Connecting computers

- Point-to-point links
  - Connect one sender with one receiver
  - No conflict for access to link
  - Not practical: limits any-any communication

Multiple Access Protocols

- Ways of enabling devices to share a network
  - Three approaches
    1. **Channel partitioning**
       - Time Division Multiplexing (TDM)
         - Each node gets a time slot
       - Frequency Division Multiplexing (FDM)
         - Each channel gets a frequency band
    2. **Taking turns**
      - Polling protocol – master polls nodes in sequence
      - Token passing protocol – node needs a token to transmit
    3. **Random access**
      - No scheduled time slots
      - Statistical multiplexing
      - Retransmit if there's a collision

Modes of connection

**Circuit-switching (virtual circuit)**
- Dedicated path (route) – established at setup
- Guaranteed (fixed) bandwidth – routers commit to resources
- Typically fixed-length packets (cells) – each cell only needs a virtual circuit ID
- Constant latency

**Packet-switching (datagram)**
- Shared connection; competition for use with others
- Data is broken into chunks called packets
- Each packet contains a destination address
- Available bandwidth ≤ channel capacity
- Variable latency

Ethernet

- Packet-based protocol
- Originally designed for shared (bus-based) links
- Each endpoint has a unique ethernet address
  - MAC address: 48-bit value
- Evolution
  - **Ethernet hub**
    - Simulates a bus-based LAN
  - Every bit received on an interface is transmitted onto every other interface
  - **Ethernet switch**
    - Forwards frames to the correct host
    - Self-learning (learns from source address)
    - Switch is transparent to hosts
    - No collisions!
Ethernet service guarantees

- Each packet (frame) contains a CRC checksum
  - Recipient will drop the frame if it is bad
- No acknowledgement of packet delivery
- Unreliable, in-order delivery

Going beyond the LAN

- LAN = Local Area Network
  - A set of devices connected to the same ethernet network is a LAN
  - Wi-Fi (802.11) is compatible with ethernet and part of the LAN
- We want to communicate beyond the LAN

The Internet

- Evolved from ARPANET (1969)
- Internet = global network of networks based on the Internet Protocol (IP) family of protocols

The Internet: Key Design Principles

1. Support interconnection of networks
   - No changes needed to the underlying physical network
   - IP is a logical network
2. Assume unreliable communication; design for best effort
   - If a packet does not get to the destination, software on the receiver will have to detect it and the sender will have to retransmit it
3. Routers connect networks
   - Store & forward delivery
   - They need not store information about the flow of packets
4. No global (centralized) control of the network

Routers tie LANs together into one Internet

A packet may pass through many networks – within and between ISPs

Protocols

- For effective communication
  - same language, same conventions
- For computers:
  - electrical encoding of data
  - where is the start of the packet?
  - which bits contain the length?
  - is there a checksum? where is it? how is it computed?
  - what is the format of an address?
  - byte ordering
Protocols

These instructions & conventions are known as protocols.

Protocols encompass data formats, order of messages, responses.

Layering

To ease software development and maximize flexibility:

- Network protocols are generally organized in layers.
- Replace one layer without replacing surrounding layers.
- Higher-level software does not have to know how to format an Ethernet packet.
  … or even know that Ethernet is being used.

Protocols

Exist at different levels

- Understand format of address and how to compute a checksum.
  vs. request web page.
- Humans vs. whales, different wavelengths, French vs. Hungarian.

Layering

Most popular model of guiding (not specifying) protocol layers is

**OSI reference model**

Adopted and created by ISO.

7 layers of protocols.

OSI Reference Model: Layer 1

Transmits and receives raw data to communication medium.
Does not care about contents.
Media, voltage levels, speed, connectors.

Deals with representing bits.
Examples: USB, Bluetooth, 802.11.

OSI Reference Model: Layer 2

Detections and corrects errors.
Organizes data into frames before passing it down. Sequences packets (if necessary).
Accepts acknowledgements from immediate receiver.

