CS 455/555

Intro to Networks and Communications

The Internet Protocol (IP)

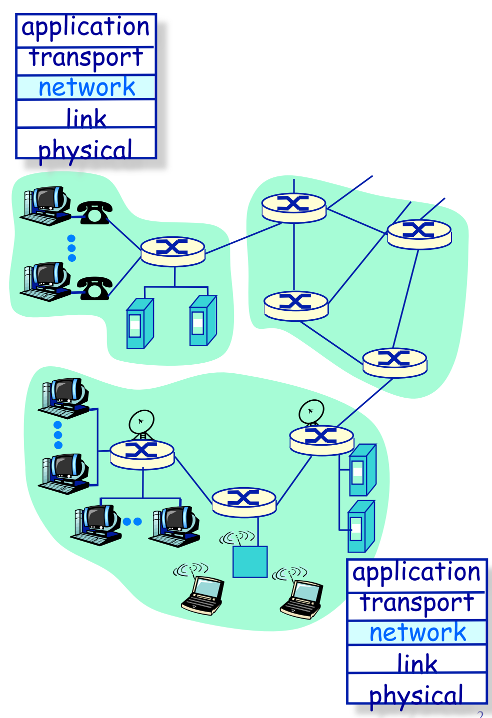
Dr. Michele Weigle
Department of Computer Science
Old Dominion University
mweigle@cs.odu.edu

http://www.cs.odu.edu/~mweigle/CS455-S13

The Network Layer: Routing & Addressing

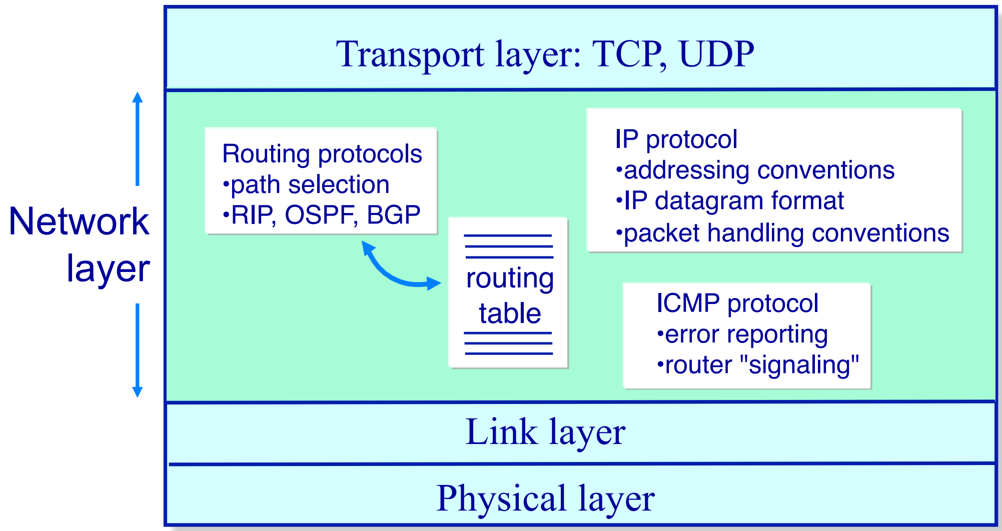
Outline

- Network layer functions
- Virtual circuits and datagram networks
- Router architecture
- IP Internet Protocol
  - » Addressing
- Routing algorithms
  - » Least cost path computation algorithms
- Hierarchical routing
  - » Connecting networks of networks
- Routing on the Internet
  - » Intra-domain routing
  - » Inter-domain routing



