Introduction to Networking

Introduction

- A distributed systems consists of a collection of computers linked together by a network
- Distributed system software enables computers to coordinate activities and share resources (software, hardware, data)

Internetworking

- Network: a set of computers and peripherals (printers, modems, etc.) connected together by a medium (direct or indirect connection).
- Network architecture:
  - Devices in same room or building (LAN) vs. WAN

Communication Protocols

To communicate any two entities requires establishing certain rules that can be expressed by a **communication protocol**.

- syntax
- semantics
- timing and synchronization

Protocol Stacks

The protocol stack separates the communicating application from aspects of communication that are irrelevant to that application.

- The International Standardization Organization (ISO) proposed the **Open System Interconnection** (OSI) communication model.
- The Internet does NOT strictly follow the OSI model!

Protocol Stacks (2)

- Each layer presents a set of services for the layer immediately above it and can use the services provided by the layer immediately below it.
- The services are accessed through **service access points** (SAP).
- At each layer, certain control information is attached to the message. This is called a **header**.
- data + header = protocol data unit (PDU)
  - the size of the message passing through the protocol stack increases at each layer!

Protocol Stacks (3)

- Message headers may contain info needed for:
  - Connection control
  - Addressing
  - Segmentation and re-assembly
  - Sequencing
  - Error recovery
  - Flow control
  - Multiplexing

Connection Control

- Two fundamental types of protocols:
  - Connection-less
    - Each PDU is handled and routed independently
    - The message is called a **datagram**
    - Many different routing algorithms (static vs adaptive)
  - Connection oriented
    - a **virtual circuit** (VC) is maintained for the duration of the communication

Addressing

- Need to indicate the destination of the message
- Other info:
  - Addressing level (e.g. IP address)
  - Mode
    - Unicast
    - Multicast
    - Broadcast

Segmentation and Re-assembly

- Long messages are cut into smaller chunks
- Header indicates chunk number

Sequencing

- Must maintain chronological sequence of messages to conform to protocol!
- Header contains sequence number

Error Recovery

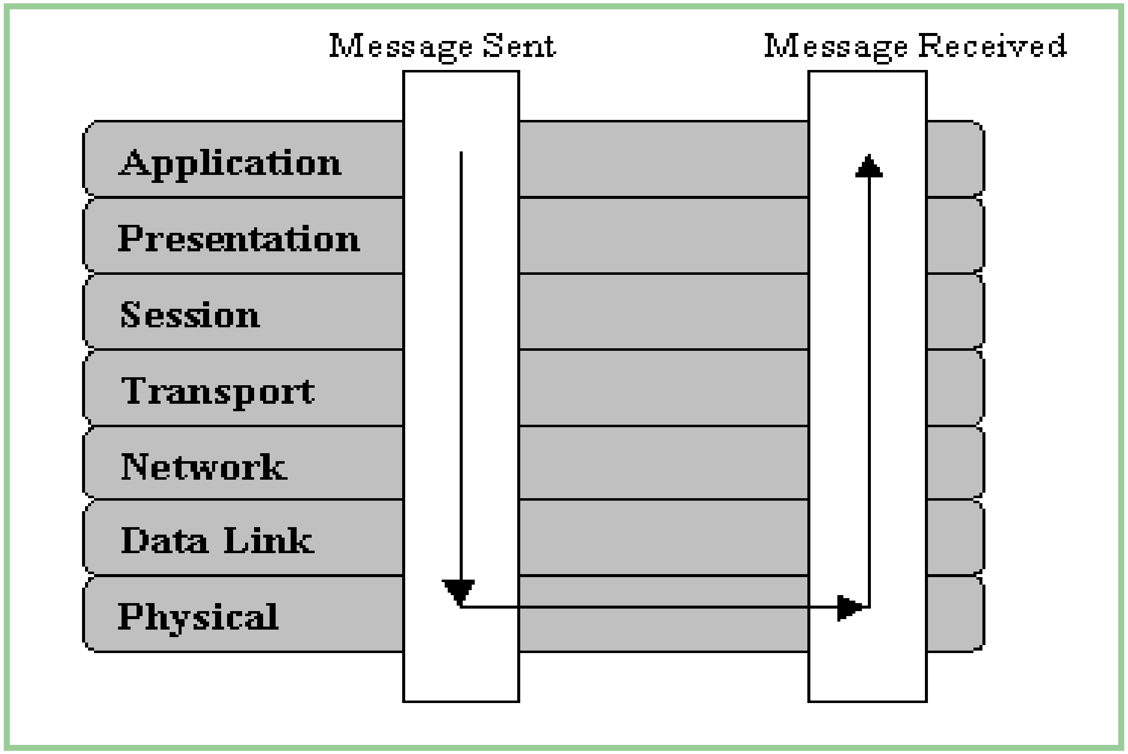
- Can use sequence number to detect loss of message
- Request for re-send
- Buffering within “window” of opportunity to limit re-send requests

Flow Control

- Producer might be too fast for consumer!
- Need to tell producer to stop or slow down
- More generally: congestion control

The OSI Reference Model

- A protocol stack with seven layer



Internet vs OSI

- Driven by Internet Engineering Task Force (IETF)
  - IETF issues Requests for Comments (RFCs)
- The Internet addresses layers 3 and up
- Layers 5, 6 and 7 are collapsed into one Application Layer
- Network Layer: the Internet Protocol (IP)
  - Routing: RIP, BGP, OSPF…
- Transport Layer: TCP/IP and UDP/IP
- Application layer: ftp, smtp, http, snmp…

- Does the number of layers matter?

