Computer Networks

Wenzhong Li, Chen Tian

Nanjing University

Material with thanks to James F. Kurose, Mosharaf Chowdhury, and other colleagues.

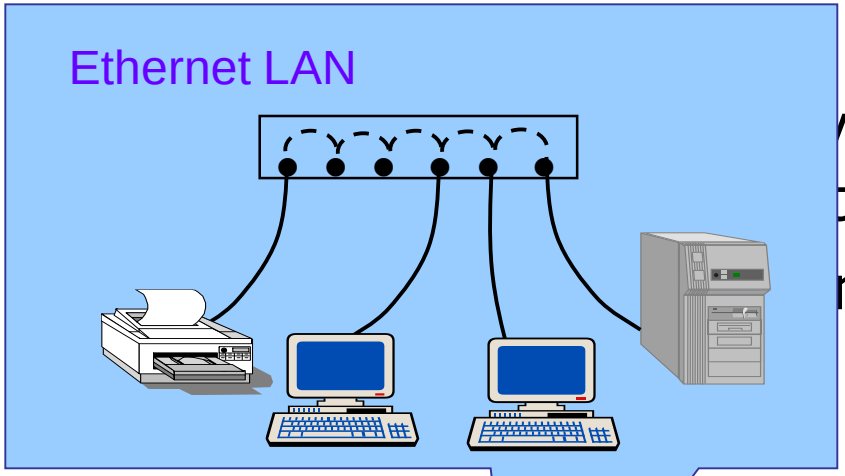
Chapter 3. Network Layer

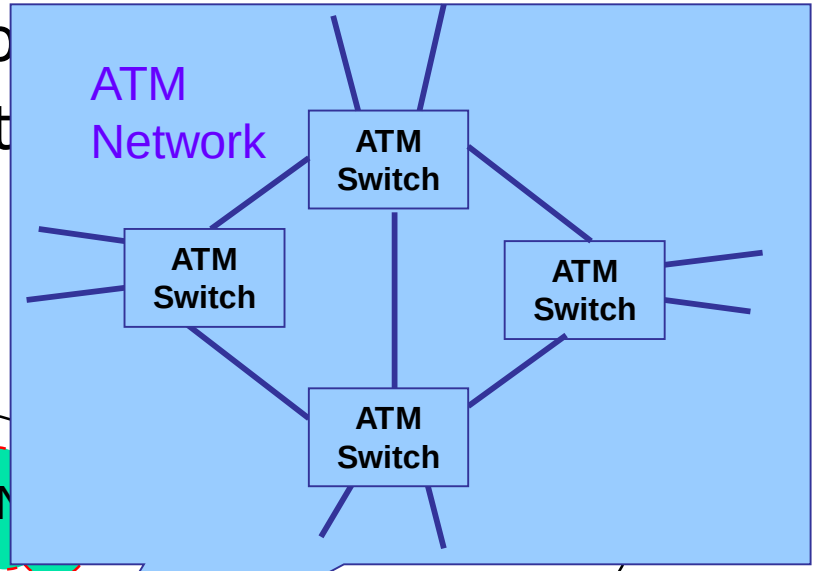
- Network Layer Functions
- IP Protocol Basic
- IP Protocol Suit
- Routing Fundamentals
- Internet Routing Protocols
- IP Multicasting