The Internet Network Layer

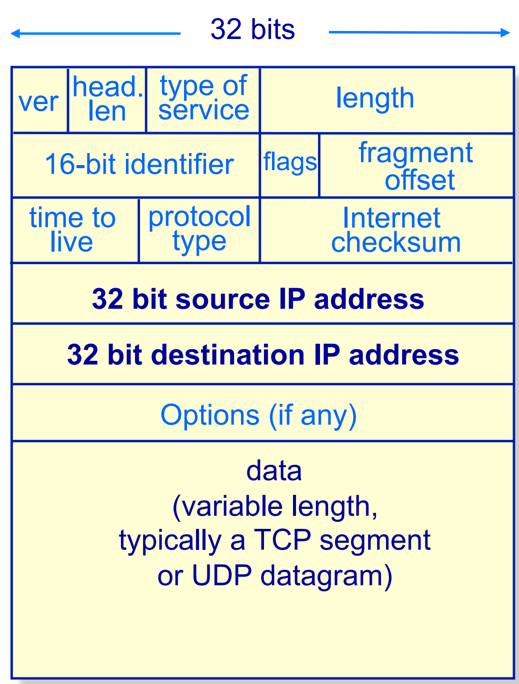
Host and router network layer functions



The Internet Network Layer

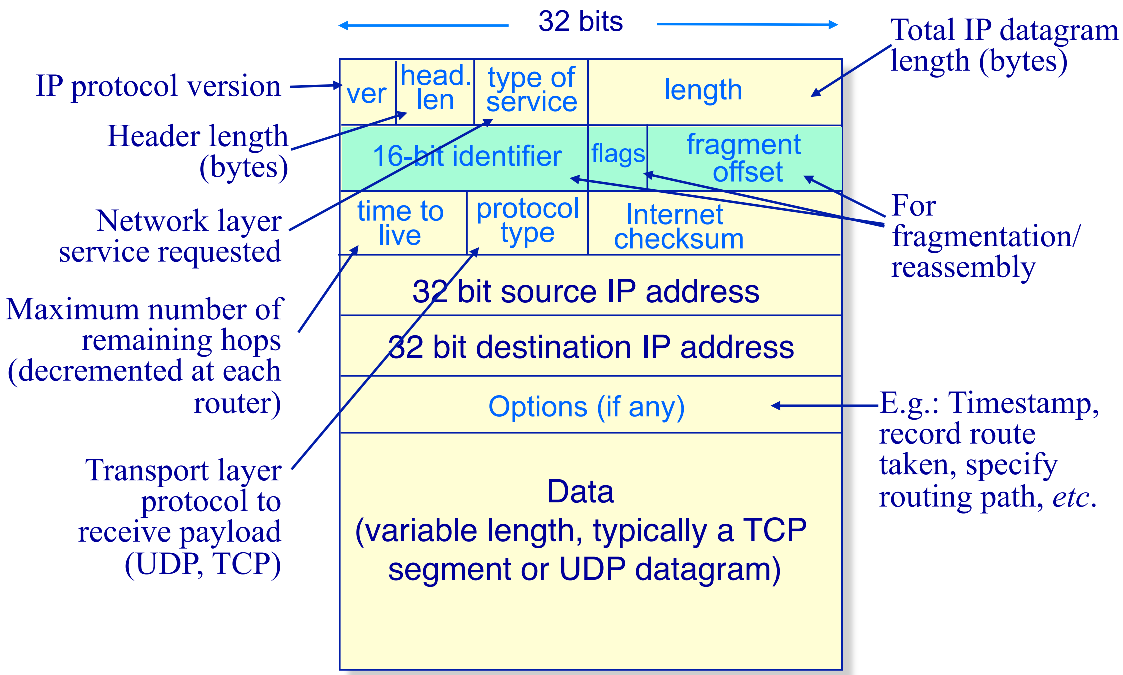
IP datagram format

- IP datagrams
  - » The protocol data units at the IP network layer)
- (Not to be confused with UDP datagrams)
  - » The protocol data units at the UDP transport layer are also called datagrams



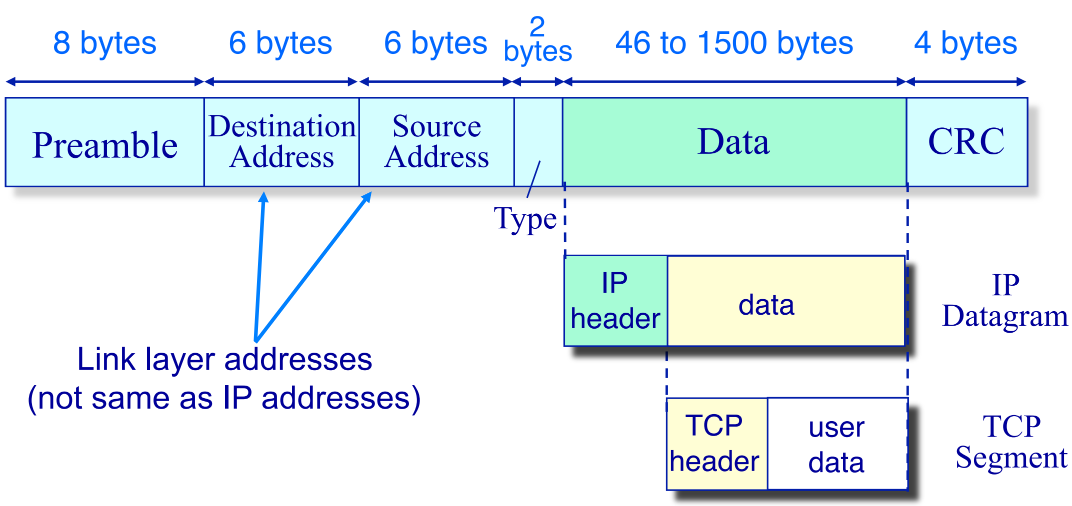
IP Datagrams

Details



IP Datagrams

Encapsulation (Ethernet)

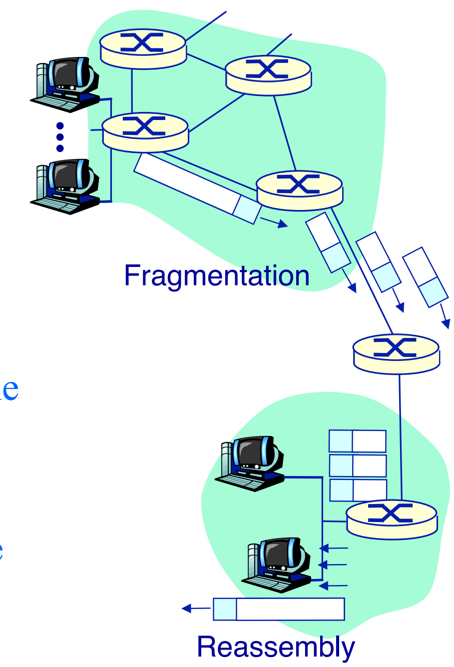


- Sending interface adapter encapsulates IP datagram (or other network layer protocol packet) in an *Ethernet frame*