Finding Things: Names/Addresses

- To make datagram delivery possible, each computer on the net is assigned a unique identifier (well, at least one…)

- How many “names” exist for a “PC”
- How to address the webserver that manages the course webpage?

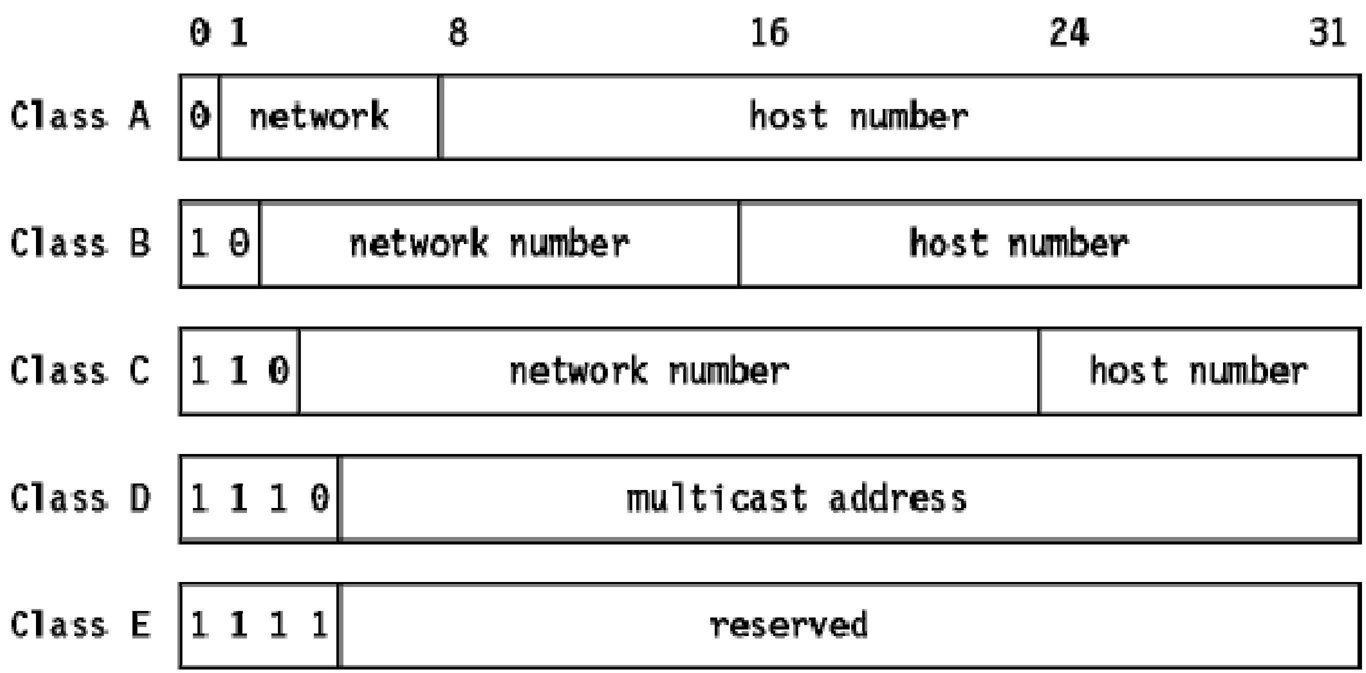
Addresses/Names

- Network Interface
  - Hardware/MAC address, often 48-bit IEEE address (but not always: Bluetooth has 3 bit address dynamically assigned, dedicated links do not need hardware addresses) Why not use IP address?
- Hosts (well, technically interface as well)
  - IP address (i.e. 134.117.63.134)
  - DNS name/symbolic name (i.e., kunz-pc.sce.carleton.ca)
- Applications
  - Port Numbers

More on Naming

- Names/Addresses better be unique (sort of)
  - ICANN/IANA, IEEE
- Names/Addresses have structures, are NOT opaque identifiers, structure is exploited in locating entities
  - IP addresses are 32-bit numeric identifiers containing network and host identifiers (see later)
  - DNS/symbolic names have a hierarchical structure, based on well-known "top-level" domains
  - IEEE MAC addresses: 24 bits identify manufacturer, 24 bits supposedly running number uniquely identifying each NIC built by manufacturer
  - Port numbers: ?

Internet Addresses



The IP addresses uniquely identify the network and a host on that network

Internet Addresses

- Different classes meet the requirements of different organizations
  - Class A for large networks (>65,536 hosts)
  - Class B for med-size networks (256<size<65,536)
  - Class C for networks with up to 256 hosts
  - Class D is used for multicasts
  - Class E for future use
- Network part of Internet address is assigned by ICANN
- Host part is assigned by owner of the network

Subnet mask

- Allows further subdivision of network
- Is achieved by making special use of host number:
- <network number><subnet number><host number>

- The Subnet Mask tells you how many bits are assigned to the subnet number

Which Addresses/Names do End Users need to know?

- IP address: numerical, hard to remember
- Symbolic name often easier to remember
- Translation from Symbolic name to IP address: Domain Name Service (DNS)
- MAC Address: only really used/needed for MAC protocol – automatic translation from/to IP address using ARP/RARP
- Port Numbers: ?

/etc/host and /etc/services

- Type: cat /etc/host (or /etc/services) in a Unix system
- What do you see?