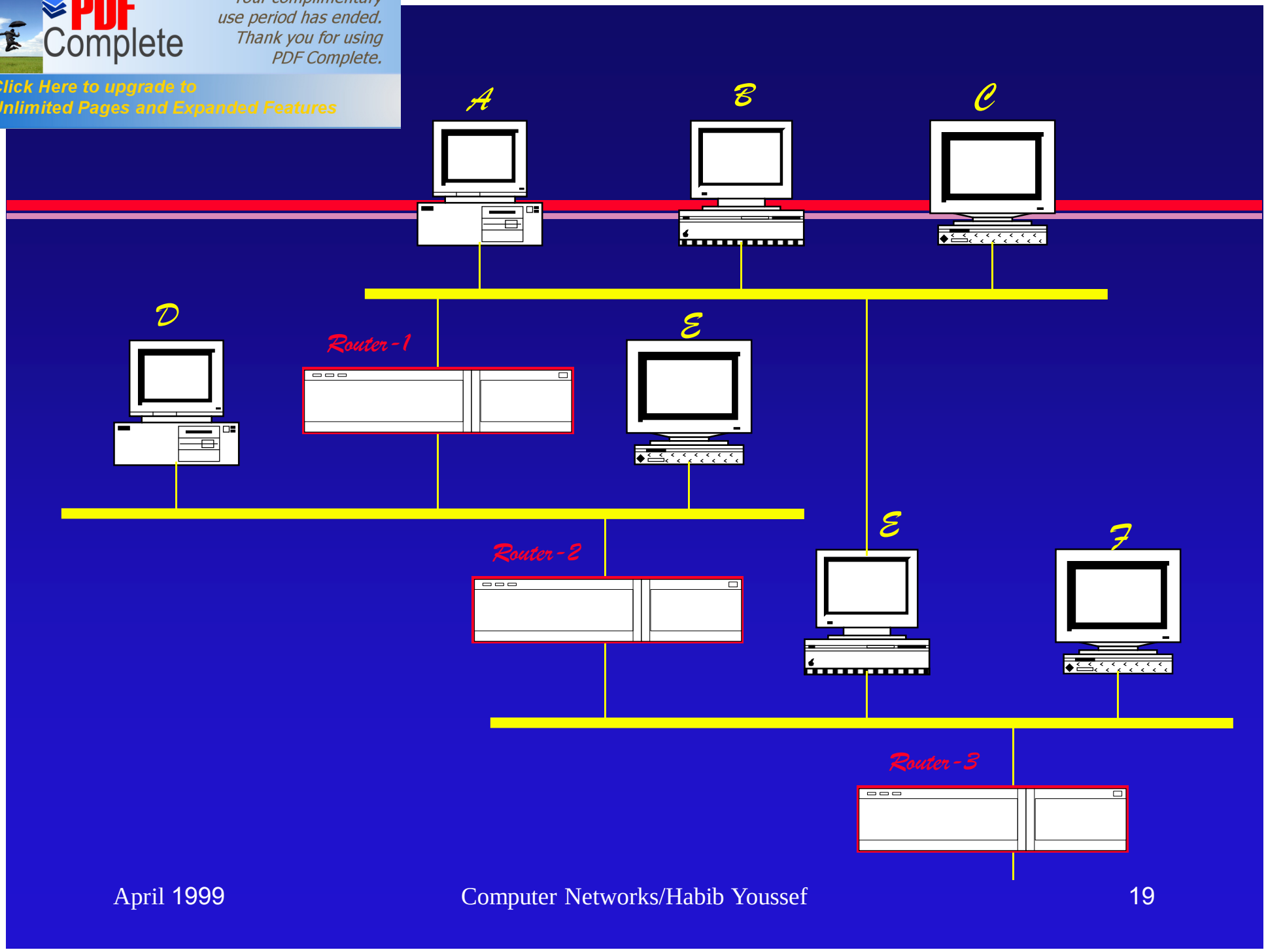
- This mechanism works at the network layer, and is handled by the IP (Internet Protocol) software.
- In TCP/IP terminology, any device that is connected to the network is referred to as a **host**.
- A host may be a computer, router, network printer, etc.

Local Message Delivery

- When IP sends a message that is directed to a device on the local network, it hands the message over to the physical layer software which tags the message with the physical address of the recipient, and sends it.
- The device that matches the physical address retrieves the message.

Message Routing

- When a message is not destined for a device on the local network, it must be routed.
- TCP/IP assigns an address to each host and to each network.
- Each host is configured with a default router to which it sends messages that must be sent to a remote network.



- The responsibility of determining how messages should be addressed is one of the tasks of the IP layer.
- IP identifies whether a message is destined for a host on the local network or it should be sent to the default router.
- It makes use of addresses called **IP addresses** to logically identify networks and hosts.

- The physical address of either a local host or the default router is added by the physical layer software to each message that is sent.
- IP receives data from the higher level protocols, and attaches to each data segment a header containing addressing information.

- The combination of data from higher layers with the IP header is referred to as a **packet**.
- Determining routing paths between routers is usually the responsibility of one of the following two protocols.
  - Routing Information Protocol (RIP)
  - Open Shortest Path First (OSPF)

Important questions

- How are the machines addressed?
- How do internet (IP) addresses relate to physical addresses?
- How do internet gateways learn about routes?

Internet addresses

- Internet is a universal communication system that uses a globally accepted addressing scheme to identify hosts connected to it.
- IP addresses uniquely identify each host
- Internet addressing helps TCP/IP software hide physical network details

Internet addresses (cont.)

- Names, addresses, and routes refer to successively lower level representations of host identifiers
  - A name identifies what an object is,
  - its address identifies where it is, and
  - a route indicates how to get to it
- TCP/IP addressing scheme analogous to physical network addressing

Internet addresses (cont.)

- Each Internet host is assigned a 32-bit integer address called its Internet address or IP address
- The integers are carefully structured for efficient routing
- IP address = {Net-ID, Host-ID}
- Gateways base routing on Net-ID

Internet addresses (cont.)

- 32-bit address number specified in each IP datagram
  - Written as 4 decimal numbers separated by dots (dotted quad notation)
  - Each number is from 0-255
  - Example: razi 196.15.69.230
- Number of bits used for Net-Id and for Host-Id depends on class of IP address

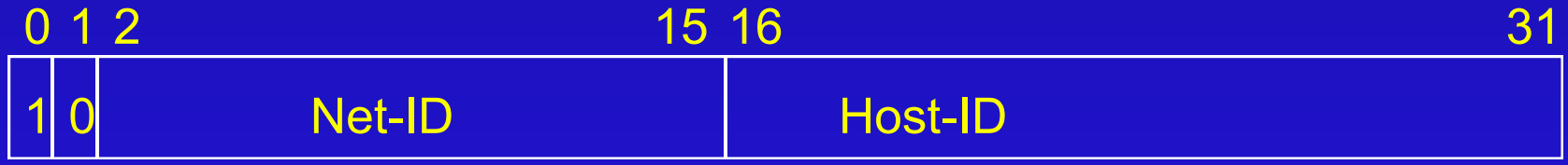
Classes of IP addresses

- Class A: Used for the very few large networks with more than 2^{16} hosts.
  First byte < 128

Classes of IP addresses (cont.)

