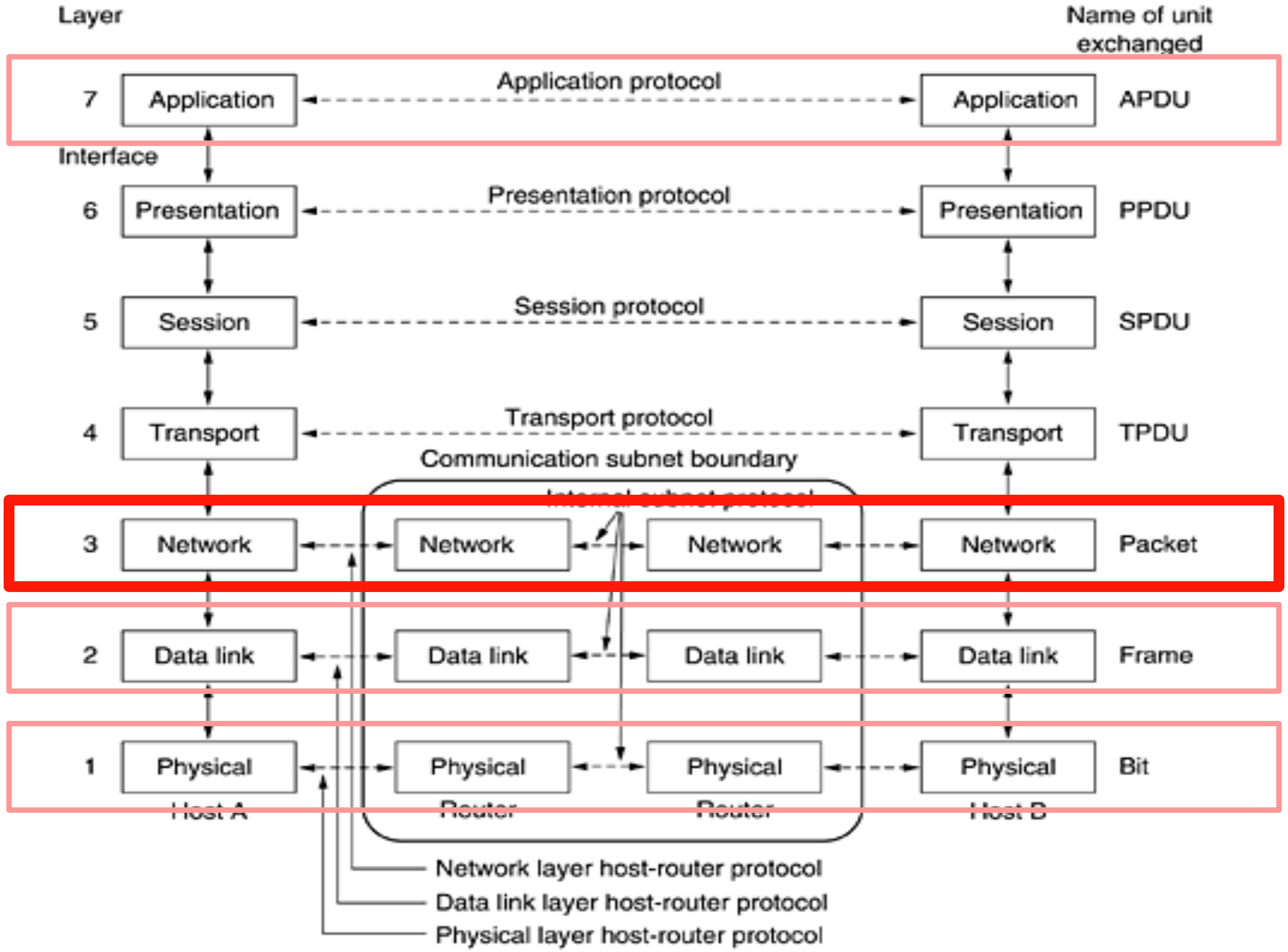
Computer Networks

The Network Layer

Adrian Sergiu DARABANT

Lecture 7

The Network Layer



The Internet Protocol -IP

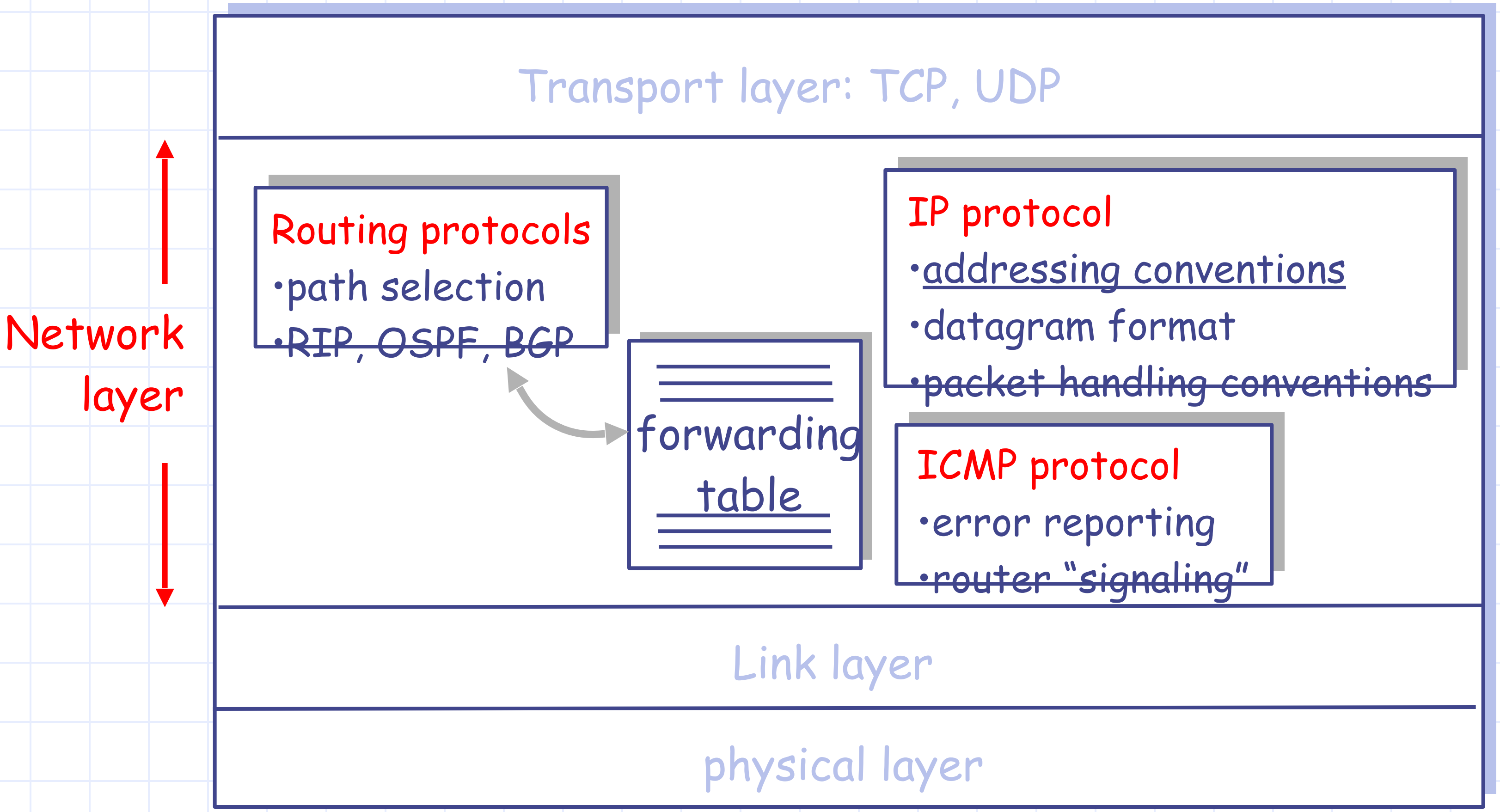
The Internet (IP) Protocol

- IPv4 addressing
- Moving a datagram from source to destination
- Datagram format
- IP fragmentation
- ICMP: Internet Control Message Protocol