IP Datagrams

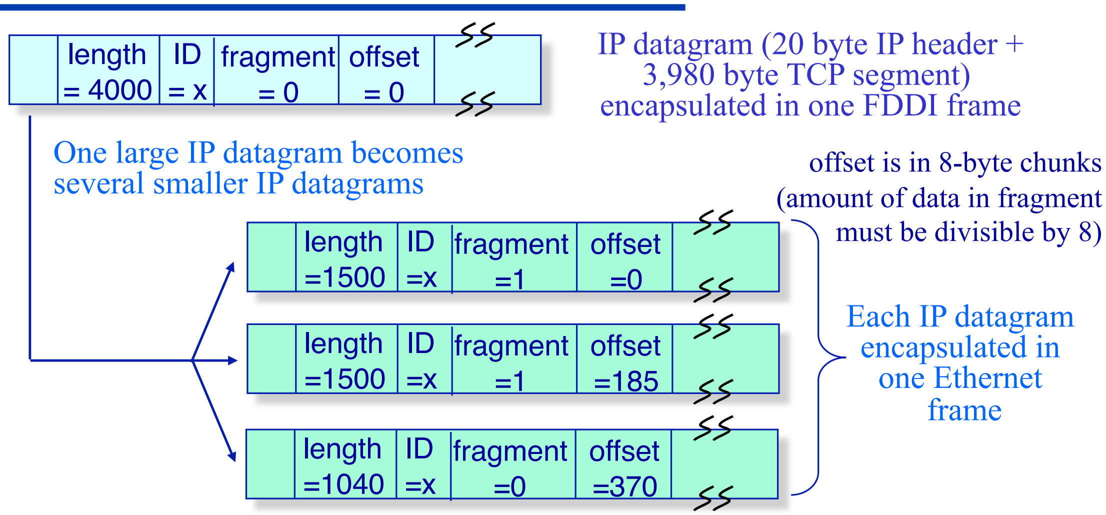
Fragmentation & Reassembly

- Network links have a maximum frame size
  - » Called the *maximum transmission unit* (MTU)
  - » Different link types, different MTUs
- Large IP datagrams must be "fragmented" to link MTU sizes
  - » One IP datagram becomes several IP datagrams as it transits networks
  - » "Fragments" reassembled only at the final destination
- All fragments carry the same IP identification number
  - » All fragments (except the last) have the fragment bit set



IP Fragmentation and Reassembly

Ethernet MTU example



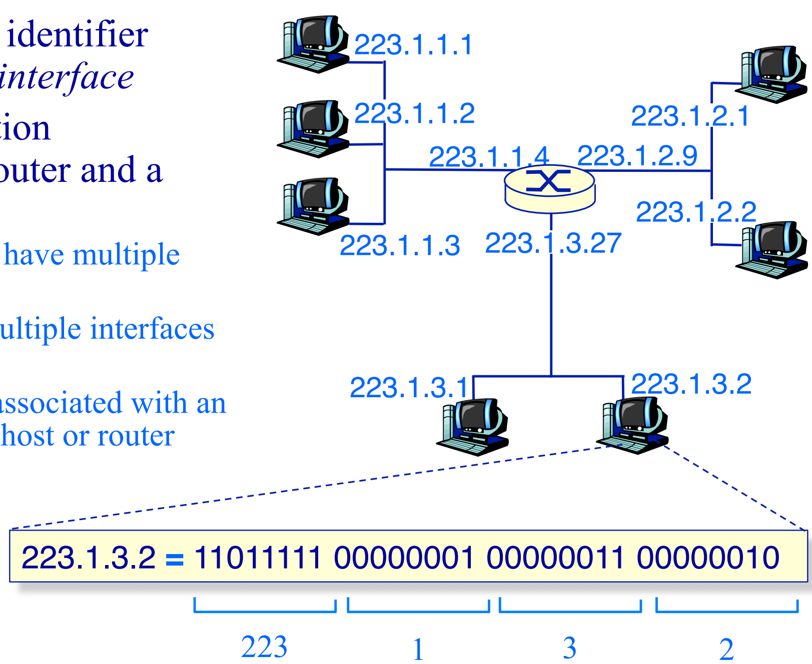
- Consider a 3,980 byte message sent in an FDDI frame (MTU 4000B)
- The message generates 3 fragments when it transits an Ethernet (MTU 1500 B)
  - » How much application data is in each fragment?

Problem

- Original MTU of 4000 B
- Sending IP datagram of 2000 B (including IP header)
- Transits over network with MTU of 576 B
- How is datagram fragmented?

IP Addressing
Introduction

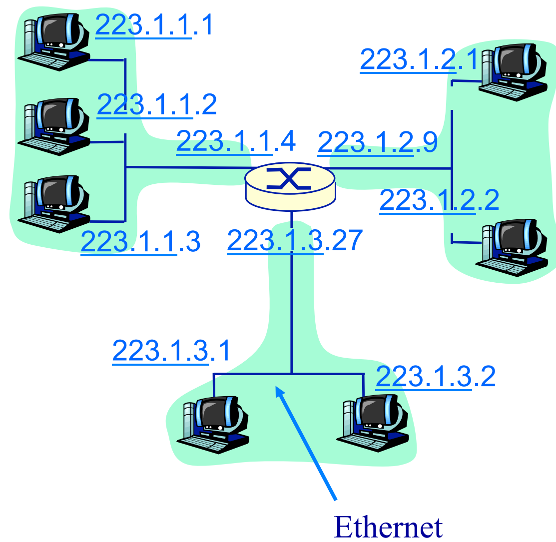
- IP address: 32-bit identifier for host or router *interface*
- Interface: connection between host or router and a physical link
  - » Routers typically have multiple interfaces
  - » Host *may* have multiple interfaces (typically not)
  - » IP addresses are associated with an interface, *not* the host or router



