Data Delivery

How packets travel hither and thither

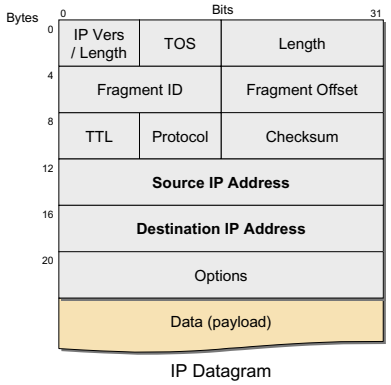
UNIX Network Administration
CIS 68C2

Data Delivery

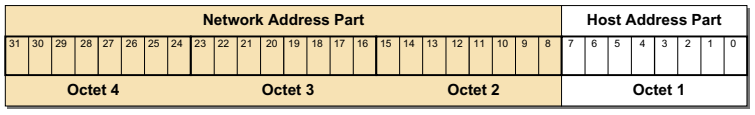
- Three crucial parts to IP data delivery
  - Addressing
    - How the destination host is uniquely identified
  - Routing
    - How data is switched across networks
  - Multiplexing
    - How data is delivered to a service running on a host
- Each part is involved in every packet's journey

Addressing

- The IP Address
  - Always 32 bits
    - Not byte-oriented
  - IP Datagram words 4 & 5
  - Comprised of two parts
    - **Network Address** part
    - **Host Address** part

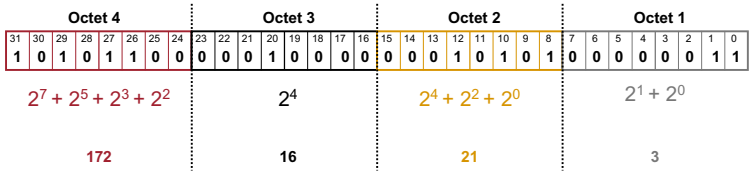


IP Datagram



Addressing

- The IP Address
  - Specifying an IP address as 32 bits of 0's and 1's is painful
  - Instead a more human readable form is used
  - Comprised of 4 integers, separated by dots
    - Each integer is the decimal value of the corresponding octet



172.16.21.3
IP Address

Addressing

- Three Forms of Addressing
  - Unicast
    - Packets are targeted to a single interface
      - **Destination Address** of the IP datagram is interface's IP address
  - Broadcast
    - Packets are targeted to all interfaces attached to the network
    - Broadcast packets are not transmitted across routers
  - Multicast
    - Frame targeted to only *subscribed* interfaces
      - An interface must be programmed to receive a multicast address
    - If routers support multicast, it may forward packet (if requested)

Addressing

- Multi-homing
  - A system with multiple IP interfaces
  - IP Addresses are …
    - …individually assigned to each NIC
    - …really interface addresses, not *host* addresses
      - Term *host address* is common, but *interface address* is more accurate
  - Consider routing system on two networks, A & B
    - Known by one IP address on network A, but by another IP address on network B

Addressing - Classes

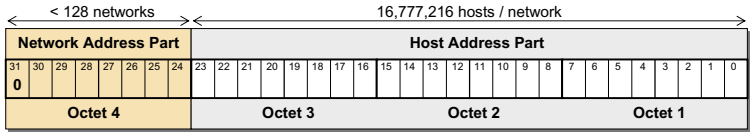
- Traditional IP Address Classes
  - IP addresses were traditionally divided into classes
    - Class A
    - Class B
    - Class C
    - Class D (Multicast)
  - The class defines:

    a) The range of valid IP addresses, and

    b) The maximum number of hosts possible on a network

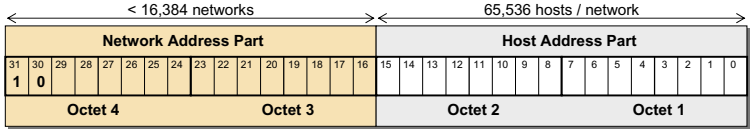
Addressing - Classes

- Class A Addresses
  - Network address part is the first 8 bits
    - Bit 31 of IP address is 0
  - Host address part is the last 24 bits
  - IP addresses from 1.x.x.x to 127.x.x.x
  - Number of networks: < 128 (2^7)
  - Maximum hosts/network: 16,777,216 (2^{24})