- DHCP: Dynamic Host Configuration Protocol
- NAT: Network Address Translation

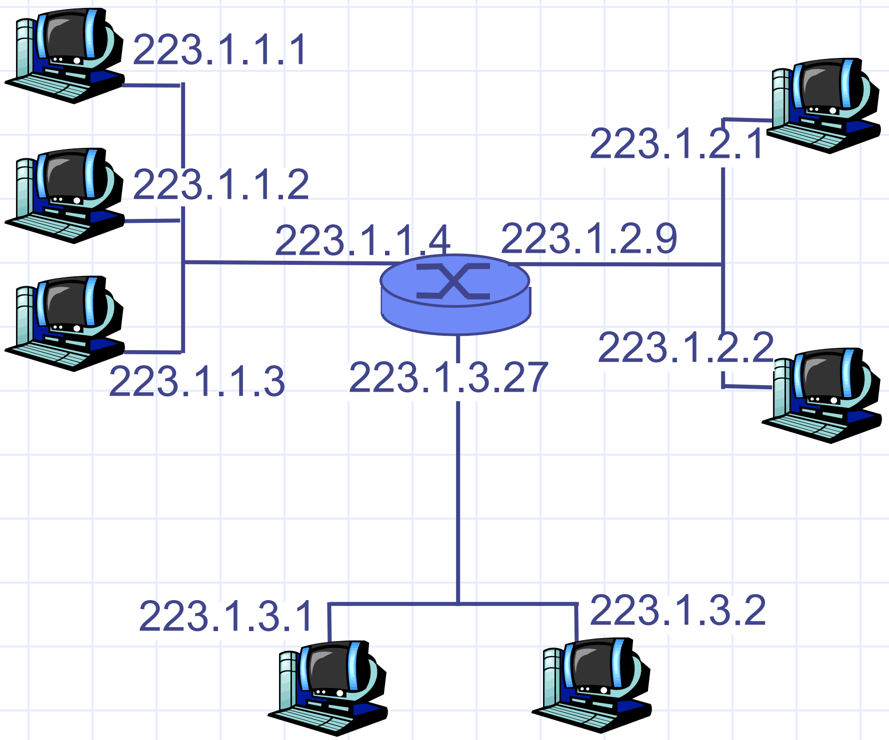
- Routing

The Internet Network Layer



IP Addressing

- IP address: 32-bit identifier for host, router interface
- interface: connection between host/router and physical link
  - router's typically have multiple interfaces
  - host may have multiple interfaces
  - IP addresses associated with each interface

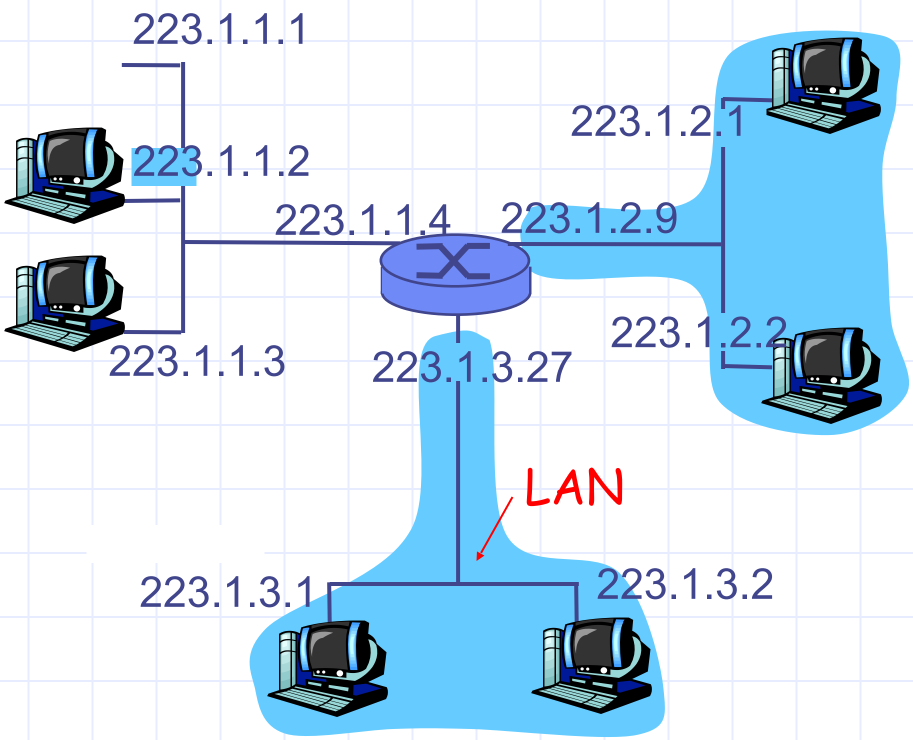


223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

IP Addressing

- IP address:
  - network part (high order bits)
  - host part (low order bits)
- What's a network ? (from IP address perspective)
  - device interfaces with same network part of IP address
  - can physically reach each other without intervening router