IP Addressing

Host address *v.* Network addresses

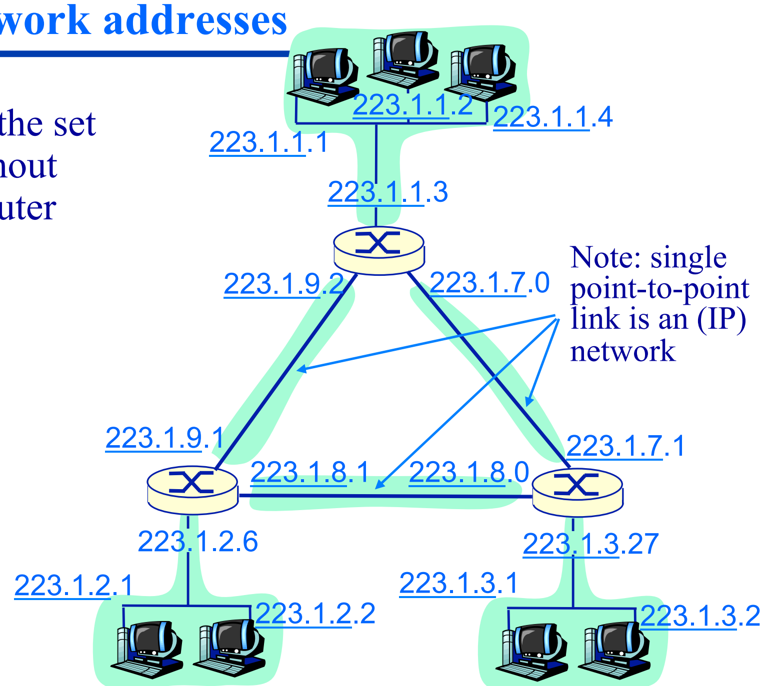
- IP address:
  - » Network part (high order bits)
  - » Host part (low order bits)
- What's a network?
  - » The set of devices that can communicate with each other without an intervening router
    - ❖ The devices attached to the same physical network
  - » From an IP address perspective its:
    - ❖ The set of device interfaces with IP addresses having a common network part



IP Addressing

Host address *v.* Network addresses

- A network (*subnet*) is the set of hosts reachable without having to traverse a router
  - » Detach each interface from router or host
  - » Create "islands" of isolated networks



Routing IP Datagrams
Example

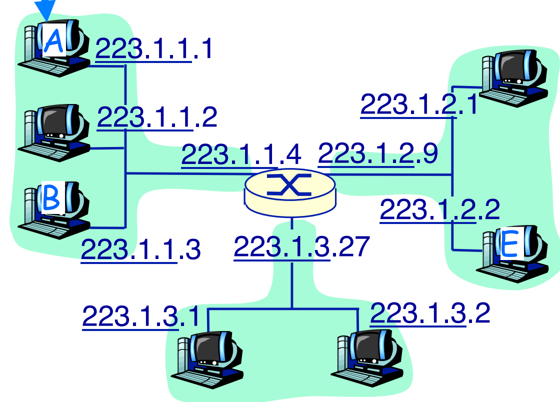
misc fields	source IP addr	dest IP addr	data

IP Datagram

- All routing is based on the IP destination address field in the IP header
- IP destination address (and data fields) never change!
  - » Delivery to intermediate hops involves link-layer addresses

Routing table in A

Dest. Net.	next router	Nhops
223.1.1/24	-	1
*(default)	223.1.1.4	2



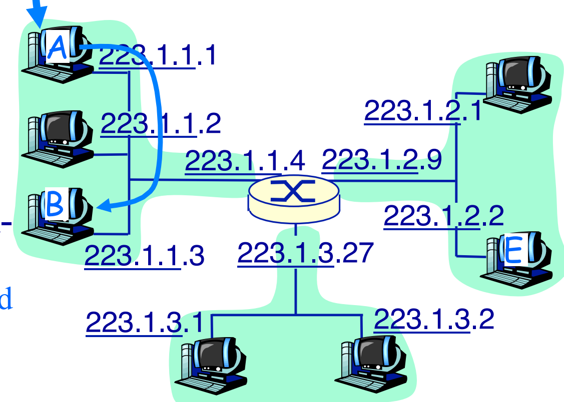
Routing IP Datagrams
Routing to a local destination

misc fields	223.1.1.1	223.1.1.3	data

- An application on A generates an IP datagram addressed to B
  - » The IP layer on A looks up the network address of B...
  - » And determines that B is on same network as A (223.1.1)
- A's link layer sends the IP datagram directly to B inside link-layer frame
  - » B and A are assumed to be connected to the same physical network

Routing table in A

Dest. Net.	next router	Nhops
223.1.1/24	-	1
*(default)	223.1.1.4	2



Routing IP Datagrams

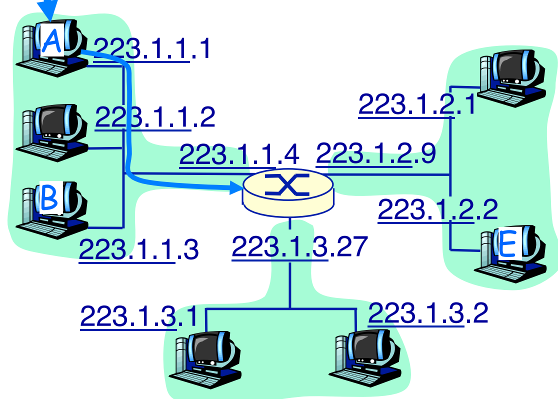
Routing to a remote destination

misc fields	223.1.1.1	223.1.2.2	data

Routing table in A

Dest. Net.	next router	Nhops
223.1.1/24	-	1
*(default)	223.1.1.4	2

- Host A generates an IP datagram addressed to E
  - » The IP layer on A looks up up network address of E (223.1.2)
  - » A determines that E is NOT on same network as A
  - » A's routing table shows router 223.1.1.4 as the default for all networks
- A's link layer sends IP datagram to router inside link-layer frame