- Class B: For medium size networks that have between 2^8 and 2^{16} hosts

  First byte is from 128 to 191



Classes of IP addresses (cont.)

- Class C: Small network < 2^8 hosts

First byte is from 192 to 223

Internet addresses (cont.)

- IP address
  - Not a host address
  - Each network interface has an IP address
  - Each IP address specifies a connection to a network not an individual machine
- A gateway connecting N networks has N distinct IP addresses, one for each physical network connection

Special Addresses

- Net-Id = 0, Host-Id = 0
  - Designates this host
  - Allowed only at startup
- Net-Id = 0
  - Host on this net
  - Allowed only at startup

Special Addresses (cont.)

- IP address all 1's
  - Limited broadcast
  - Never a valid source address
- Host-Id = all 1's
  - Broadcast address
  - Never a valid source address

Special Addresses (cont.)

- Net-Id = 127
  - Loopback address (Class A address)
  - Used for testing
  - Interprocess communication on local host
  - Allows local host to be addressed in the same manner as a remote host
  - Should never appear on a network

Weaknesses of IP addressing

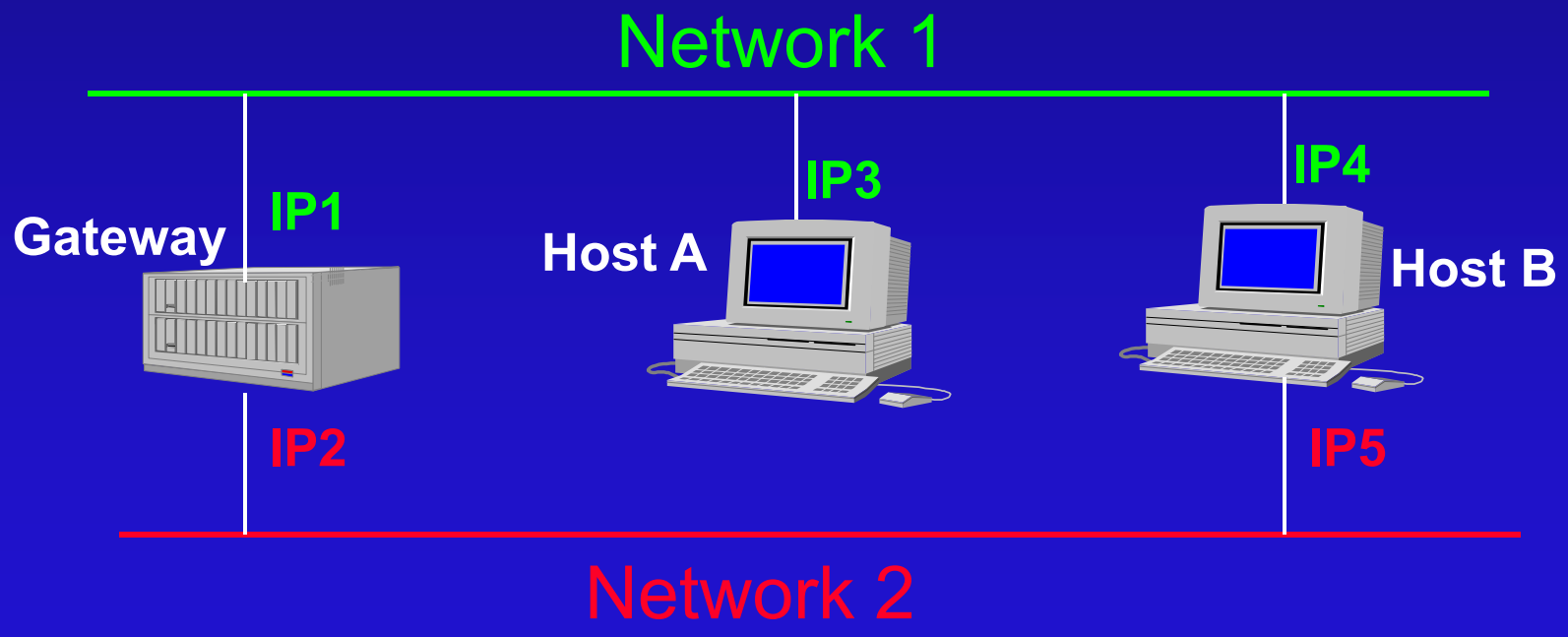
- Addresses refer to physical connections not to hosts
  - This disallows computer mobility because the IP address assigned to that computer also identifies the network it is attached to
  - If a host moves from one network to another, its IP address must be changed

Weaknesses of IP addressing (cont.)

- When any Class C network grows to more than 255 hosts, it must have its address changed to a Class B address
- Routing decisions are made on the basis of the Net-Id part of IP address
  - The path taken by packets traveling to a host with multiple IP addresses depends on the IP address used

Weaknesses of IP addressing (cont.)

If connection of Host B to Network 1 fails, users on Host A who specify IP4 can no longer reach B, where those that specify IP1 can still reach Host B