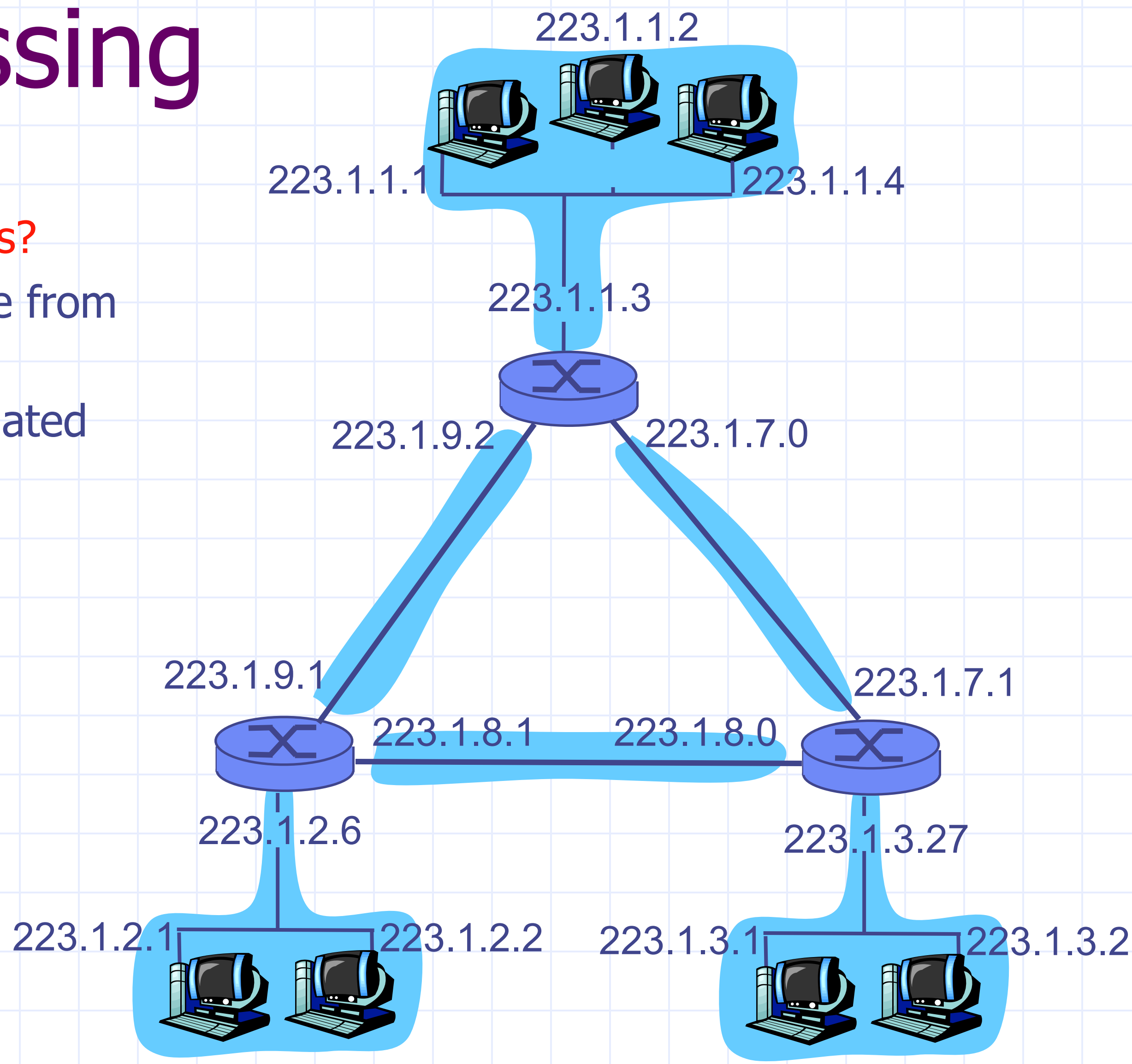


network consisting of 3 IP networks (for IP addresses starting with 223, first 24 bits are network address)

IP Addressing

How to find the networks?

- Detach each interface from router, host
- create "islands of isolated networks



Interconnected system consisting of six networks

IP Addresses

given the notion of "network", let's re-examine IP addresses:

"class-full" addressing:

class		
A	0 network | host	1.0.0.0 to 127.255.255.255
B	10 network | host	128.0.0.0 to 191.255.255.255
C	110 network | host	192.0.0.0 to 223.255.255.255
D	1110 multicast address	224.0.0.0 to 239.255.255.255

← 32 bits →

IP Addressing: CIDR

Classful addressing:

- inefficient use of address space, address space exhaustion
- e.g., class B net allocates enough addresses for 65K hosts, even if we only have 2K hosts in that network

CIDR: Classless InterDomain Routing

- network portion of address of arbitrary length
- address format: a.b.c.d/x, where x is # bits in network portion of address

network part ← → host part

11001000 00010111 00010000 00000000

200.23.16.0/23

IP Subnet

- Basic concept:
  - A subset of a class A, B or C network.
- IP addresses that do not use subnets consists of
  - A <u>network portion</u>, and
  - A <u>host portion</u>.
- Represents a static two-level hierarchical addressing model.

IP Subnet (cont)

- IP subnets introduces a third level of hierarchy.
  - A <u>network</u> portion
  - A <u>subnet</u> portion
  - A <u>host</u> portion

usually handled together as network but with substructure

- Allow more efficient (and structured) utilization of the addresses.
- Uses network masks.

CIDR – Introduction

- The size of the global routing tables have grown very fast in recent years.
  - Caused routers to become saturated.
- CIDR is a new concept to manage IP networks.
  - Classless Inter Domain Routing.
  - No concept of class A, B, C networks.
  - Reduces sizes of routing tables.

CIDR - Basic Idea

- An IP address is represented by a prefix, which is the IP address of the network.
- It is followed by a slash, followed by a number **M**.
  - **M**: number of leftmost contiguous bits to be used for the network mask.
  - Example: 144.16.192.57 / 18

CIDR - Rules

- The number of addresses in each block must be a power of 2.

- The beginning address in each block must be divisible by the number of addresses in the block.
  - A block that contains 16 addresses cannot have beginning address as 193.226.40.36.
  - But the address 193.226.40.64 is possible !

IP/Netmask - examples

209.220.186.8/255.255.255.252=>

209.220.186.8
209.220.186.9
209.220.186.10
209.220.186.11

209.220.186.8/255.255.255.248=>

209.220.186.8
209.220.186.9
209.220.186.10
209.220.186.11
209.220.186.12
209.220.186.13
209.220.186.14
209.220.186.15

Invalid combination: 209.220.186.8/255.255.255.240

Network masks

- Network mask 255.0.0.0 is applied to a class A network 10.0.0.0;
  - Mask = series of contiguous 1's followed by a series of contiguous 0's

11111111 00000000 00000000 00000000

NETWORK

HOST

Natural Masks

- Provide a mechanism to split the IP address 10.0.0.20 into:
  - A network portion – **10**;
  - A host portion – **20**;

IP Address:	10.0.0.20	00001010 00000000 00000000 00010100
Mask:	255.0.0.0	11111111 00000000 00000000 00000000
		Network Host

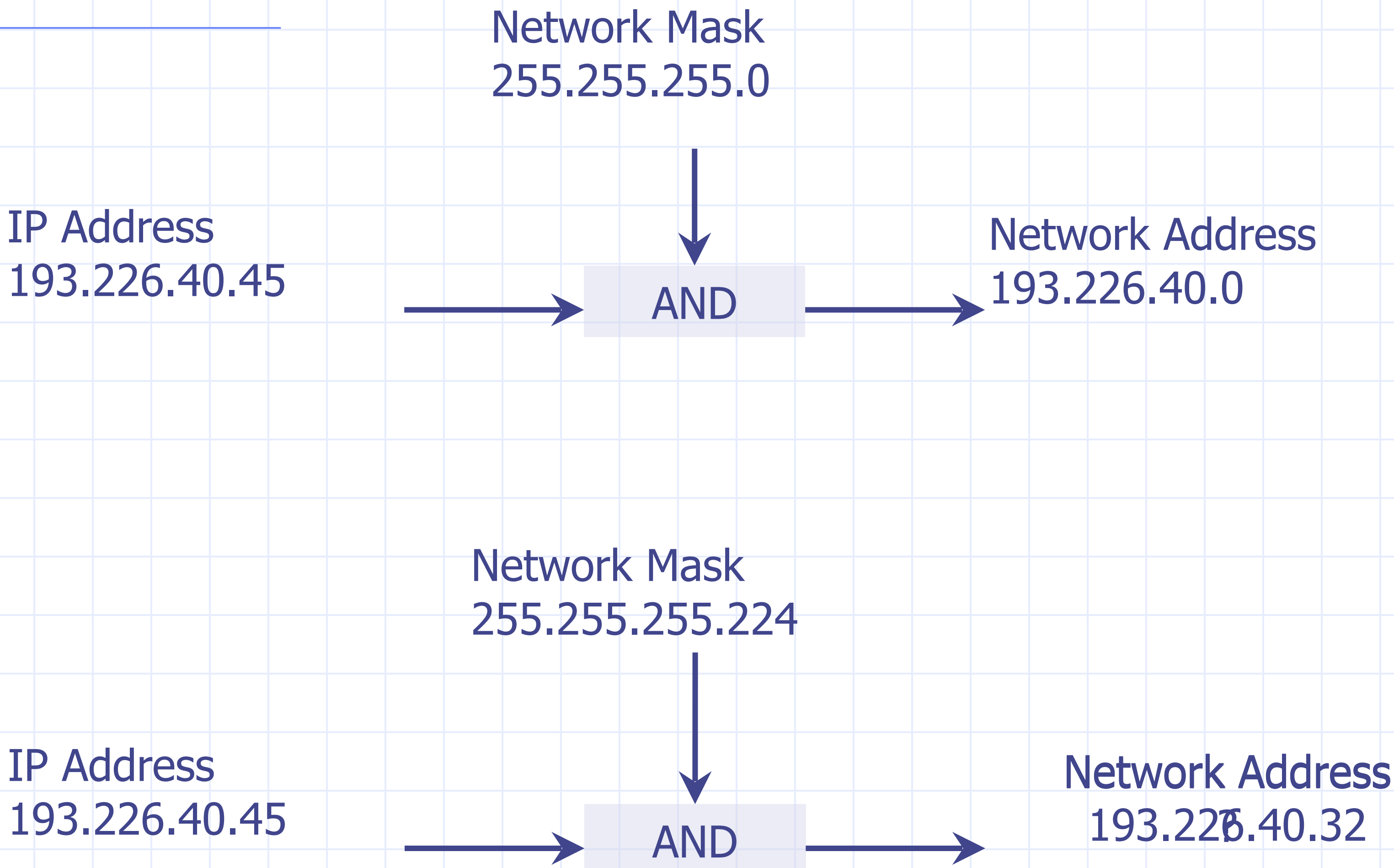
Natural masks

- Class A, B and C addresses
  - Have fixed division of network and host portions
  - Can be expressed as masks
- Natural Masks
  - Class A: 255.0.0.0
  - Class B: 255.255.0.0
  - Class C: 255.255.255.0