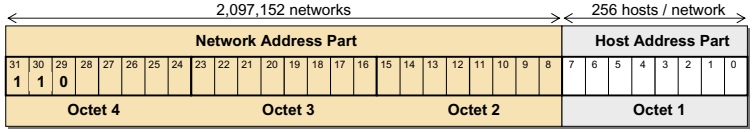
Addressing - Classes

- Class B Addresses
  - Network address part is the first 16 bits
    - Bits 31-30 of IP address is 1 0
  - Host address part is the last 16 bits
  - IP addresses from 128.x.x.x to 191.x.x.x
  - Number of networks: < 16,384 (2^{14})
  - Maximum hosts/network: 65,536 (2^{16})



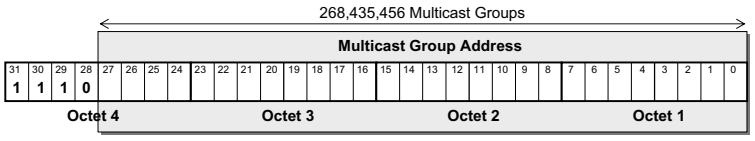
Addressing - Classes

- Class C Addresses
  - Network address part is the first 24 bits
    - Bits 31-29 of IP address is 1 1 0
  - Host address part is the last 8 bits
  - IP addresses from 192.x.x.x to 223.x.x.x
  - Number of networks: < 2,097,152 (2^{21})
  - Maximum hosts/network: 256 (2^{8})



Addressing - Classes

- Multicast Addresses (*aka* Class D)
  - No network part
    - Bits 31-28 of IP address is 1 1 1 0
    - Entire address specifies multicast group
  - Multicast group address is last 28 bits
  - IP addresses from 224.x.x.x to 239.x.x.x
  - Number of multicast groups: < 268,435,456 (2^{28})



Addressing - Classes

- Problems with Address Classes
  - Class A
    - Too many hosts per network; too few networks
  - Class B
    - Enough hosts per network; too few networks
  - Class C
    - Provides plenty of networks; too few hosts per network
  - IP addresses were not distributed geographically
    - Creates routing inefficiencies and aggregating problems
      - Today, IP ranges are given to large ISPs

Addressing - CIDR

- Classless Internet Domain Routing - CIDR
  - Allows flexible division between network and host part of an IP address
  - Temporary solution until IPv6 deprecates IPv4
  - Requires a network mask (**netmask**)
    - An IP address that differentiates network part from host part
    - The **network mask** that corresponds to class A, B, or C addresses is called the **default mask** or **natural mask**
  - Required router & routing protocols modifications
    - CIDR unsupported in older operating systems and older routing protocols such as RIP

Addressing - CIDR

- Specifying Network Addresses
  - IP address / network mask pair is cumbersome
  - An abbreviated notation is used instead
    - Specify IP address followed by the number of bits used for the network part
      - *IP-address/network-length*
    - Example
      - 192.4.0.0/16
        - Netmask 255.255.0.0
      - 172.16.26.32/27
        - Netmask 255.255.255.224