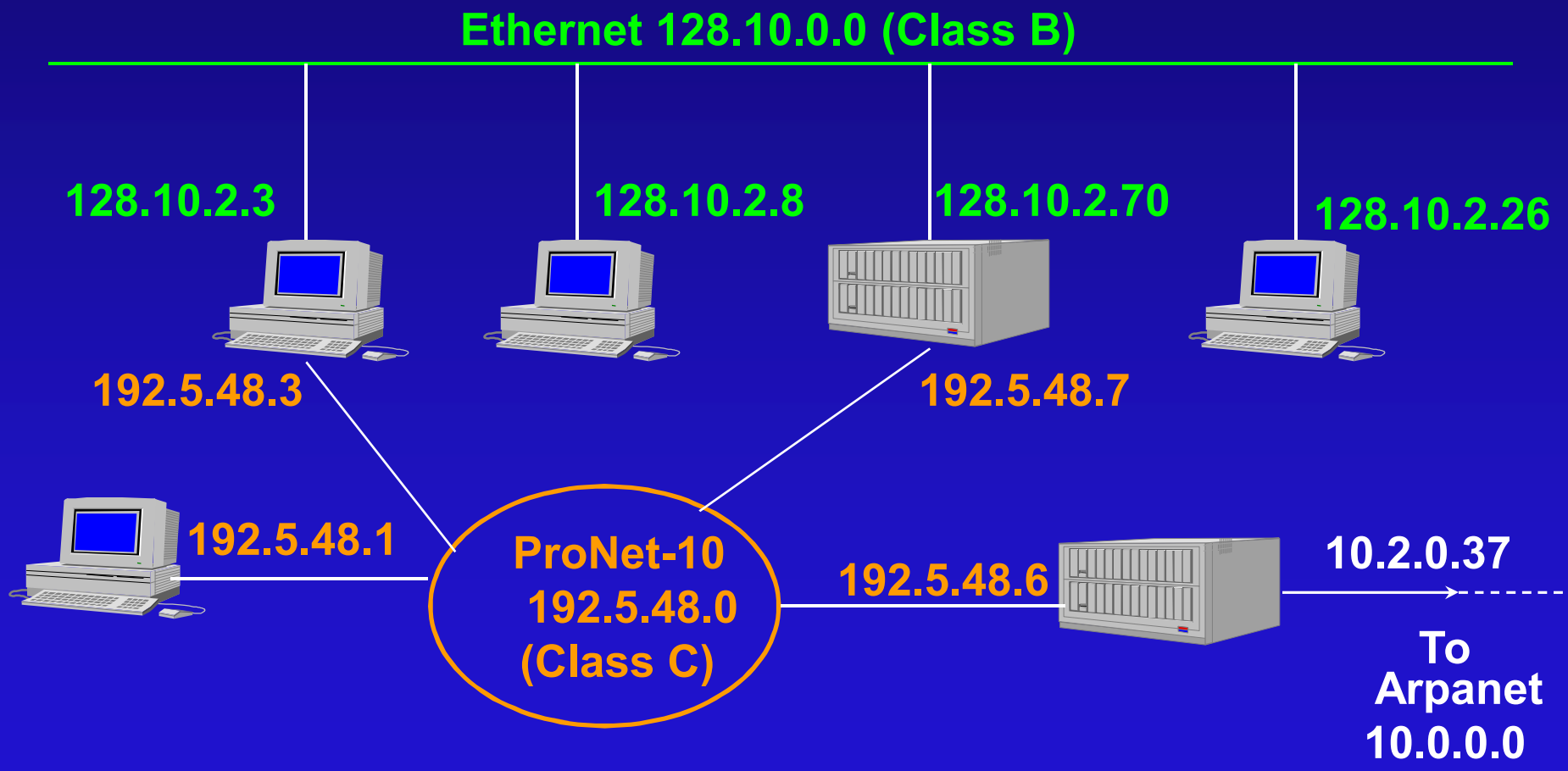


Internet Addressing Authority

- All internet addresses are assigned by a central authority:

  The network Information Center (NIC)
- The NIC assigns the Net-Id portion
  - Small networks (< 255 hosts) are assigned Class C addresses, since many LANs are expected
  - Large networks are assigned Class A addresses since only few such networks are expected

Example



IP Addresses to Physical Addresses

- How does a machine map its IP address to its physical network address?
  - Example:
    - Machines A and B connected to the same network, with IP addresses IA and IB and physical addresses PA and PB.
    - Suppose A has has only B's IP address, then how does A map IB to PB?

Address Resolution

- Some protocol suites adopt one of the following:
  - Keep mapping tables in each machine
  - Hardware (physical) addresses are encoded in the high level addresses
- Both are ad-hoc, awkward solutions

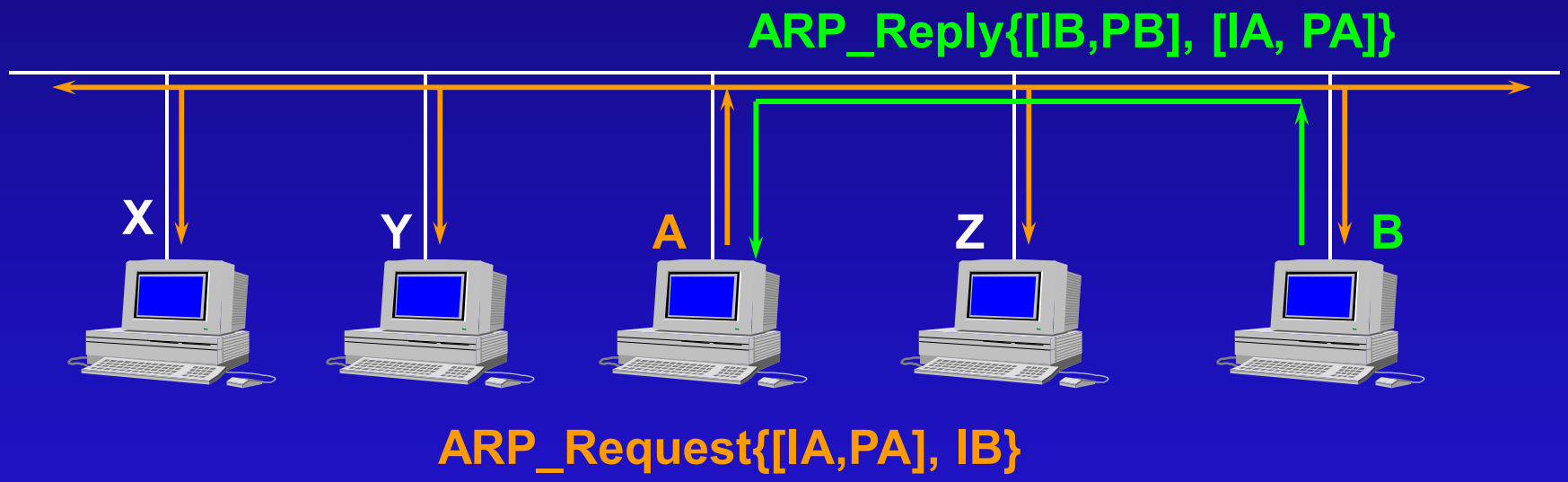
tion Through Dynamic Binding

- Ethernet uses 48-bit physical addresses
  - » Addresses assigned by manufacturers
  - » Replacing a faulty interface card meant a change to the machine physical address
- Can't encode 48-bit long address into a 32-bit long IP address
- TCP/IP solution: Address Resolution Protocol (ARP)

ARP

- Exploits broadcast capability of Ethernet
- Allows a host to find the Ethernet address of a target host on the same network, given the target's IP address
- Builds and maintains dynamically a table to translate IP addresses into Ethernet physical addresses

ARP (cont.)



ARP (cont.)