Subnets out of masks

- Masks are very flexible.
  - Using masks, networks can be divided into smaller subnets.
- How?
  - By extending the network portion of the address into the host portion.
- Advantage gained:
  - We can create a large number of subnets from one network.
  - Can have less number of hosts per network.

Network Address



How does one get IP Addresses ?

- Q: How does a network get the network part of IP addr?
- A: it gets allocated from the 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

Reserved Addresses

CIDR address block	Description	Reference
0.0.0.0/8	Current network (only valid as source address)	RFC 1700
10.0.0.0/8	Private network	RFC 1918
14.0.0.0/8	Public data networks (per 2008-02-10, available for use[1])	RFC 1700
127.0.0.0/8	Loopback	RFC 3330
128.0.0.0/16	Reserved (IANA)	RFC 3330
169.254.0.0/16	Link-Local	RFC 3927
172.16.0.0/12	Private network	RFC 1918
191.255.0.0/16	Reserved (IANA)	RFC 3330
192.0.0.0/24	Reserved (IANA)	RFC 3330
192.0.2.0/24	Documentation and example code	RFC 3330
192.88.99.0/24	IPv6 to IPv4 relay	RFC 3068
192.168.0.0/16	Private network	RFC 1918
198.18.0.0/15	Network benchmark tests	RFC 2544
223.255.255.0/24	Reserved (IANA)	RFC 3330
224.0.0.0/4	Multicasts (former Class D network)	RFC 3171
240.0.0.0/4	Reserved (former Class E network)	RFC 1700
255.255.255.255	Broadcast	

Private Addreses

Name	IP address range	number of IPs	*classful* description	largest CIDR block	defined in
24-bit block	10.0.0.0 – 10.255.255.255	16,777,216	single class A	10.0.0.0/8	RFC 1597 (obsolete), RFC 1918
20-bit block	172.16.0.0 – 172.31.255.255	1,048,576	16 contiguous class Bs	172.16.0.0/12	
16-bit block	192.168.0.0 – 192.168.255.255	65,536	256 contiguous class Cs	192.168.0.0/16	

Not routed in Internet

Why ?

Routing tables (static)

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
172.16.25.1	172.30.0.4	255.255.255.255	UGH	0	0	0	Eth1
193.226.40.128	0.0.0.0	255.255.255.224	U	0	0		Eth0
193.0.225.0	0.0.0.0	255.255.255.0	U	0	0		Eth0
193.231.20.0	0.0.0.0	255.255.255.0	U	0	0		Eth0
172.30.0.0	0.0.0.0	255.255.0.0	U	0	0		Eth1
169.254.0.0	0.0.0.0	255.255.0.0	U	0	0		Eth1
0.0.0.0	193.0.225.9	0.0.0.0	UG	0	0		Eth0

The route command – (Windows/Linux/other OS)

IP Datagram

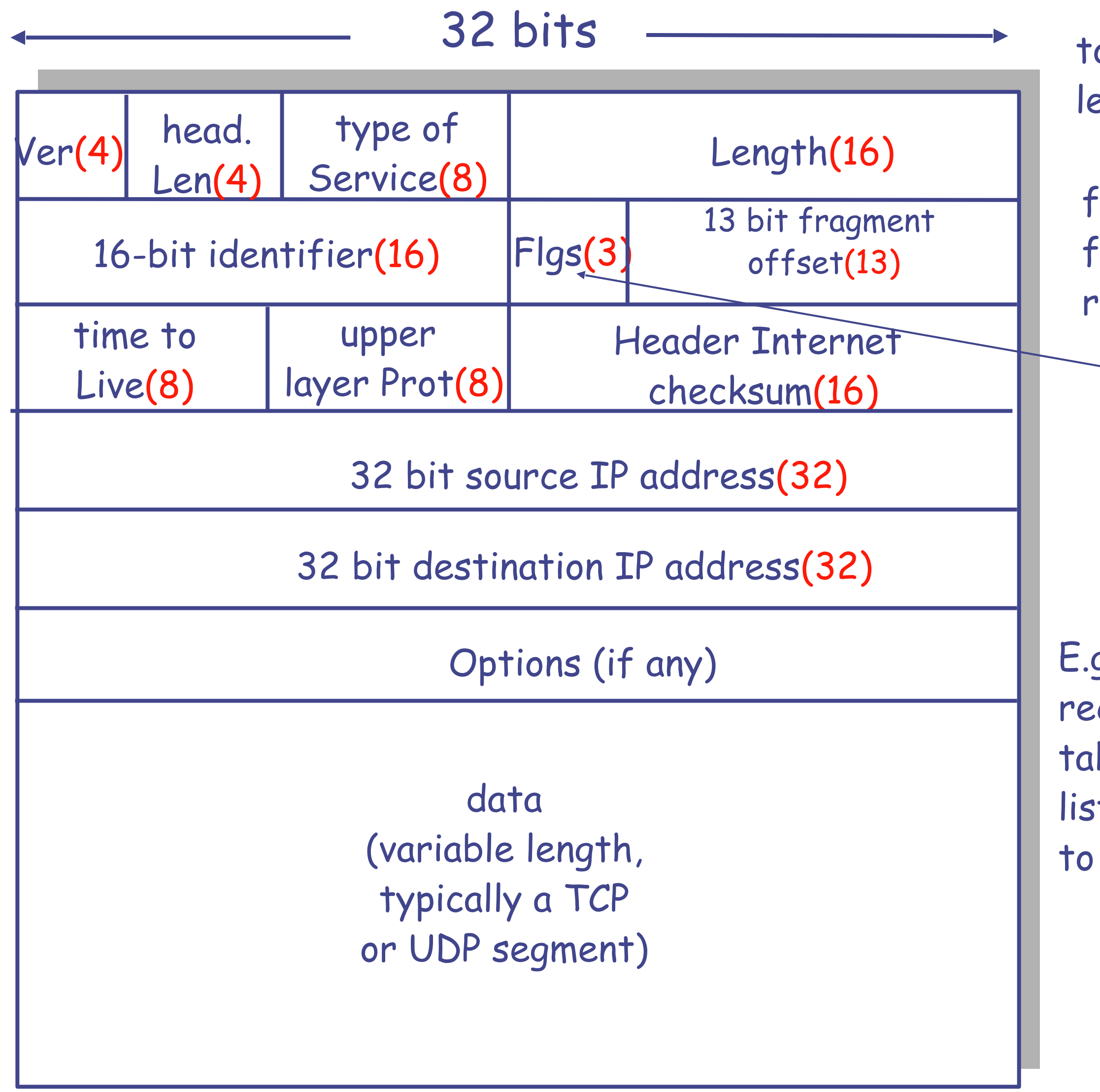
IP protocol version number

header length (bytes)

"type" of data

max number remaining hops (decremented at each router)

upper layer protocol to deliver payload to



total datagram length (bytes)

for fragmentation/ reassembly

E.g. timestamp, record route taken, specify list of routers to visit.