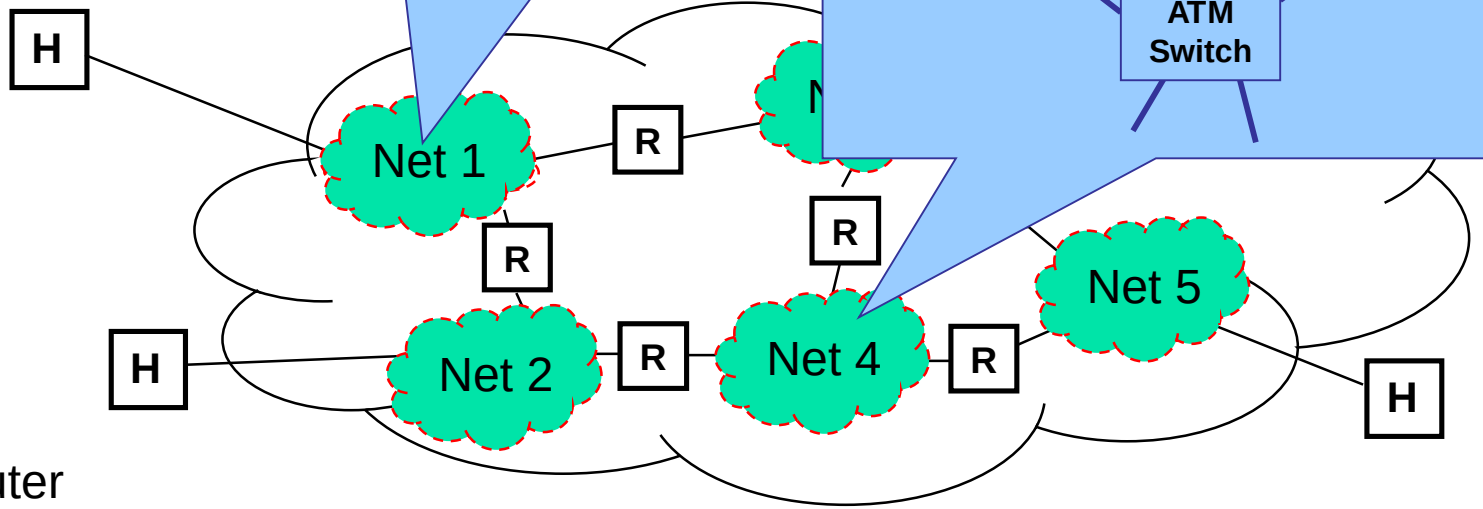
- The Internet Protocol
- IP Operations
- IP Packet Structure
- IP Fragmentation
- IP Address