- Hosts that use ARP maintain a small cache of recently acquired (IP,P) address bindings
- Cache is updated dynamically
  - Timer for each entry
  - Whenever a new binding is received, update the corresponding table entry and reset the associated timer

ining an IP Address at Startup

- Diskless machines use IP addresses to communicate with the file server
- Also, many diskless machines use TCP/IP FTP protocols to obtain their initial boot image, thus requiring that they obtain and use IP addresses
- Designers keep both the bootstrap code and initial OS images free from specific IP addresses for portability

ng an IP Address at Startup (cont.)

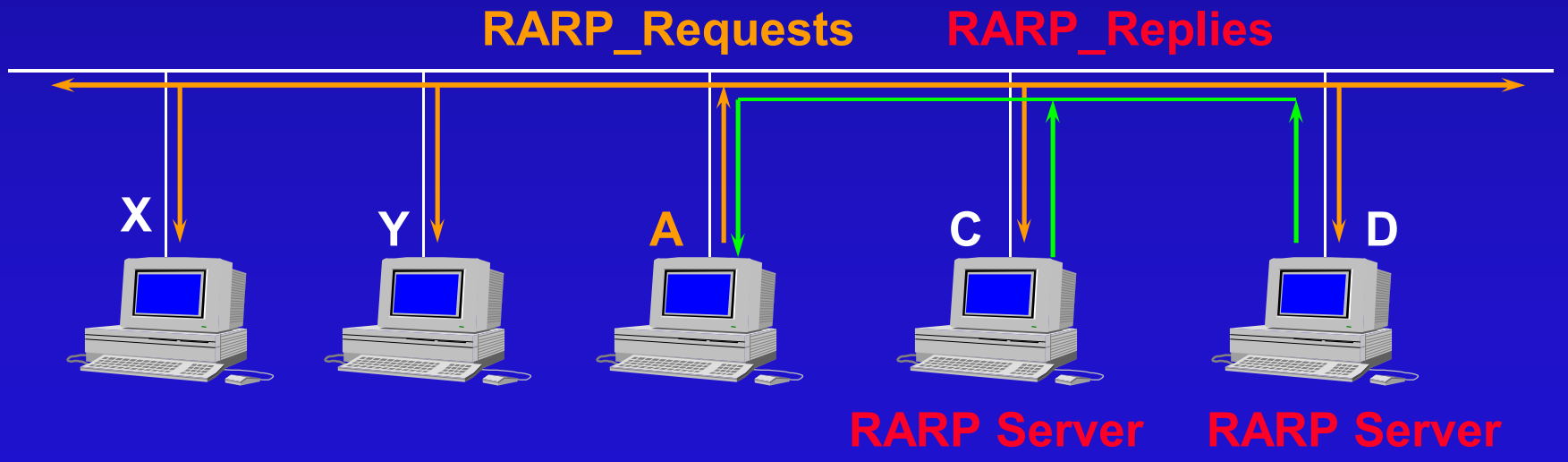
- How does a diskless machine determine its IP address?
- When bootstrap code starts execution on a diskless machine, it must use the network to contact a server to obtain the machine's IP address
- Usually, a machine's IP address is kept on disk where OS finds it at startup

e Address Resolution Protocol (RARP)

- RARP is the protocol used to solve the reverse problem solved by ARP
  - Given a physical address, get the corresponding IP address
- RARP uses the same message format as ARP
- RARP messages are sent encapsulated in Ethernet frames

RARP (cont.)

- RARP allows a host to ask about an arbitrary target
  - The sender supplies its HA separate from the target HA, and the server is careful to reply to the sender's HA



TCP/IP-Based Applications

Remote Login (TELNET)

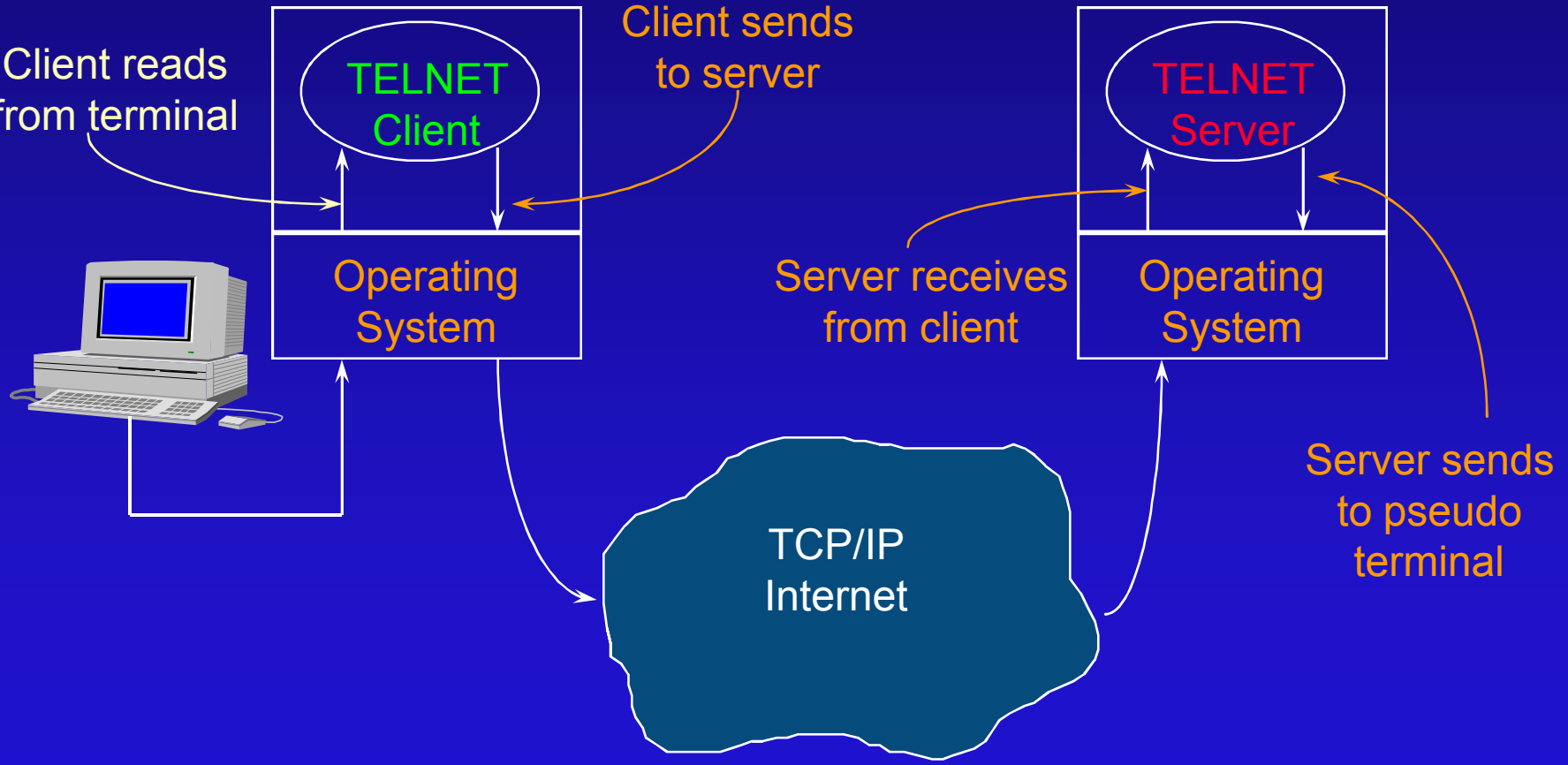
TELNET (cont.)