<u>how much overhead with TCP?</u>

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

Datagram: from source to destination

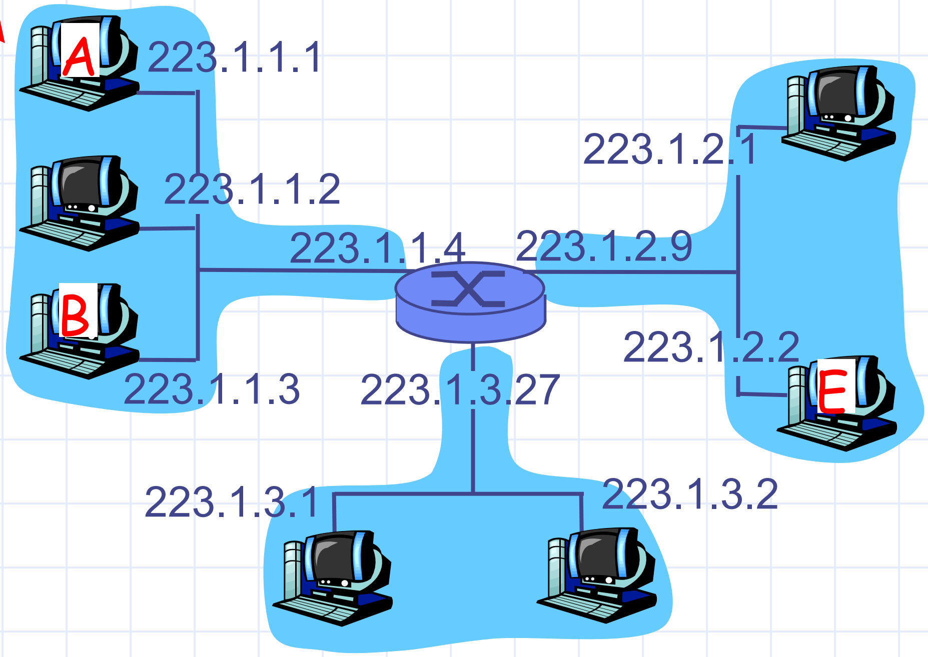
forwarding table in A

Dest Net	Mask	Nxt Router	Metric
223.1.1.0	255.255.255.0		1
223.1.2.0	255.255.255.0	223.1.1.4	2
223.1.3.0	255.255.255.0	223.1.1.4	2
64.8.32.1	255.255.255.255	223.1.1.10	2

IP datagram:

misc fields	source IP addr	dest IP addr	data

- datagram remains unchanged, as it travels source to destination
- Addresses are fields of interest here



Datagram: from source to destination

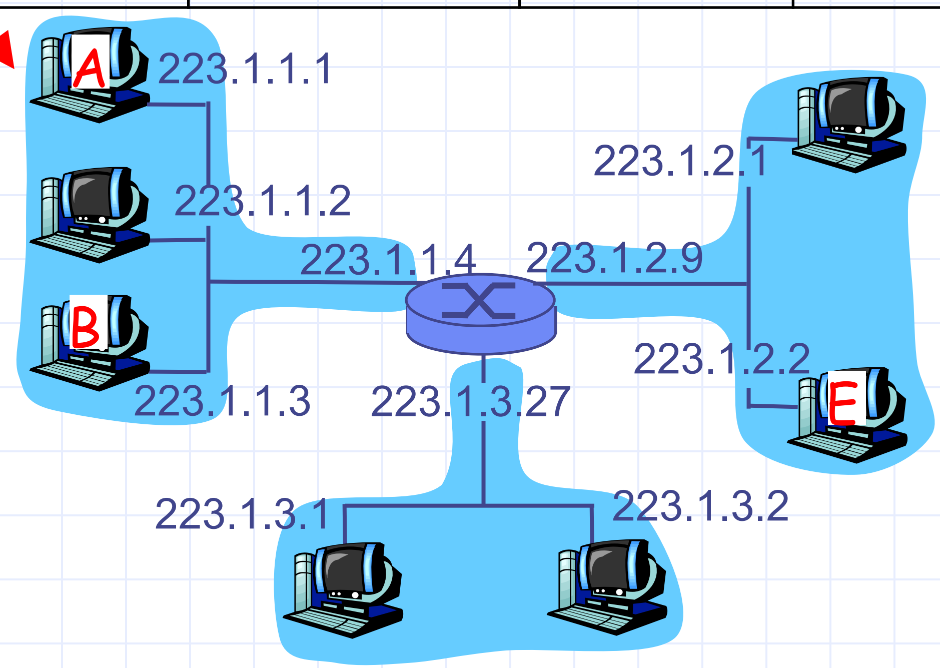
misc fields	223.1.1.1	223.1.1.3	data

forwarding table in A

Dest Net	Mask	Nxt Router	Metric
223.1.1.0	255.255.255.0		1
223.1.2.0	255.255.255.0	223.1.1.4	2
223.1.3.0	255.255.255.0	223.1.1.4	2
64.8.32.1	255.255.255.255	223.1.1.10	2

Starting at A, send IP datagram addressed to B:

- look up net. address of B in forwarding table
- find B is on same net. as A
- link layer will send datagram directly to B inside link-layer frame
  - B and A are directly connected



Datagram: from source to destination

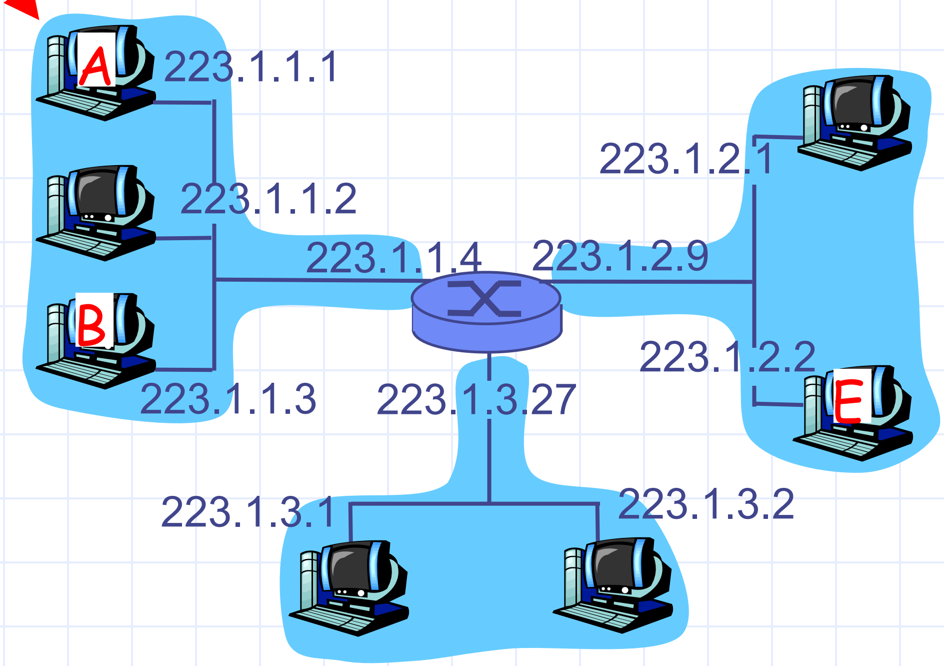
misc fields	223.1.1.1	223.1.2.3	data

forwarding table in A

Dest Net	Mask	Nxt Router	Metric
223.1.1.0	255.255.255.0		1
223.1.2.0	255.255.255.0	223.1.1.4	2
223.1.3.0	255.255.255.0	223.1.1.4	2
64.8.32.1	255.255.255.255	223.1.1.10	2

Starting at A, dest. E:

- look up network address of E in forwarding table
- E on different network
  - A, E not directly attached
- routing table: next hop router to E is 223.1.1.4
- link layer sends datagram to router 223.1.1.4 inside link-layer frame
- datagram arrives at 223.1.1.4
- continued…..