Deals with frames.
Examples: Ethernet MAC, PPP.
An *ethernet switch* is an example of a device that works on layer 2. It forwards *ethernet frames* from one host to another as long as the hosts are connected to the switch (switches may be cascaded). This set of hosts and switches defines the *local area network (LAN)*.

**OSI Reference Model: Layer 2**

- **1 Physical**
- **2 Data Link**

**OSI Reference Model: Layer 3**

An *IP router* is an example of a device that works on layer 3. A router takes an incoming IP packet and determines which interface to send it out. It enables multiple networks to be connected together.

**OSI Reference Model: Layer 3**

- **1 Physical**
- **2 Data Link**
- **3 Network**

**OSI Reference Model: Layer 4**

Provides an interface for end-to-end (application-to-application) communication: sends & receives *segments* of data. Manages flow control. May include end-to-end reliability.

**OSI Reference Model: Layer 4**

- **1 Physical**
- **2 Data Link**
- **3 Network**
- **4 Transport**

**OSI Reference Model: Layer 5**

Services to coordinate dialogue and manage data exchange. Software implemented switch manages multiple logical connections. Keep track of who is talking; establish & end communications.

**OSI Reference Model: Layer 5**

- **1 Physical**
- **2 Data Link**
- **3 Network**
- **4 Transport**
- **5 Session**

**OSI Reference Model: Layer 6**

Data representation concerned with the meaning of data bits. Convert between machine representations.

**OSI Reference Model: Layer 6**

- **1 Physical**
- **2 Data Link**
- **3 Network**
- **4 Transport**
- **5 Session**
- **6 Presentation**

Examples:
- Ethernet switch
- IP router
- HTTP 1.1
- Session
- Presentation
- Data representation
- XDR, ASN.1, MIME, XML
- Physical
- Data Link
- Network
- Transport
- Session
- Presentation
- Data representation
OSI Reference Model: Layer 7

A layer communicates with its counterpart

But really traverses the stack

IP vs. OSI stack

Collection of application-specific protocols

Deals with app-specific protocols

Examples:
- web (HTTP)
- email (SMTP, POP, IMAP)
- file transfer (FTP)
- directory services (LDAP)

A layer communicates with its counterpart

A layer communicates with its counterpart

What's really happening

Internet protocol stack
OSI protocol stack

Middleware

Physical
Data Link
Network
Transport
Session
Presentation
Application

Logical View

Logical View

Logical View

Logical View

Logical View
**Protocol Encapsulation**

At any layer:
- The higher level protocol headers are just treated like data
- Lower level protocol headers can be ignored

An ethernet switch or ethernet driver sees this:
- Ethernet Header
- Ethernet payload
- Ethernet Header
- Ethernet payload

A router or IP driver sees this:
- IP payload
- IP header
- IP payload
- IP header

A TCP driver sees this:
- TCP payload
- TCP header
- TCP payload
- TCP header

An application sees this:
- TCP payload
- TCP header
- TCP payload
- TCP header

**Addressing machines (data link layer)**

Each interface on a host has a unique MAC address
- E.g., aramis.rutgers.edu: 48-bit ethernet address = 00:03:ba:09:1b:b0
- This isn’t too interesting to us as programmers
  - We can send ethernet frames to machines on the same LAN

**Addressing machines (network layer)**

Each interface on a host is given a unique IP address
- IPv4 (still the most common in the U.S.): 32-bit number
  - Example, cs.rutgers.edu = 128.6.4.2 = 0x80060402
- IPv6: 128-bit number
  - Example, cs.rutgers.edu = 0:0:0:0:0:FFFF:128.6.4.2 = ::FFFF:8006:0402
- This also isn’t too interesting to us as programmers
  - We can send IP packets to machines on the Internet
  - BUT … we want to talk to applications