- Internet services are provided through application level programs
- **Telnet is a** Terminal emulation application program.
- Allows a user to remote-login on to another computer.

TELNET (cont.)

- **TELNET**
  - Allows a user at one site to establish a TCP connection to a login server at another
    - TELNET client software allows the user to specify a remote machine by giving its domain name or IP address
  - Passes keystrokes from the user terminal (client site) to the remote machine (server)
  - Carries output from the remote machine back to the user's terminal

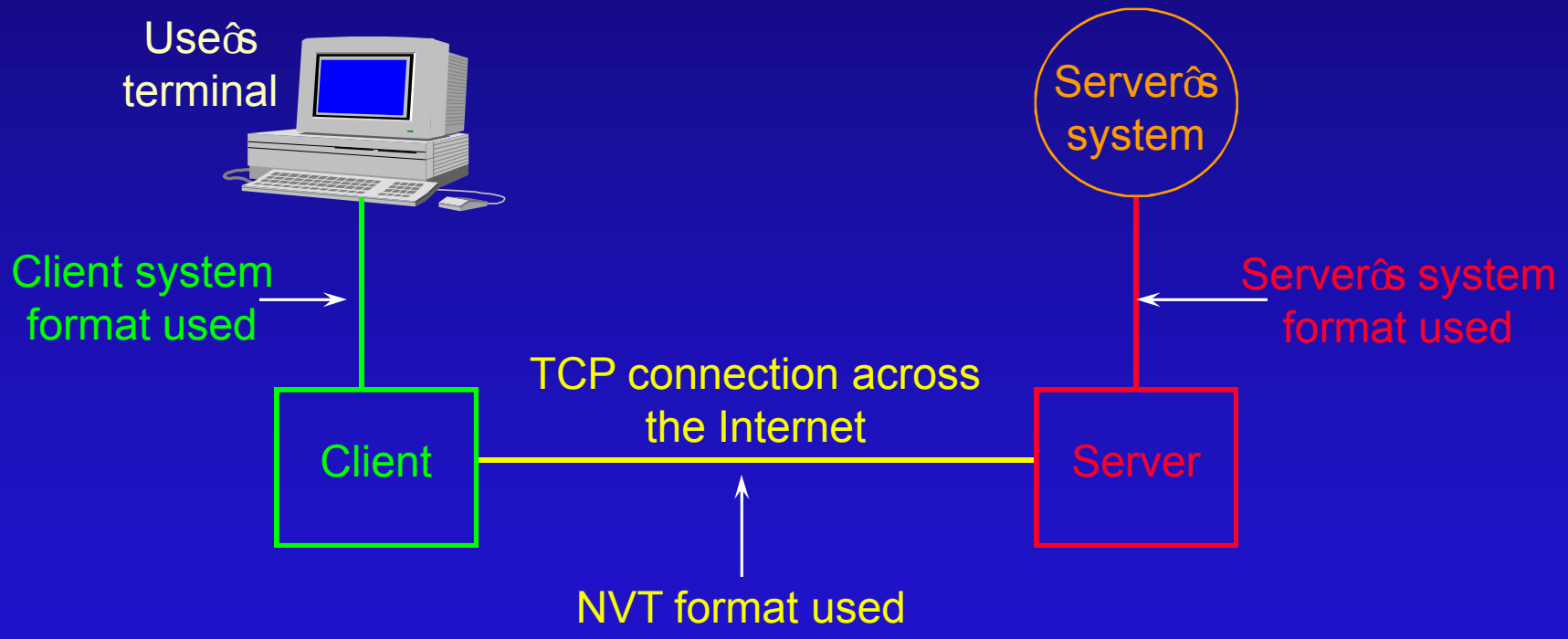
TELNET (cont.)



TELNET (cont.)

- To accommodate heterogeneity, Telnet defines how data and commands are sent across the Internet. The definition is known as the Network Virtual Terminal (NVT)

TELNET (cont.)