Routing IP Datagrams

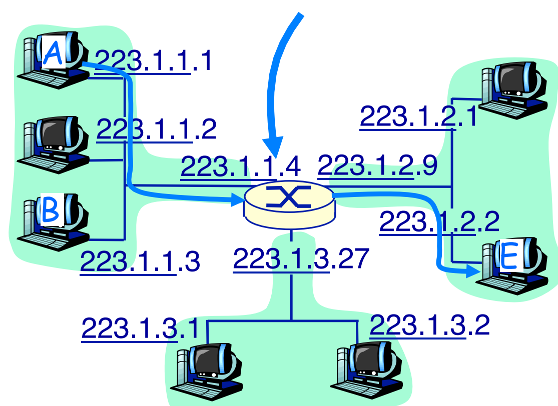
Routing to a remote destination

misc fields	223.1.1.1	223.1.2.2	data

Routing table in router

Dest. network	next router	Nhops	interface
223.1.1/24	-	1	223.1.1.4
223.1.2/24	-	1	223.1.2.9
223.1.3/24	-	1	223.1.3.27

- A's datagram addressed to E arrives at the router
  - » The router looks up network address of E (223.1.2)
  - » E has the *same* network address as router's interface 223.1.2.9
  - » Router is directly attached to the same network (223.1.2) as E
- Router's link layer sends the datagram to 223.1.2.2 inside a link-layer frame *via* interface 223.1.2.9
  - » Datagram arrives at 223.1.2.2



IP Addressing

Class-Based (or classful) Addressing

Class A 1.0.0.0 - 127.255.255.255

0	N N N N N N N	.H H H H H H H H.H H H H H H H H.H H H H H H H H
	7 bits **(128 networks)**	24 bits **(16,777,216 hosts)**

Class B 128.0.0.0 - 191.255.255.255

1 0	N N N N N N.N N N N N N N N	.H H H H H H H H.H H H H H H H H
	14 bits **(16,384 networks)**	16 bits **(65,536 hosts)**

Class C 192.0.0.0 - 223.255.255.255

1 1 0	N N N N N.N N N N N N N N.N N N N N N N N	.H H H H H H H H
	21 bits **(2,097,152 networks)**	8 bits **(256 hosts)**

IP Addressing

Special IP Addresses

Prefix	Suffix	Address Type	Purpose
all-0s	all-0s	this computer	bootstrap
network	all-0s	network	network ID
network	all-1s	directed bcast	bcast on specified net
all-1s	all-1s	limited bcast	bcast on local net
127	any	loopback	testing

- 127.0.0.1 is the most popular address for loopback
- Private addresses available
  - » 10.0.0.0 - 10.255.255.255
  - » 172.16.0.0 - 172.31.255.255
  - » 192.168.0.0 - 192.168.225.255

  (see RFC 1597)

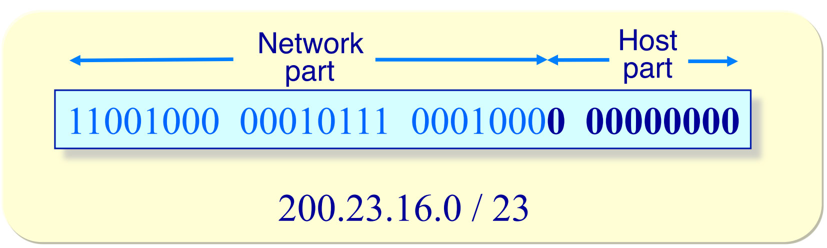
IP Addressing

Classless InterDomain Routing (CIDR)

- Why don't we use class-based addressing anymore?
- Today, we use *subnet addressing*, or *classless addressing* (CIDR)
- Netmask
  - » 32-bit number to identify which bits are network prefix in the IP address
  - » the network part bits are all 1 and the host part bits are all 0

IP Addressing

Classless InterDomain Routing (CIDR)

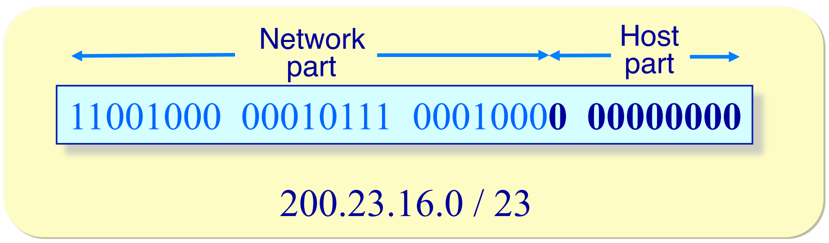


- Network portion of address has an arbitrary length
- Address format: a.b.c.d/x, where x is the number of bits in network portion of address
  - » used only in routing tables, not IP datagram source/ destination

IP Addressing
Classless Addressing (CIDR)

- With CIDR, 128.10.0.0 Class B address becomes 128.10.0.0/16
  - » first 16 bits are network prefix
  - » second 16 bits are host suffix

10000000	128
00001010	10
00000000	0
00000000	0



/8	**255.0.0.0**	**16777216**	**Class A**
/9	255.128.0.0	8388608	
/10	255.192.0.0	4194304	
/11	255.224.0.0	2097152	
/12	255.240.0.0	1048576	
/13	255.248.0.0	524288	
/14	255.252.0.0	262144	
/15	255.254.0.0	131072	
/16	**255.255.0.0**	**65536**	**Class B**
/17	255.255.128.0	32768	ISP / large business
/18	255.255.192.0	16384	ISP / large business
/19	255.255.224.0	8192	ISP / large business
/20	255.255.240.0	4096	Small ISP / large business
/21	255.255.248.0	2048	Small ISP / large business
/22	255.255.252.0	1024	
/23	255.255.254.0	512	
/24	**255.255.255.0**	**256**	**Class C**
/25	255.255.255.128	128	Large LAN
/26	255.255.255.192	64	Small LAN
/27	255.255.255.224	32	Small LAN
/28	255.255.255.240	16	Small LAN
/29	255.255.255.248	8	
/30	255.255.255.252	4	