**IP transport layer protocols**

IP gives us two transport-layer protocols
- TCP: Transmission Control Protocol
  - Connection-oriented service
  - Operating system keeps state
  - Full-duplex connection: both sides can send messages over the same link
  - Reliable data transfer: the protocol handles retransmission
  - In-order data transfer: the protocol keeps track of sequence numbers
  - Flow control: receiver stops sender from sending too much data
  - Congestion control: “plays nice” on the network
- UDP: User Datagram Protocol
  - Connectionless service: lightweight transport layer over IP
  - Data may be lost
  - Data may arrive out of sequence
  - Checksum for corrupt data: operating system drops bad packets
  - Sender can keep blasting data to the receiver even if it can’t keep up

**Reliable data transfer**

- Simplest approach: stop & wait
  - while (no ack)
    - send message
    - wait for ack
- Problem
  - In the best case, the next packet is delayed for the latency of the entire path
  - Inefficient use of data channel
- Solution
  - Cumulative acknowledgments
TCP acknowledgements

- **Sequence number**
  - Starting byte sequence # of the data chunk being sent

- **Acknowledgement number**
  - Receiver responds with the next byte it's expecting
  - **Cumulative acknowledgments**
    - A receiver may receive a bunch of messages — respond only with the first missing byte #
    - Out-of-order segments may be buffered at the receiver
    - Will be sent to the application in order

The sender can send multiple messages without waiting for an acknowledgement

Controlling traffic: flow control

- Sending big chunks of data is more efficient than small chunks
  - Especially if we have to wait for acknowledgements
- Sending & sending & sending is more efficient than sending & waiting for acknowledgements
  - But we don’t want to have a fast sender overwhelm a slow receiver

- **Flow control**
  - Allow a receiver to control the rate of transmission
  - TCP receive window: # bytes receiver is willing to accept

Data rate = window size ÷ round-trip-time

Controlling traffic: congestion control

- **Additive increase**
  - Increase window size by one segment each round-trip time

- **Multiplicative decrease**
  - If we lose a segment (timeout occurs)
  - Decrease window size by 2

- **Additive Increase, Multiplicative Decrease (AIMD)**

Addressing applications (transport layer)

Communication endpoint at the machine

- **Port number**: 16-bit value
  - Port number = transport endpoint
  - Identifies a specific data stream
  - Some services use well-known port numbers (0 – 1023)
  - IANA: Internet Assigned Numbers Authority (www.iana.org)
  - Also see the file etc/services

- Ports for proprietary apps: 1024 – 49151
- Dynamic/private ports: 49152 – 65535

To communicate with applications, we use a transport layer protocol and an IP address and port number

Network API

- App developers need access to the network
- A **Network Application Programming Interface (API)** provides this
  - Core services provided by the operating system
    - Operating System controls access to resources
  - Libraries may handle the rest
- We will only look at IP-based communication

Programming: connection-oriented protocols

- **analogous to phone call**
  1. establish connection
  2. [negotiate protocol]
  3. exchange data
  4. terminate connection

- **virtual circuit service**
  - Provides illusion of having a dedicated circuit
  - Messages guaranteed to arrive in-order
  - Application does not have to address each message

Not to be confused with virtual circuit networks

- Which provide constant latency & guaranteed bandwidth
- TCP simulates a virtual circuit network... sort of
Programming: connectionless protocols

- no call setup
- send/receive data
- (each packet addressed)
- no termination

Datagram service

- client is not positive whether message arrived at destination
- no state has to be maintained at client or server
- cheaper but less reliable than virtual circuit service

Sockets

- Dominant API for transport layer connectivity
- Created at UC Berkeley for 4.2BSD Unix (1983)
- Design goals
  - Communication between processes should not depend on whether they are on the same machine
  - Communication should be efficient
  - Interface should be compatible with files
  - Support different protocols and naming conventions
  - Sockets is not just for the Internet Protocol family

What is a socket?

Abstract object from which messages are sent and received

- Looks like a file descriptor

- Application can select particular style of communication
  - Virtual circuit (connection-oriented), datagram (connectionless), message-based, in-order delivery

- Unrelated processes should be able to locate communication endpoints
  - Sockets can have a name
  - Name should be meaningful in the communications domain
    - E.g., Address & port for IP communications

Connection-Oriented (TCP) socket operations

Connectionless (UDP) socket operations

The end