TCP/IP-Based Applications

File Transfer & Access

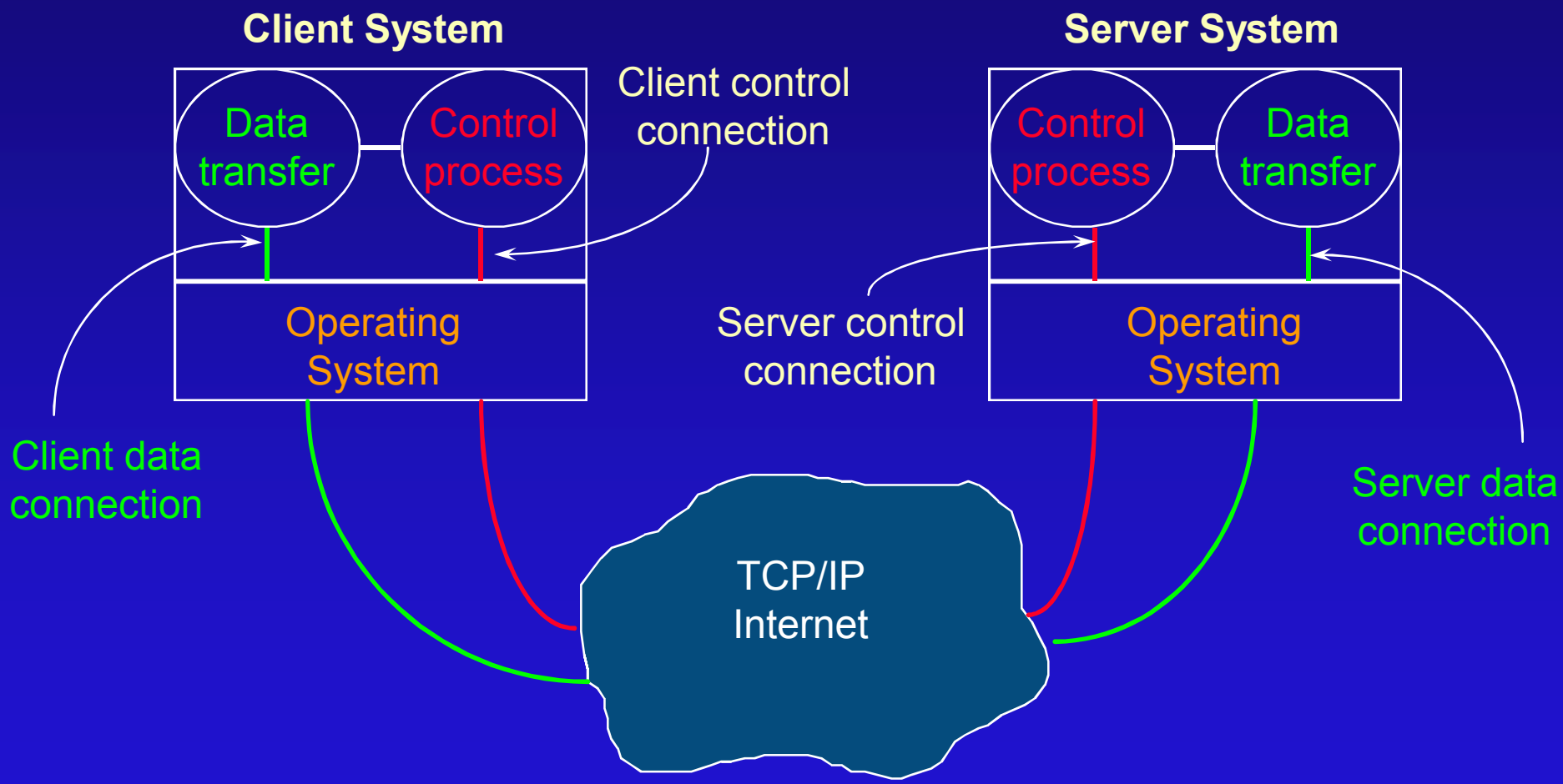
File Transfer

- A facility to access files on remote machines
- FTP is the major TCP/IP file transfer protocol
- File transfer is among the most frequently used TCP/IP applications
- Anonymous downloading of files.

File Access Model

- Like most other servers, most FTP implementations allow concurrent access to multiple clients
  - Clients use TCP to connect to the server
    - Control connection carries commands telling the server which file to transfer
    - Data transfer connection carries data transfers
  - A single master server process awaits connections and creates a slave process to handle each connection

File Access Model (cont.)



TCP Port Number Assignment

- When a client forms a connection to a server
  - The client uses a random, locally assigned, protocol port number
  - But, the client contacts the server at a well known port number (Port 21)
- Once the control connection is established, future TCP connections established for data transfers use other port numbers on the client machine, and Port 20 on the server machine

User's View of FTP

- FTP viewed as an interactive system
- Once invoked, a client performs the following operations repeatedly
  - Read a line of input
  - Parse the line and extract command and its arguments
  - Execute the command

Example of FTP Session

```
% ftp spice.ccse.kfupm.edu.sa  -- Invokes ftp
    ……..
    ……..
Name (spice:youssef) CR
Password: ******CR
ftp> help CR   -- lists various ftp commands
ftp> help bell
bell      beep when command completed
ftp> bell
Bell mode on
ftp> ls              -- lists remote directory
    ……..
```

Example of FTP Session (cont.)

```
ftp> cd shortcourse/tcpip  -- move to indicated directory
    ……..
ftp> get  RemoteFile  LocalFile
    ……..
ftp> put Localfile RemoteFile
    ……..
ftp> close
    ……..
ftp> quit
%
```

TCP/IP-Based Applications

Electronic Mail

Introduction

- Email is the first encounter of users with computer networks
- Millions connected to the Internet use it.
- Low cost and fast communication.
- Encourages collaboration.
- "A person ... can say HELP to 10,000 people ... The next morning he may have 15 answers to his problem."

Introduction (cont.)

- E-mail is delivered in few minutes.
- E-mail costs half that of regular postal mail (SNAIL MAIL) and ONLY 15% that of Fax.

Email address

youssef@ccse.kfupm.edu.sa

youssef : User name

@ : Connects the who to where

ccse : subdomain name

kfupm : domain

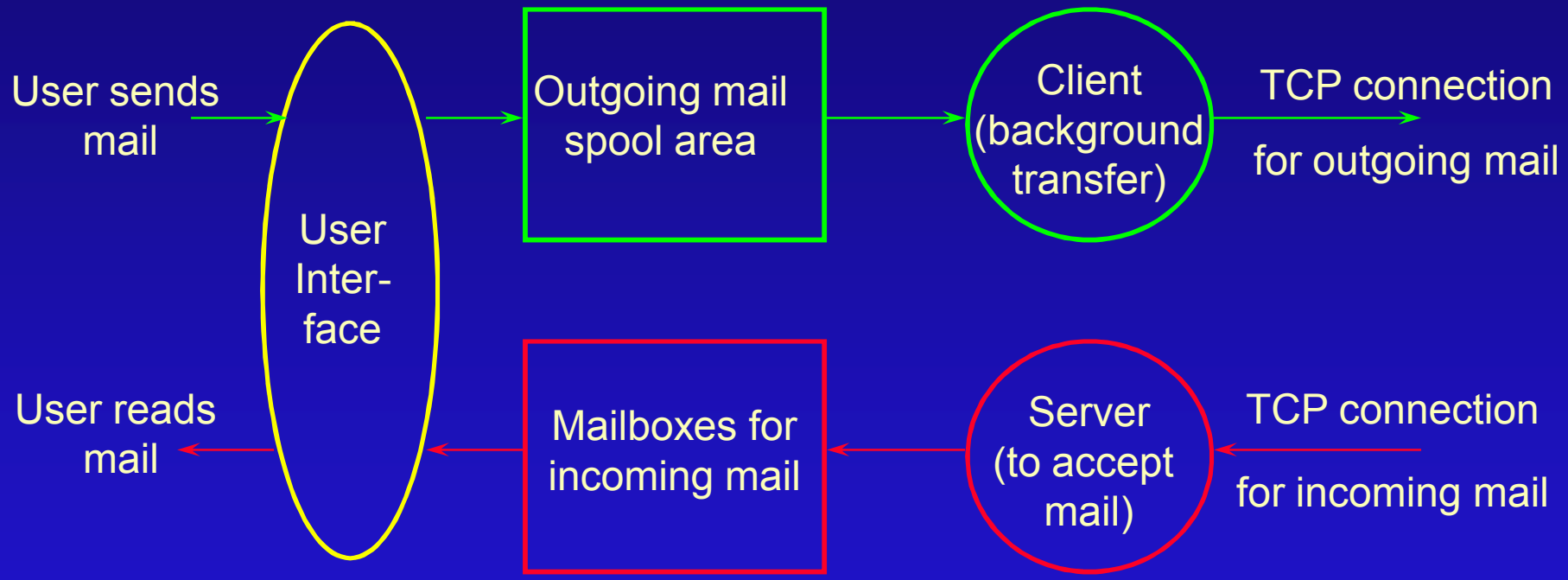
edu : segment type

sa : final where segment (sa= Saudi Arabia, tn= Tunisia, ca: Canada)