Datagram: from source to destination

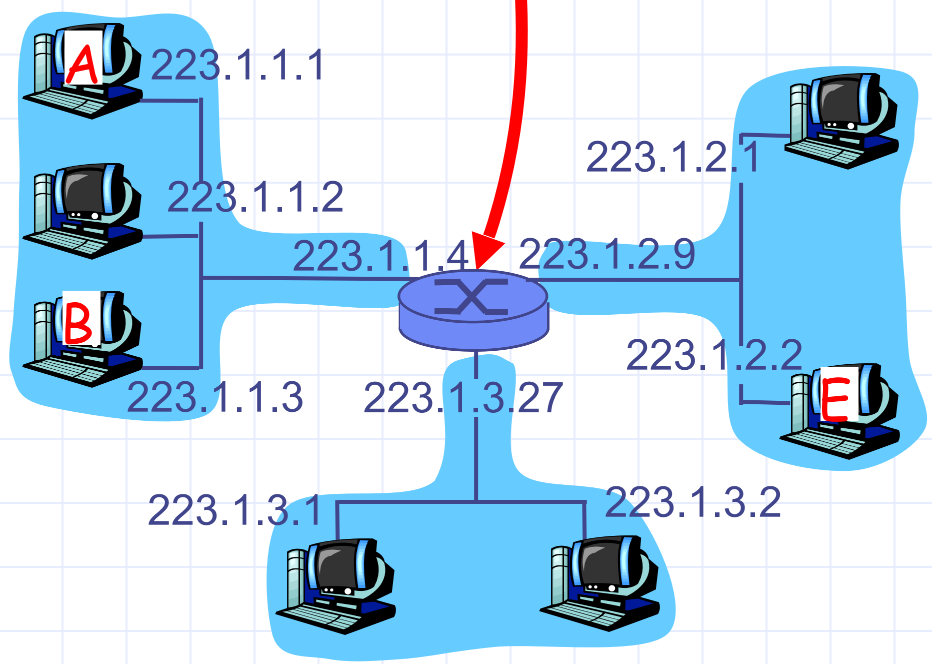
misc fields	223.1.1.1	223.1.2.3	data

forwarding table in router

Dest Net	Mask	Nxt R	Metric	Interface
223.1.1.0	255.255.255.0	-	1	223.1.1.4
223.1.2.0	255.255.255.0	-	1	223.1.2.9
223.1.3.0	255.255.255.0	-	1	223.1.3.27

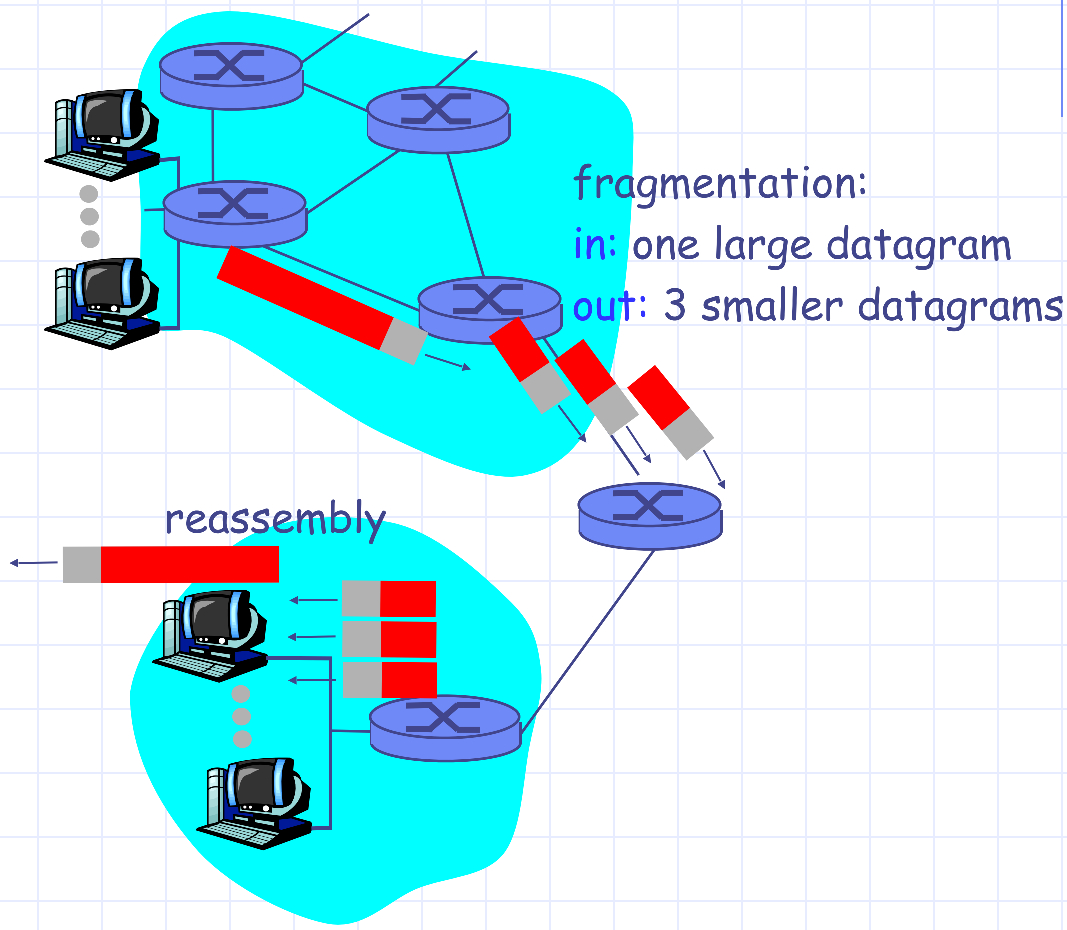
Arriving at 223.1.4, destined for 223.1.2.2

- look up network address of E in router's forwarding table
- E on same network as router's interface 223.1.2.9
  - router, E directly attached
- link layer sends datagram to 223.1.2.2 inside link-layer frame via interface 223.1.2.9
- datagram arrives at 223.1.2.2!!! (hooray!)