Addressing - CIDR

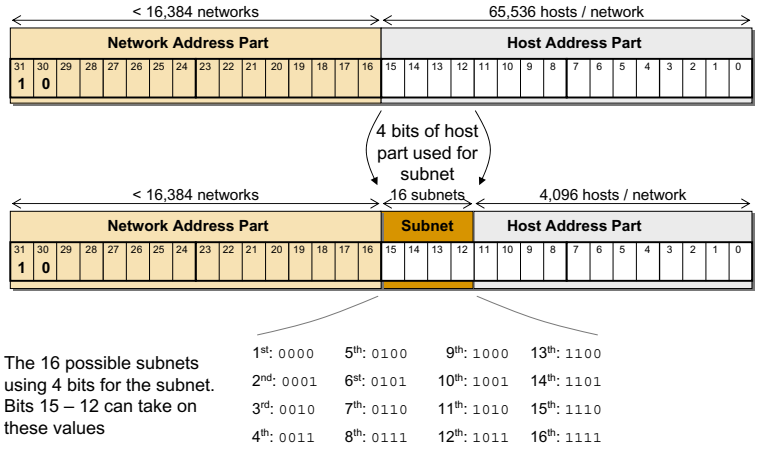
- Supernetting
  - Creating a network with more hosts than the traditional class address allows
  - For IP address range 195.4.0.x -> 195.4.255.x
    - As class C addresses, limited to
      - 256 networks, with 256 hosts/network
    - Supernetting allows more options:
      - 1 network; 65536 hosts/network
      - 2 networks; 32768 hosts/network
      - 16 networks; 4096 hosts/network
      - 128 networks; 512 hosts/network
      - etc...

Addressing - CIDR

- Subnetting
  - Using the upper bits of the host part of an IP address to subdivide a single network into several independent networks
  - The upper host bits become the internal subnet number
  - Used locally to increase the number of networks
    - At the cost of reducing the number of hosts per network
  - Benefits
    - Allows departmental management of IP addresses
    - Allows routers to seemingly join networks
      - Routers must address two different networks
      - Overcomes distance limitations, or hardware differences

Addressing - CIDR

- Example: Subnetting a Class B Address



Addressing - CIDR

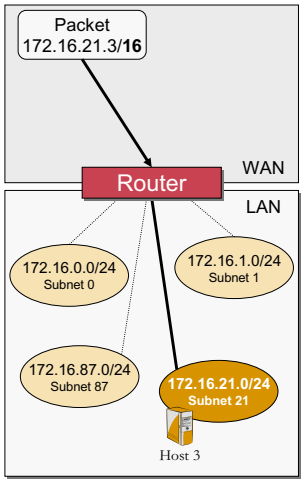
- Subnetting
  - Subnet mask vs. Network mask
    - The term **subnet mask** is used to refer to the network mask applied to an organization's internal routes for a subnetted network
    - Network mask is the more general term, used typically for non-subnetted networks
  - Either way, it is a **route** that has a network mask

Addressing - CIDR

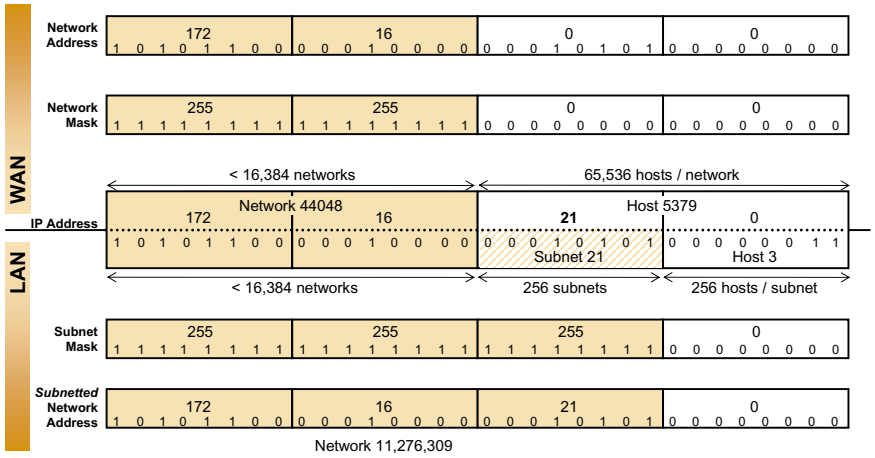
- Subnetting
  - Example
    - IP Address: 172.16.21.3/16
    - Traditionally considered a Class B address
    - Natural network mask is 255.255.0.0
      - Allows up to 65536 hosts
    - Using subnet mask of 255.255.255.0
      - Reduces number of hosts/network to 256
      - Provides 256 subnetted networks
      - LAN routers route packets from one subnet to another