CIDR Addressing
Example

- 130.127.128.0/20
  - » 20 bits network
  - » 12 bits host, 2^{12} = 4096, 4094 assignable IPs

network:	130	127	128	0
	1000 0010	0111 1111	1000 0000	0000 0000
first addr:	130	127	128	1
	1000 0010	0111 1111	1000 0000	0000 0001
last addr:	130	127	143	254
	1000 0010	0111 1111	1000 1111	1111 1110
netmask:	255	255	240	0
	1111 1111	1111 1111	1111 0000	0000 0000

CIDR Addressing
Problems

- 152.2.136.0/26
  - » How many assignable IP addresses?
  - » What is the range of assignable IP addresses?
  - » What is the proper netmask?

- range: 130.127.64.129 - 130.127.64.134 and netmask: 255.255.255.248
  - » What is the CIDR address?

IP Addresses

How are addresses assigned to a network?

How does *network* get subnet part of IP addr?

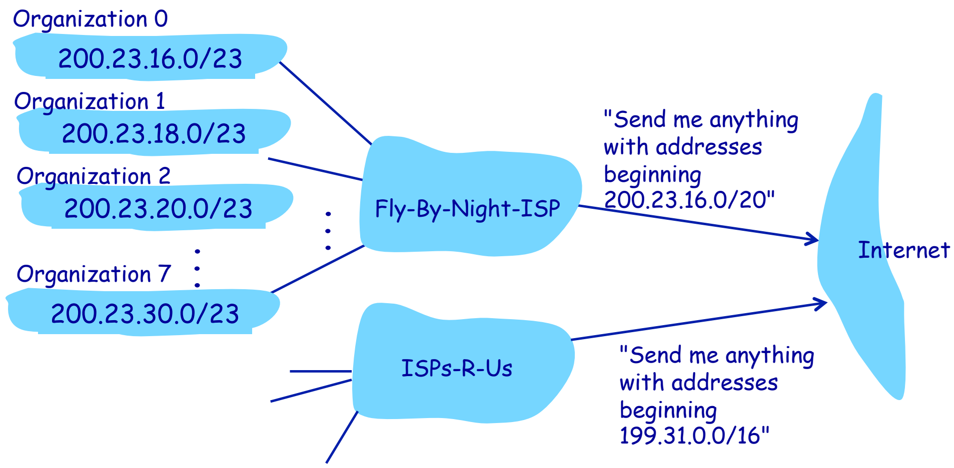
>> gets allocated portion of its provider ISP's address space

ISP's block	11001000 00010111 00010000	00000000	200.23.16.0/20
Organization 0	11001000 00010111 00010000	00000000	200.23.16.0/23
Organization 1	11001000 00010111 00010010	00000000	200.23.18.0/23
Organization 2	11001000 00010111 00010100	00000000	200.23.20.0/23
...	.....	....	....
Organization 7	11001000 00010111 00011110	00000000	200.23.30.0/23

IP Addresses

Hierarchical Addressing: route aggregation

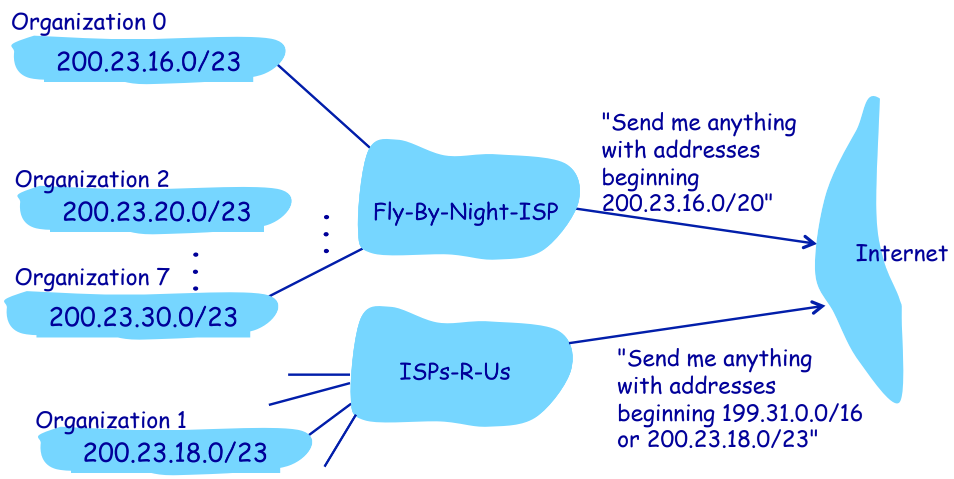
Hierarchical addressing allows efficient advertisement of routing information



IP Addresses

Hierarchical Addressing: more specific routes

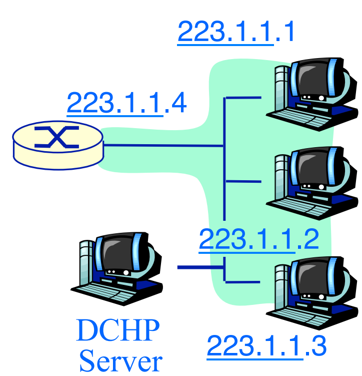
ISPs-R-Us has a more specific route to Organization 1



IP Addresses

How are addresses assigned to a host?