Fragmentation/Reassembly

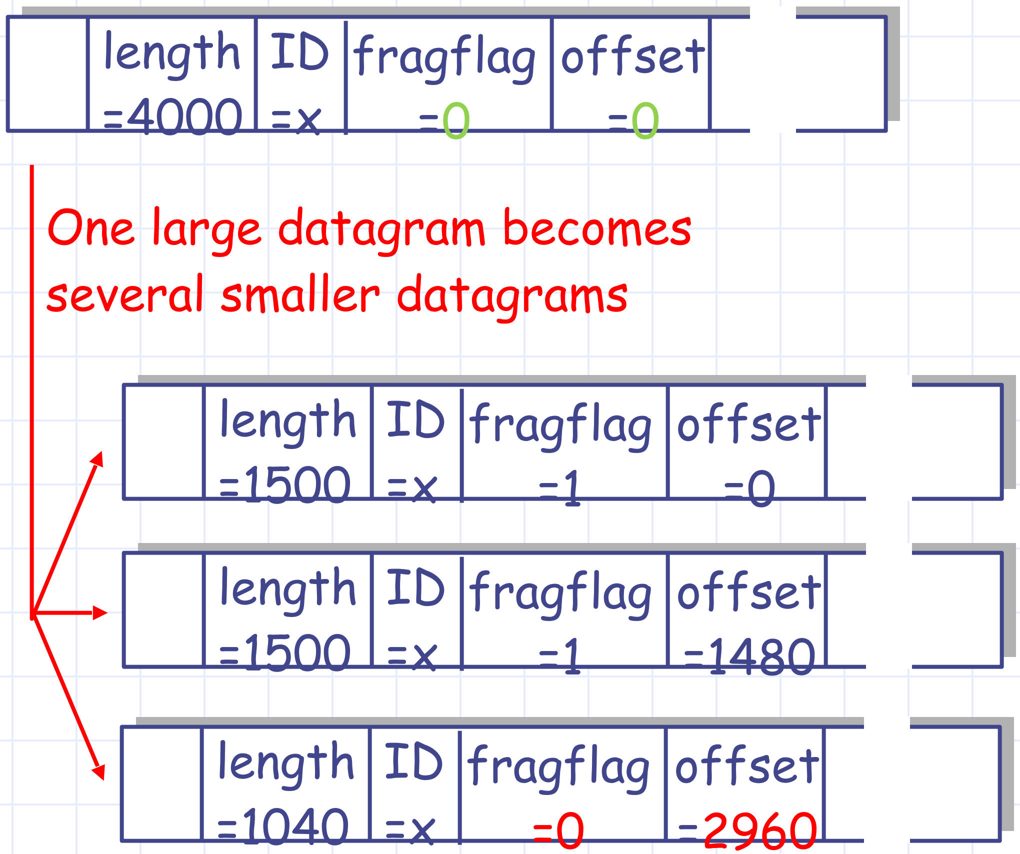
- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



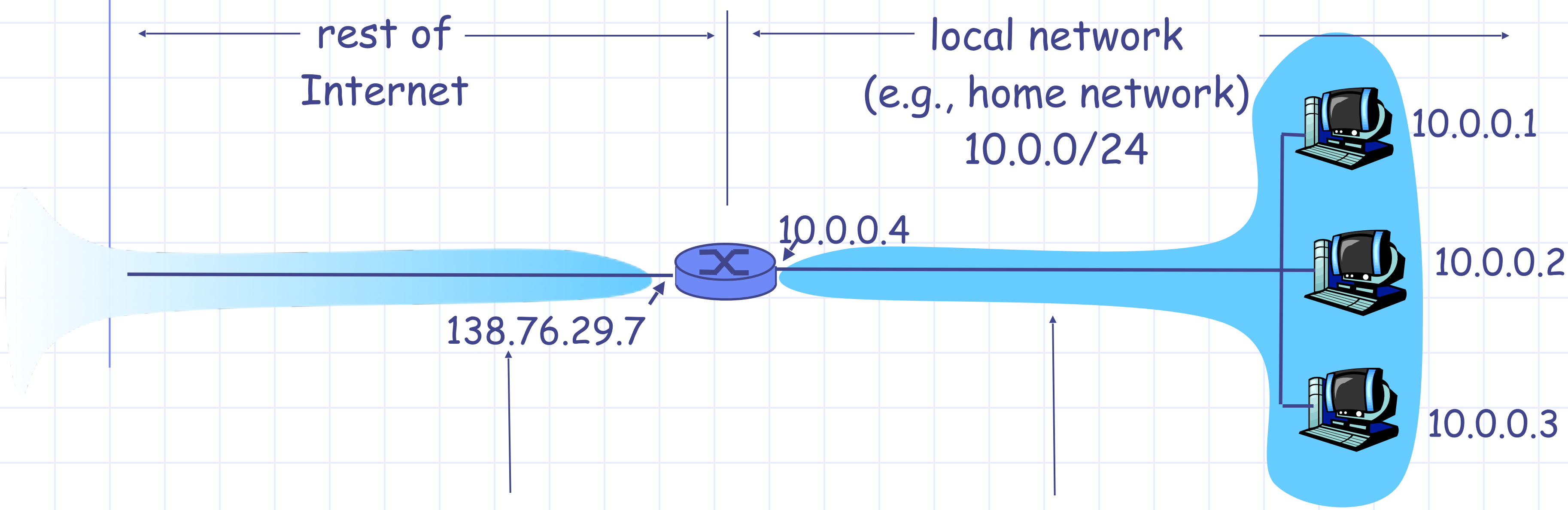
Fragmentation/Reassembly

Example

- 4000 byte datagram
- MTU = 1500 bytes



NAT – Network Address Translation



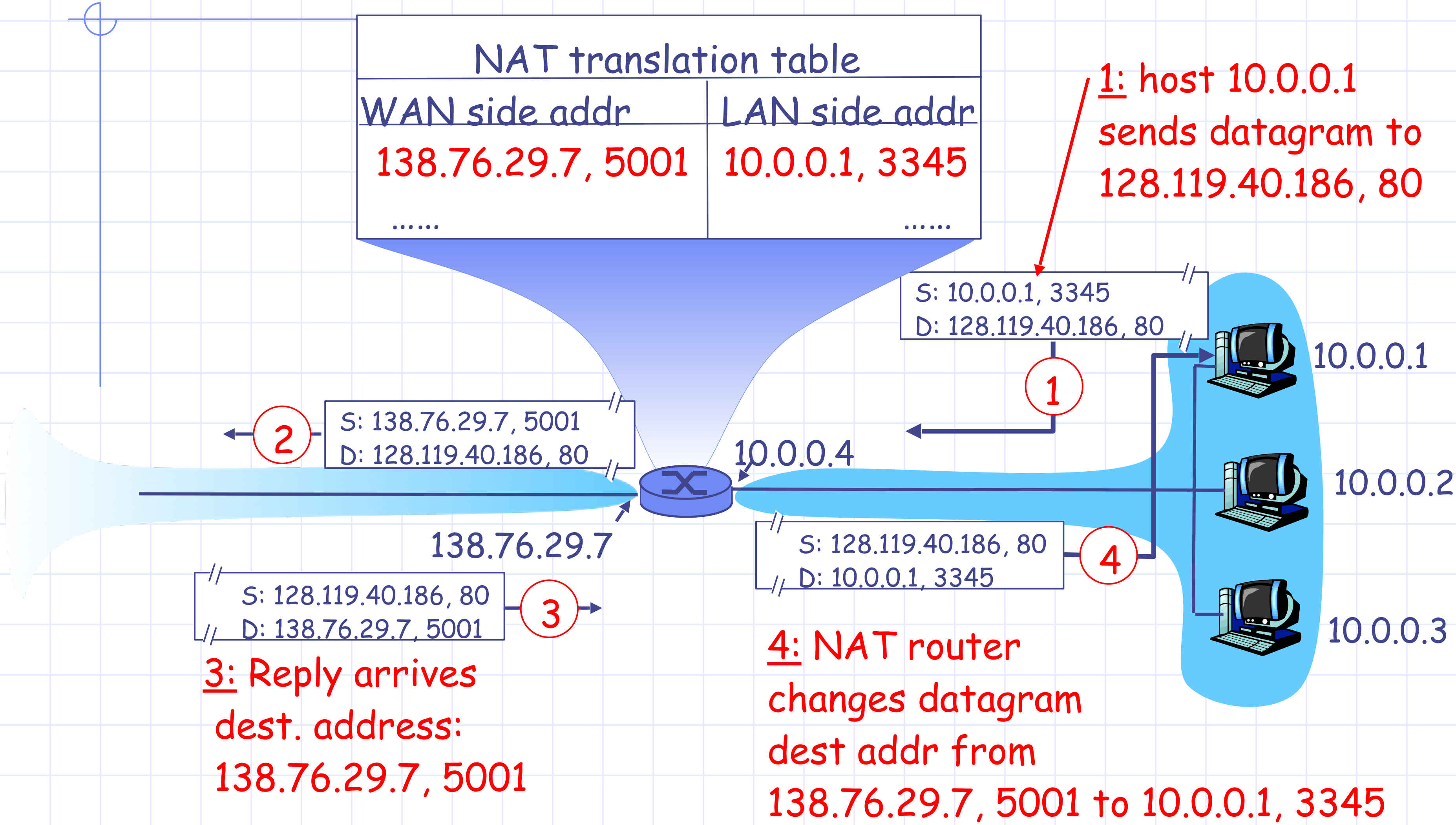
All datagrams leaving local network have same single source NAT IP address: 138.76.29.7, different source port numbers

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 word is concerned:
  - no need to be allocated range of addresses from ISP: - just one IP address is used 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).