Addressing - CIDR

- Subnetting
  - IP Address 172.16.21.3/16
  - On WAN
    - Appears as host number 5379 (x.x.**21.3**) on network **172.16**.0.0/16
    - Reaches net using network mask of 255.255.0.0
  - Within LAN
    - Packet is routed to subnetted network 172.16.21.x
      - Host 3 on the subnet receives packet
      - Subnet mask of 255.255.255.0



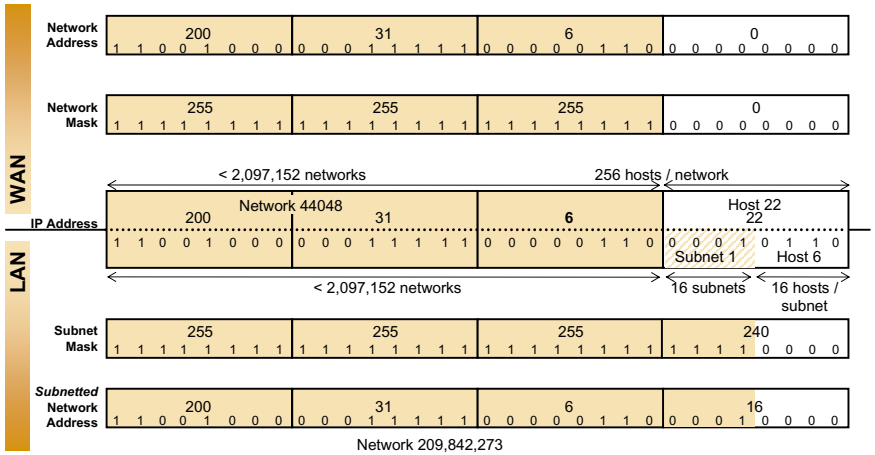
Addressing - CIDR



Addressing - CIDR

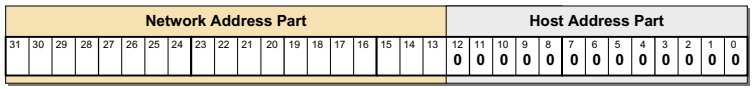
- Subnetting
  - Example
    - IP Address: 200.31.6.22
    - Class C
    - Natural network mask is 255.255.255.0
      - Allows up to 256 hosts
    - Using subnet mask of 255.255.255.240
      - Reduces number of hosts/network to 16
        - Only 4 bits now for the host part
      - Provides 16 subnetted networks
        - 4 extra bits for the network part

Addressing - CIDR

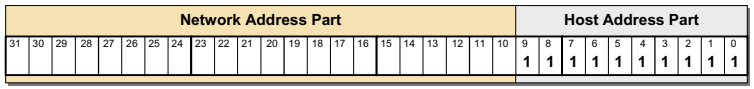


Addressing – Reserved IPs

- Reserved Host Numbers
  - All 0's host part of IP address
    - Refers to the **network** itself
    - Used by routers and routing software



  - All 1's host part of IP address
    - Is the **broadcast** address for the network



Addressing – Reserved IPs

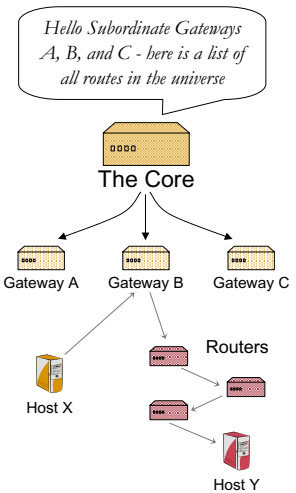
- Reserved Network Numbers
  - 0.0.0.0 (Class A)
    - The network itself
    - Simplifies routing table entries
  - 127.0.0.0 (Class A)
    - The **loopback** network address
    - Simplifies programming network applications
    - The standard hostname for the loopback interface is **localhost**
  - 223.x.x.x to 239.x.x.x (Multicast)
  - 240.x.x.x to 255.x.x.x are reserved for future use