- Static assignment:
  - » Configuration parameter (manually) set during system installation
- Dynamic assignment at boot/ wake-up time
  - » DHCP: Dynamic Host Configuration Protocol

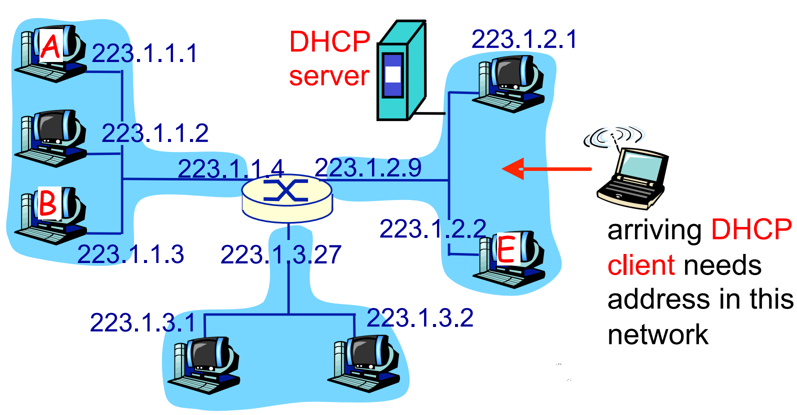


IP Addresses
DHCP: Dynamic Host Configuration Protocol

- Goal: allow host to dynamically obtain its IP address from network server when it joins network
  - » can renew its lease on address in use
  - » allows reuse of addresses (only hold address while connected and "on")
  - » support for mobile users who want to join network

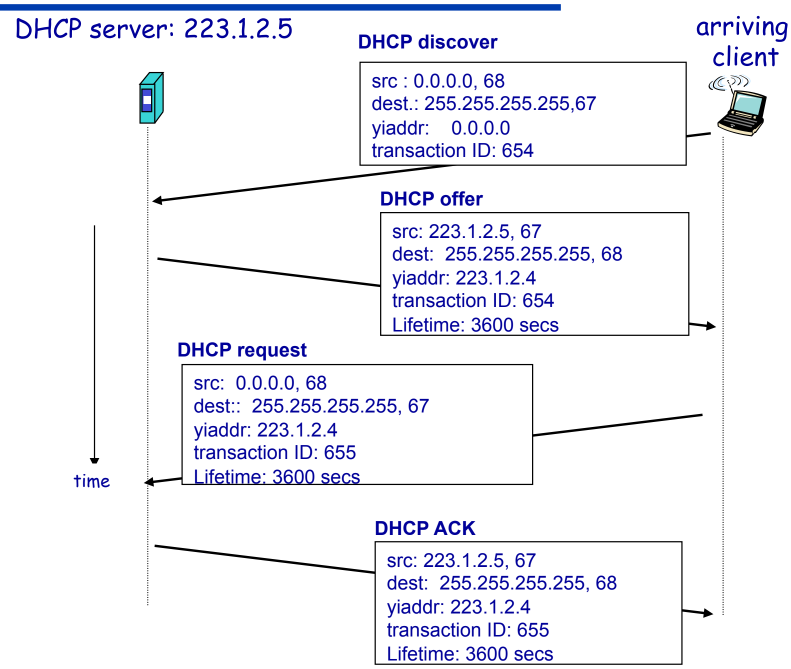
- DHCP overview
  - » host broadcasts "DHCP discover" msg
  - » DHCP server responds with "DHCP offer" msg
  - » host requests IP address: "DHCP request" msg
  - » DHCP server sends address: "DHCP ack" msg

DHCP
Client-Server Scenario



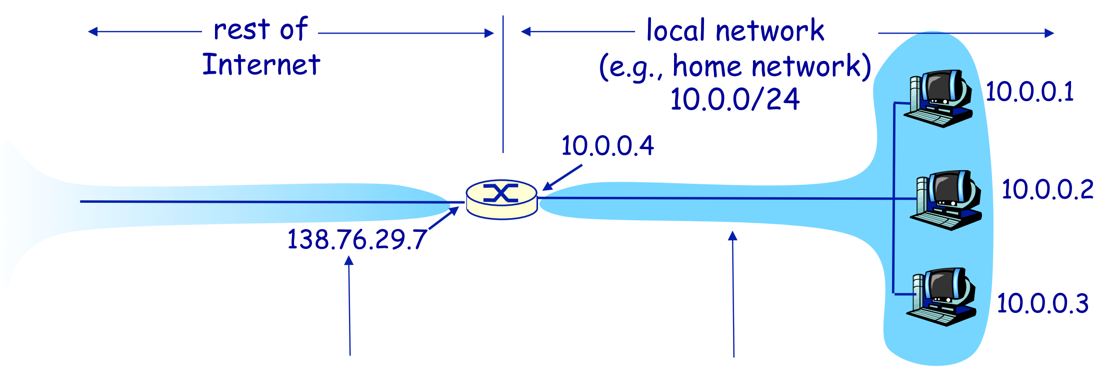
DHCP

Client-Server Scenario



The Internet Network Layer

NAT: Network Address Translation



All datagrams *leaving* local network have same single source NAT IP address: 138.76.29.7, different source port numbers

Datagrams with source or destination in this network have 10.0.0/24 address for source, destination (as usual)