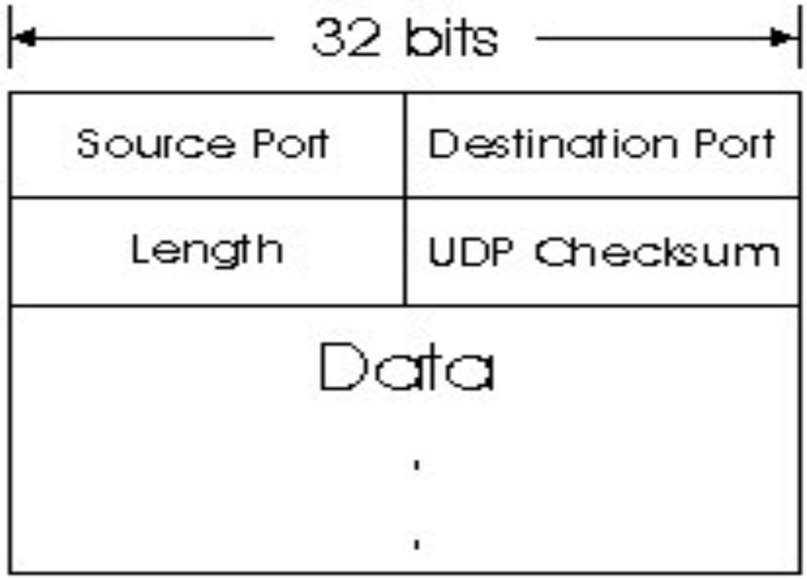
NAT – Network Address Translation



NAT – Network Address Translation

- 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, e.g., P2P applications
  - address shortage should instead be solved by IPv6

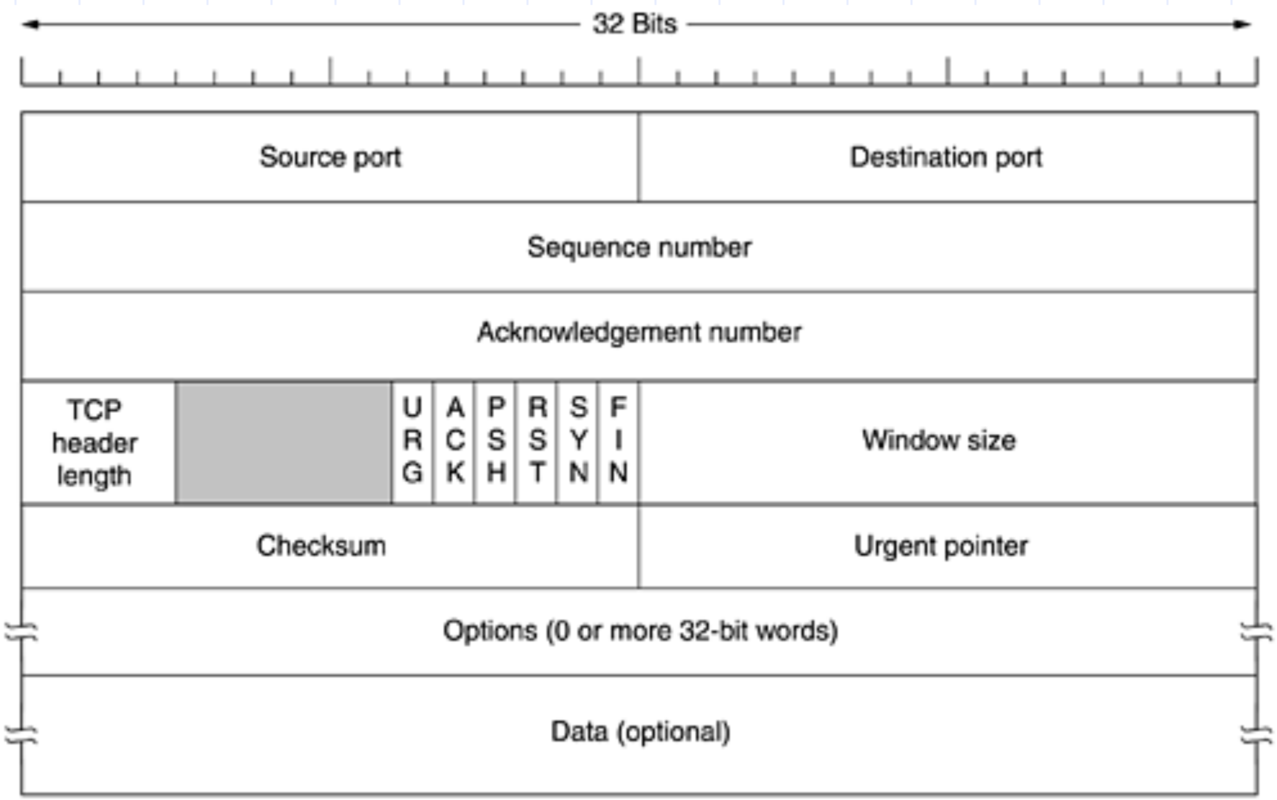
UDP



Checksum – for the entire datagram (header + data)

Length >=8 – entire datagram

TCP Datagrams

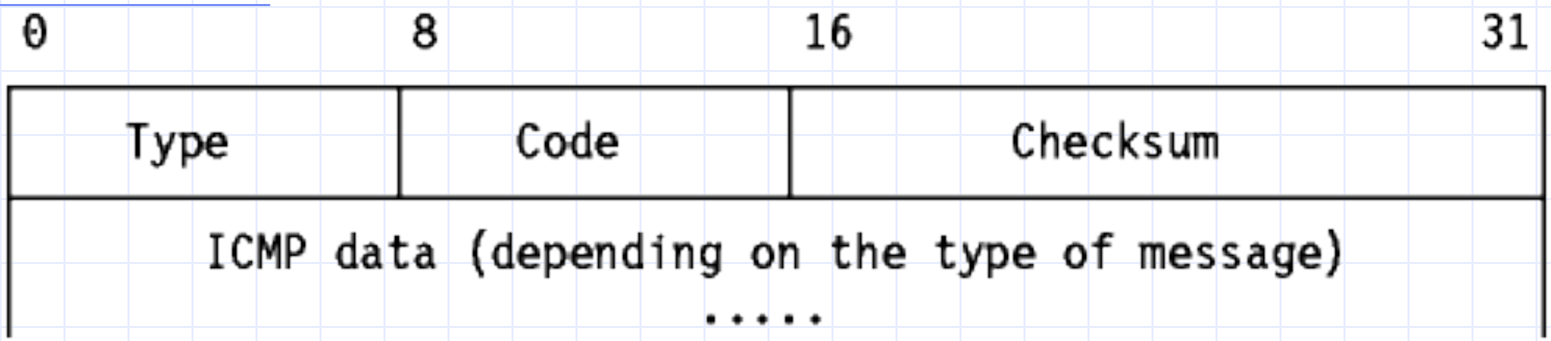


Sequence No – ACK No

ICMP

- Used by hosts, routers, gateways to communication network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- Network-layer "above" IP:
  - ICMP msgs carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing error

ICMP



Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown

Type	Code	description
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