Addressing – Private IPs

- Addresses for Private Internets
  - Class A: 10.0.0.0
  - Class B: 172.16.0.0 to 172.31.255.255
  - Class C: 192.168.0.0 to 192.168.255.255
  - For use in private internets only
    - Requires no coordination with in Internet registry
  - Must not be connected outside the *enterprise*
    - These systems cannot connect directly to the Internet
  - Reduces demands on limited IP address pool
  - See also: RFC 1918

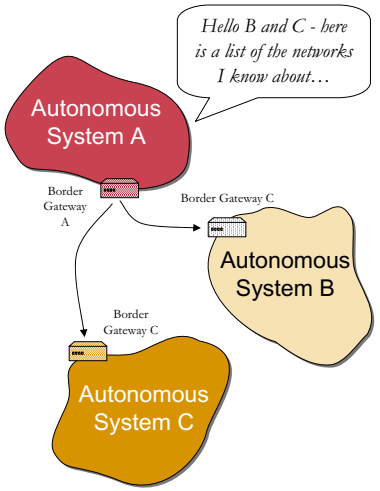
Routing

- Routing
  - How do packets find their way?
    - Hosts only send packets to hosts on the same network
  - Old way - Hierarchical System
    - The Core – the central system
    - Subordinate Core Gateways
      - Each contained **all** information about internet
    - Gateway to Gateway Protocol - **GGP**
    - Major Weakness
      - All routes processed by the Core!



Routing

- New model
  - Routing Domains
    - Collections of co-equal, autonomous systems
    - Non-hierarchical
      - A central core system is not required for best routes
    - Border Gateway Protocol - **BGP**
      - The BGP protocol is used to exchange routing (**reachability**) information
      - Includes associated **cost**
    - Problem still remains – who decides best routes?



Routing

- Routing Table
  - All IP network devices make routing decisions
    - For hosts, decision is simple
      - Local network packets are addressed to local host
      - Packets destined for other networks are addressed to the local gateway
    - Recall this occurs at the IP layer
      - Routing decision based on network part of IP address
        - Network part determined from the network mask
      - If packet is for local network
        - Either the subnet mask or default mask is used to determine destination network

Routing

- Routing Table
  - Also called the **forwarding table**
  - Every IP device has a routing table
  - Tells IP where to deliver packets
    - Uses network number to look up routes in table
    - Assigns packet's destination IP address
  - Entries in table
    - Are considered routes to other networks
    - Can be created dynamically or statically
  - Routing is just a table lookup!

Routing

- Routing Table
  - The **netstat** utility shows a host's routing table
    - Flags: U = up; G = gateway; H = host
    - Gateway: means *goes through given interface*
    - Ref: number of active uses of route (eg. by TCP)
    - Use: number of packets sent through route

```
$ netstat -rn

Routing Table:
  Destination           Gateway           Flags  Ref   Use   Interface
-------------------- -------------------- ----- ----- ------ ---------
153.18.0.0           153.18.75.207         U        3    186  elx0
224.0.0.0            153.18.75.207         U        3      0  elx0
default              153.18.75.254         UG       0    140
127.0.0.1            127.0.0.1             UH       0   6877  lo0
```


Address Resolution

- The problem
  - Physical networks use physical addresses, not IP addresses
  - Need the physical address of the destination host
- The Solution
  - Mapping of IP address to physical hardware address
  - Performed by TCP/IPs Network Access layer
  - Address Resolution Protocol - **ARP**
    - Performs IP to Ethernet translation
    - Maintains dynamic table of IP/Ethernet pairs

Address Resolution

- ARP
  - Requests are made to translate IP addresses
  - ARP looks in its ARP Table
    - If IP address is found, Ethernet address is returned
    - If not found, ARP broadcasts request to network
      - Receiving host then returns its Ethernet address
      - IP/Ethernet address pair is cached in ARP table for future
  - ARP Requests
    - Only published addresses are returned
      - A host normally publishes and returns its own Ethernet address
      - A host may return Ethernet address for other hosts
        - Called Proxy ARP