Spooling

- Mail systems use Spooling technique to handle delayed delivery
  - When a user sends a message, the system places a copy in its private storage (spool) area along with the identification of sender, recipient, dest machine, and time of deposit
  - The transfer is initiated in the background, allowing the sender to proceed with other activities

ceptual Components of an Email System



Email concepts (cont.)

- The background mail transfer process becomes a client
  - It maps the dest machine name to an IP address
  - It forms a TCP connection to the mail server on dest machine
  - It passes a copy of the message to the remote server, which stores a copy in the remote's system spool area

Email concepts (cont.)

- Once the client and server agree that the copy has been accepted and stored, the client removes the local copy
- If TCP connection fails, the transfer process records the time it tried delivery and terminates

Email concepts (cont.)

- The background transfer process sweeps through the spool area periodically

  For each undelivered or new outgoing mail
  - It attempts delivery again
  - If a mail message cannot be delivered after an extended time (3 days), it returns the mail message to the sender

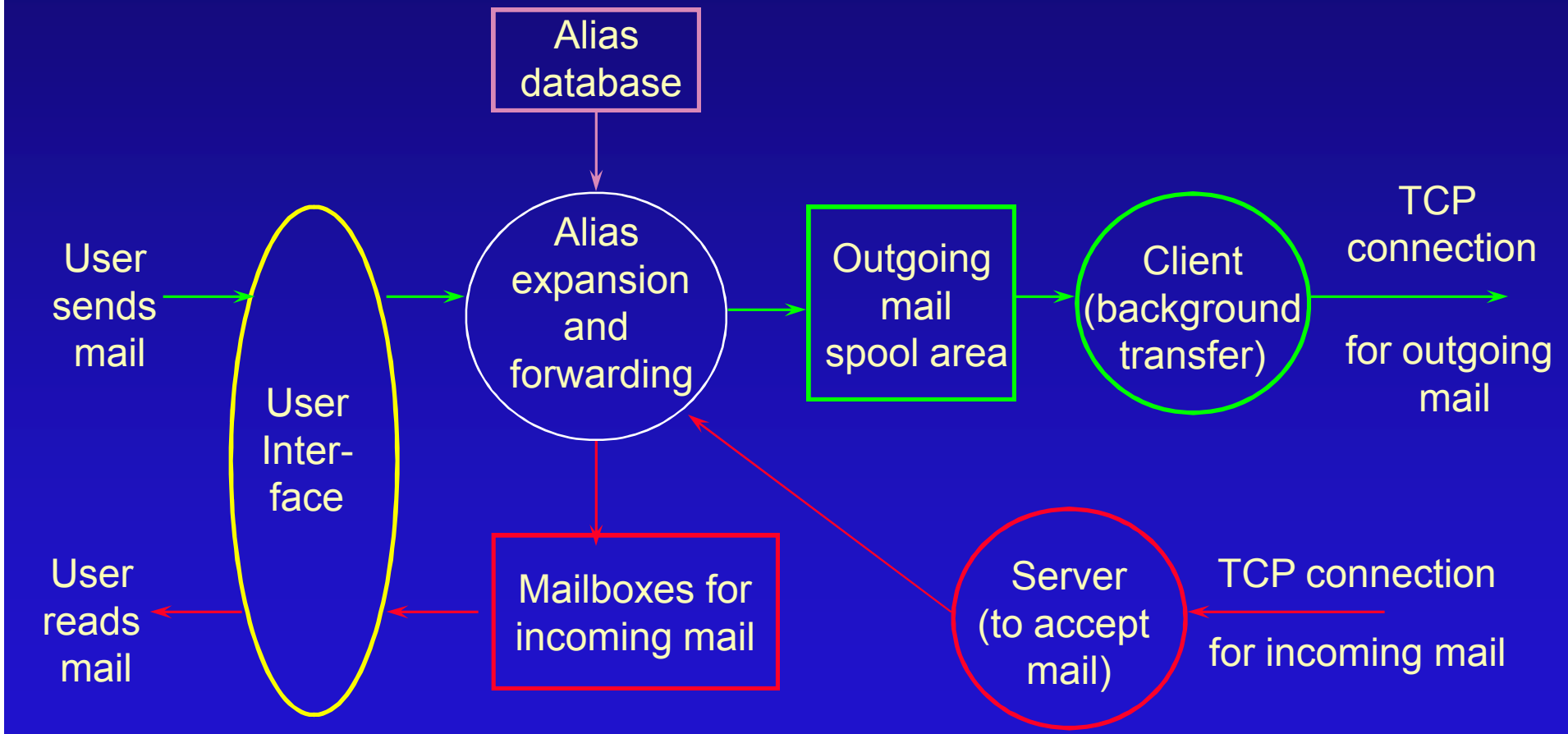
Mailbox names and Aliases

- Users specify
  - the mail destination machine (usually the machine's domain name)
  - a mailbox at that machine (usually the user's login Id)
- Most systems provide mail forwarding software that includes alias expansion mechanism

s Expansion and Mail Forwarding

- A mail forwarder allows the local site to map Ids used in mail addresses to a set of one or more new mail addresses
- After a user composes a message and names a recipient
  - » the mail interface consults the local aliases to perform necessary mappings before passing the message to the delivery system

Conceptual Model of a Mail System



TCP/IP Standard for Email Service

- TCP/IP divides its mail standard into two sets
  - One standard specifies the format for mail messages (RFC 822)
  - The other specifies the details of electronic mail exchange between two computers
- This division makes it possible to build mail gateways to non TCP/IP networks while still using the same format

Standard Format

- Headers contain readable text, divided into lines that consist of
  - a keyword
  - a colon ":"
  - a value
- Some keywords are required, others are optional, and the rest are uninterpreted

Standard Format (contd.)