NAT: Network Address Translation
Motivation

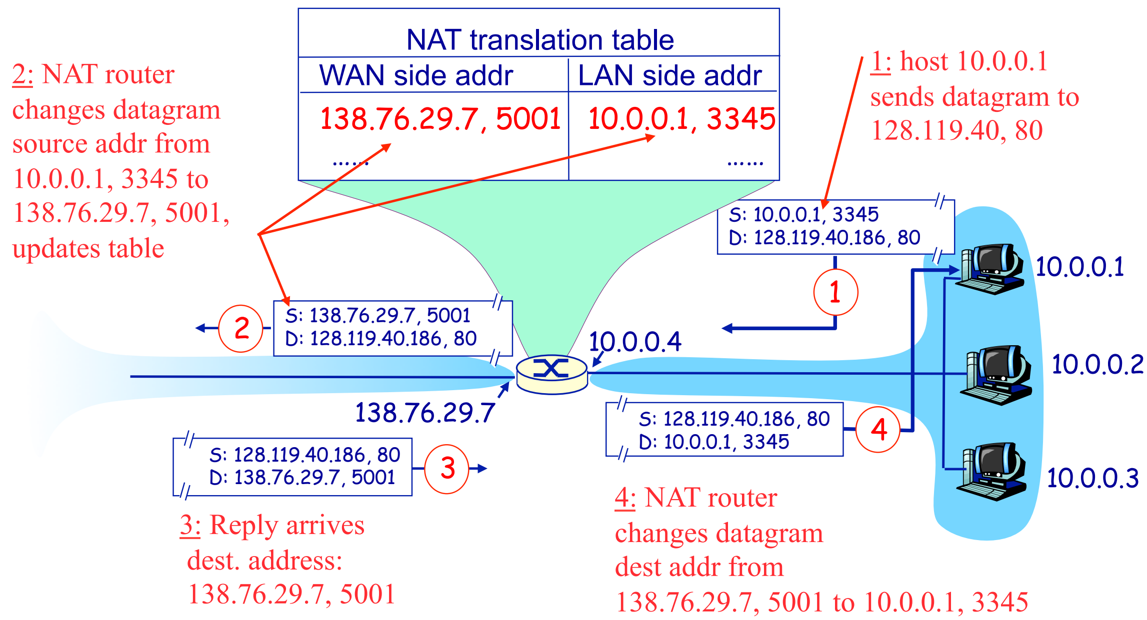
- Local network uses just one IP address as far as outside world is concerned:
  - » range of addresses not needed from ISP: just one IP address for all devices
  - » can change addresses of devices in local network without notifying outside world
  - » can change ISP without changing addresses of devices in local network
  - » devices inside local net not explicitly addressable, visible by outside world (a security plus).

NAT: Network Address Translation
Implementation

- NAT router must:
  - » *outgoing datagrams*: replace (source IP address, port #) of every outgoing datagram with (NAT IP address, new port #)
    - remote clients/servers will respond using (NAT IP address, new port #) as destination addr.
  - » *remember (in NAT translation table):* every (source IP address, port #) to (NAT IP address, new port #) translation pair
  - » *incoming datagrams:* replace (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

The Internet Network Layer

NAT: Network Address Translation

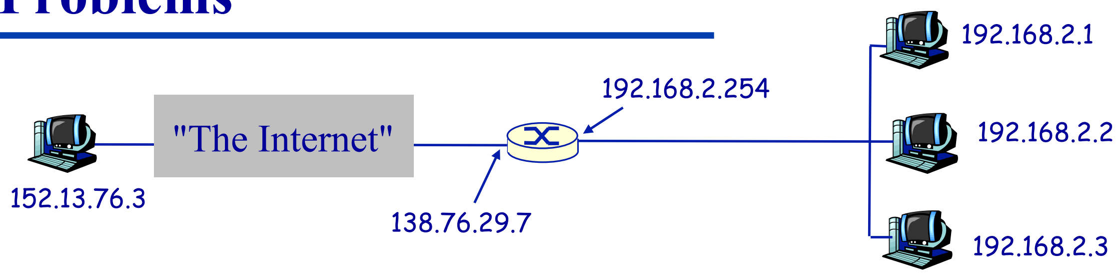


NAT: Network Address Translation

Controversy

- 16-bit port-number field:
  - 60,000 simultaneous connections with a single LAN-side address!
- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, eg, P2P applications
  - address shortage should instead be solved by IPv6

Problems



1) You are using 192.168.2.3 and want to access a web server running on 152.13.76.3.
   a) What data is filled into the NAT table during connection setup?
   b) What information is changed in the SYN and SYN/ACK packets leaving and entering your network?

2) You've setup a web server on 192.168.2.2 and want your friend at 152.13.76.3 to be able to access it.
   a) What should you put in the NAT table at your router?
   b) What address and port should you tell your friend to contact in order to reach your web server?

The Internet Network Layer

The Internet control message protocol ICMP

- Used by hosts, routers, gateways to communicate network-level information
  - » Error reporting: e.g., unreachable {host, network, port, protocol}
  - » Echo request/reply (used by *ping*)
- Provides network-layer functions logically "above" IP
  - » ICMP is encapsulated in IP datagrams
  - » ICMP is assigned a protocol number in the IP header just like TCP and UDP

Type	Code	description
0	0	echo reply (ping)
3	0	network unreachable
3	1	host unreachable
3	2	protocol unreachable
3	3	port unreachable
3	6	network unknown
3	7	host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

- ICMP message: type + code + first 8 bytes of IP datagram triggering the ICMP message

IP Addressing

- **Problem:** We're running out of addresses
- Researchers at USC have mapped out the IP address space
  - » video/slides describing the experiment and graph: http://www.isi.edu/ant/address/video/
  - » interactive graph: http://www.isi.edu/ant/address/browse/

The Network Layer: Routing & Addressing
Outline

- Network layer functions
- Virtual circuits and datagram networks
- Router architecture
- IP Internet Protocol
  - » Addressing
- Routing algorithms
  - » Least cost path computation algorithms
- Hierarchical routing
  - » Connecting networks of networks
- Routing on the Internet
  - » Intra-domain routing
  - » Inter-domain routing