Address Resolution

- Proxy ARP
  - Allows servers to respond for remote hosts that may not be able to respond
    - Server responds with its own Ethernet address
    - Packets destined for remote host are sent to server
    - Server forwards packets to remote host in proxy
      - Proxy means – to act on another's behalf

Address Resolution

- The arp command shows the ARP table
  - Flags: p = Publish; s = Static; m = Mapping

```
$ arp taipei
taipei (153.18.75.207) at 0:10:4b:35:7e:18 permanent published
$ arp -a
Net to Media Table
Device   IP Address               Mask      Flags   Phys Addr
------ -------------------- --------------- ----- ---------------
elx0   losaltos             255.255.255.255       00:10:4b:21:00:c1
elx0   153.18.75.254        255.255.255.255       00:10:f6:ac:58:00
elx0   tiptoe.fhda.edu      255.255.255.255       00:00:0c:45:bb:32
elx0   taipei               255.255.255.255 SP    00:10:4b:35:7e:18
elx0   153.18.71.251        255.255.255.255       00:00:0c:45:bb:32
elx0   153.18.82.12         255.255.255.255       00:04:ac:49:d5:95
elx0   BASE-ADDRESS.MCAST.NET 240.0.0.0       SM    01:00:5e:00:00:00
```


Multiplexing

- Protocols Numbers
  - Which protocol above IP receives packet
    - UPD, TCP, etc.
    - 3rd word of IP datagram
    - **/etc/protocols**

```
# Internet (IP) protocols
#
ip              0       IP              # internet protocol
icmp            1       ICMP            # internet control message protocol
ggp             3       GGP             # gateway-gateway protocol
tcp             6       TCP             # transmission control protocol
egp             8       EGP             # exterior gateway protocol
pup             12      PUP             # PARC universal packet protocol
udp             17      UDP             # user datagram protocol
```

Multiplexing

- Port Numbers
  - Allows routing data within the destination host
    - to a specific network service or application process
      - E.g. FTP, HTTP, etc.
    - Source port / Destination port
      - 1st word of TCP segment or UDP message
  - Port numbers are listed in the **/etc/services** table

Multiplexing

- /etc/services
  - Table of port numbers / protocols / service names
  - Ports 1-1023 are the **well-known** ports
    - Well-known ports simplify the connection process
    - Should not be bound to non-privileged processes
    - Below 256 are standard internet services such as FTP, telnet, http
    - Ports 256 – 1024 were originally UNIX-specific services
      - rlogin, lpr, etc.
      - No longer UNIX-specific
  - Ports 1024-49151 are the **registered** ports
  - Ports from 49152 – 65536 are the **private** ports

Multiplexing

- /etc/services

```
# Network services, Internet style
#
tcpmux          1/tcp
systat          11/tcp          users
daytime         13/tcp
daytime         13/udp
netstat         15/tcp
chargen         19/tcp          ttytst source
chargen         19/udp          ttytst source
ftp-data        20/tcp
ftp             21/tcp
telnet          23/tcp
smtp            25/tcp          mail
time            37/tcp          timserver
time            37/udp          timserver
http            80/tcp          httpd
```


Multiplexing

- Portmapper
  - Multiplexes the Remote Procedure Call service - **RPC**
    - Port 111
  - Eliminates need for using well-known port
  - **/etc/rpc**

```
rpcbind         100000  portmap sunrpc rpcbind
rstatd          100001  rstat rup perfmeter
rusersd         100002  rusers
nfs             100003  nfsprog
ypserv          100004  ypprog
mountd          100005  mount showmount
ypbind          100007
walld           100008  rwall shutdown
yppasswdd       100009  yppasswd
```

Multiplexing

- Sockets
  - An IP address / port number pair
  - A one-way connection
  - Two sockets are required for connection-oriented services
  - A socket is unique on the Internet
- Concurrent usage of the same service
  - Host opens connection to well-known destination port, using dynamically-allocated random, high-numbered source port