- Examples

TO: ics.faculty@ccse.kfupm.edu.sa
from: youssef
Reply to: elleithy@ccse.kfupm.edu.sa
cc: coe.faculty, se.faculty
subject: Farewell party for Dr. Osman

Electronic Mail Addresses

- Email addresses have a simple, easy to remember form

  local-part@domain-name

  domain-name: mail exchanger of the mail destination
  local-part: address of a mailbox on that machine

  youssef@ccse.kfupm.edu.sa

Mail Transfer Protocol (SMTP)

- SMTP is the standard mail transfer protocol of TCP/IP
- SMTP focuses on how the underlying mail delivery system passes messages across a link from one machine to another
- SMTP is simple.

ng Resources on the Internet

- **Archie.**
  - Used to search for files available via anonymous ftp.
- **Gopher.**
  - Friendly menu-driven search tool for browsing resources and displaying the requested information.

WAIS

- **WAIS : Wide Area Information Server**
- Software used to index large text files in servers.
- On the client side, it finds and retrieves documents in databases, based on user-defined keywords.
- Works on an index. The index is searched and the data tied to the index is retrieved.

WWW

- **WWW: World-Wide Web**
- Hypermedia-based system for storing and accessing hypermedia documents anywhere on the Internet.
- Each Web site has a Web server.
- Users (clients) access information in a Web site using a Web browser such as Netscape or Mosaic.

WWW (Cont.)

- WWW is the most popular tool to publish on the Internet.
- Already all major computer manufacturers, businesses, airlines, embassies, retail stores, etc., have Web pages.
- Ex:http://www.kfupm.edu.sa/~youssef
  http://www.cnn.com

TCP/IP and the Internet

Connecting to the Internet

Requirements

- Connecting to the Internet requires the following.
  - Establishing physical connections to the Internet
  - Registering the Internet addressing scheme
  - Registering a domain name
  - Optional types of registration which might be needed

Getting Connected

- The first thing any organization must do to get connected to the Internet is pick an approved **Internet Service Provider** .
- The InterNIC strongly encourages all interested parties to select an ISP rather than trying to establish a direct link into the Internet.

ments for Full-Service Links

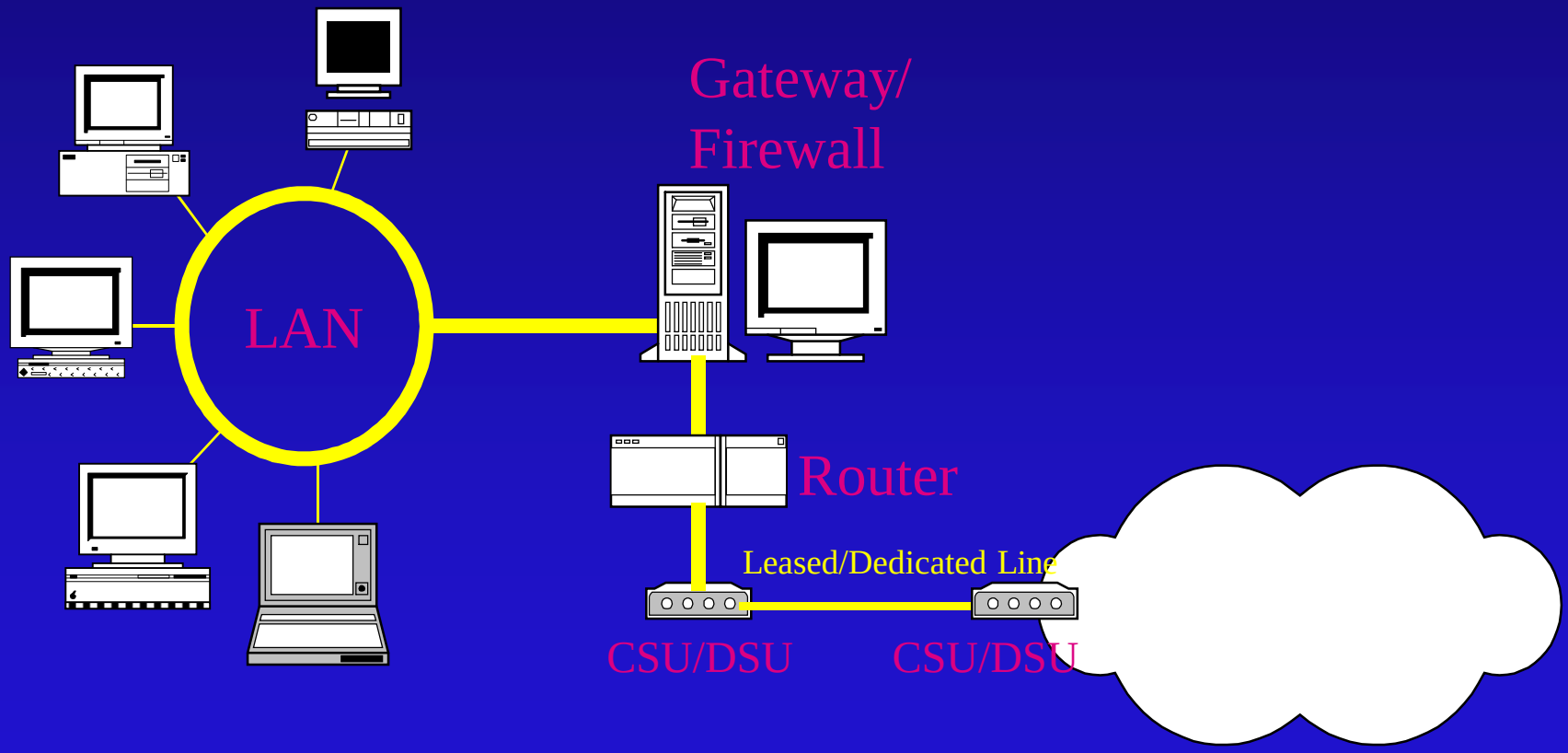
- Full-service connections use full-time, dedicated telecommunications circuits between a subscribing organization and an ISP.
- At least one **Internet Server** must be installed at the site to support the primary Internet services such as
  - » electronic mail, file transfer, and information retrieval using tools like Gopher, WWW, and WAIS.

- For small scale connections this server can provide IP routing as well, acting as a gateway between the organization's local area network and the Internet.
- Larger networks will probably need to install a dedicated router instead.
- In addition, security concerns might require the installation of a "firewall".

Ty

- At a minimum most organizations will require a dedicated analog dialup connection using either the "SLIP" or "PPP" protocols from an Internet Access Provider.
- SLIP (the Serial Line Internet Protocol) and PPP (the Point-to-Point Protocol) are two methods to provide an Internet connection over dialup telephone lines.
- Higher-speed (i.e., greater bandwidth) connections are available for organizations expecting heavier Internet usage.

Dedicated Internet Access



Personal Internet Access