Internetworking

- ides transfer of packets across
- two
- net



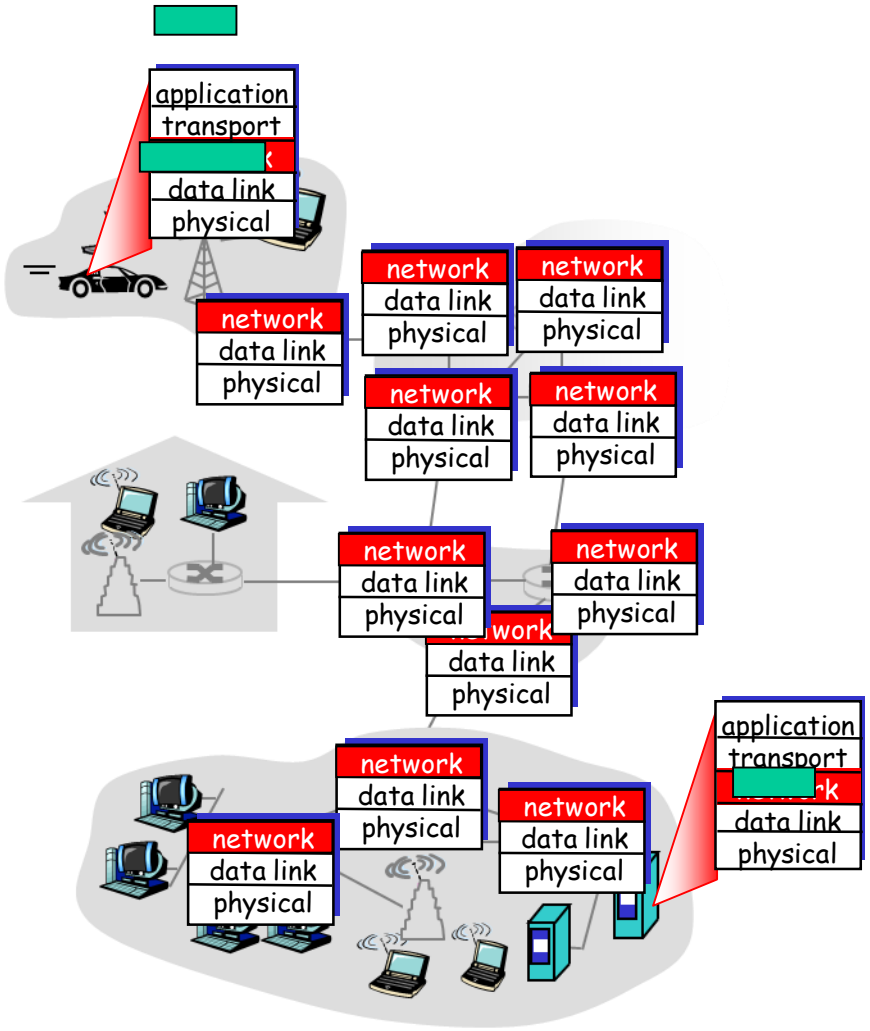




R = router
H = host

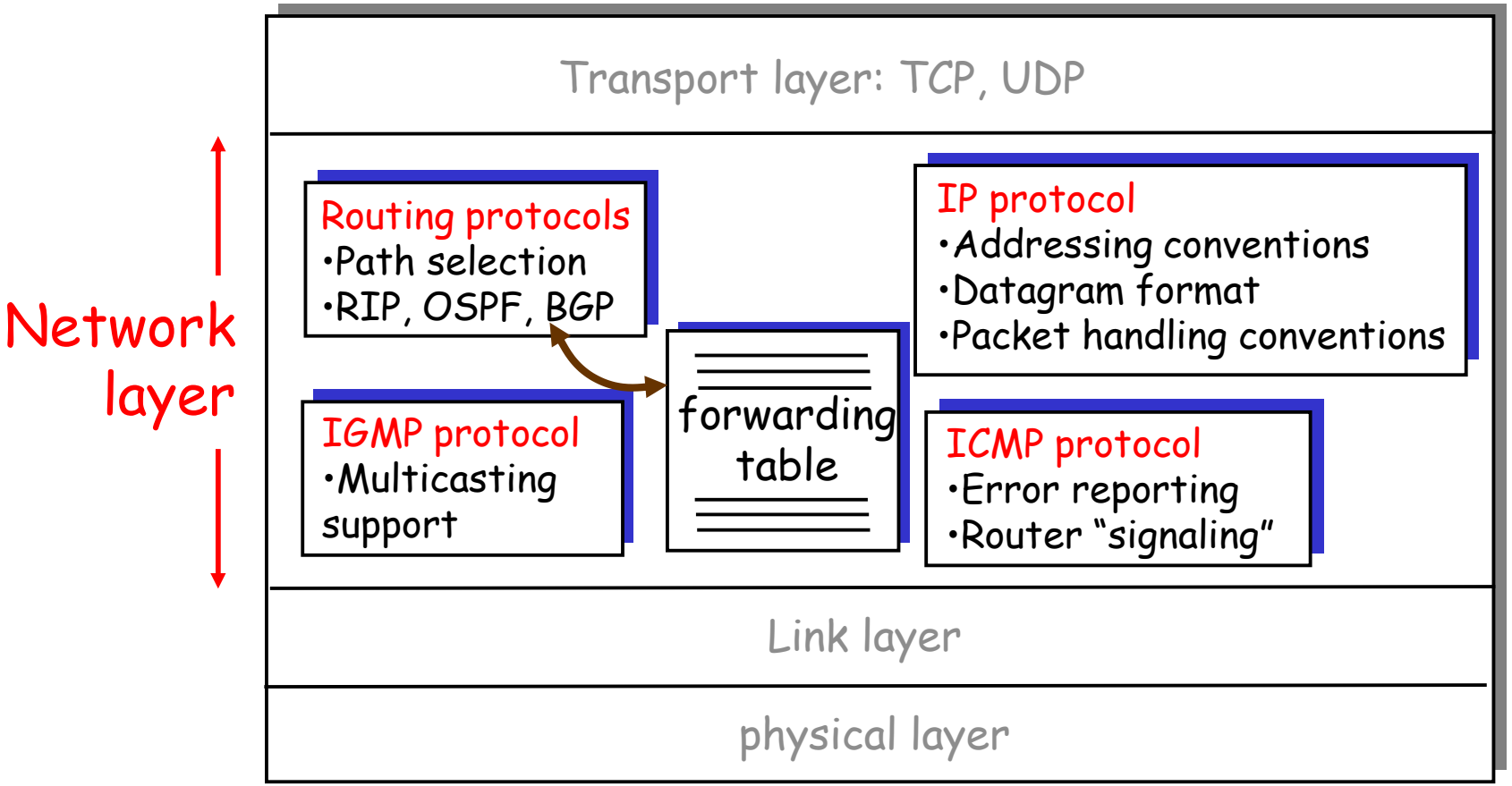
Positions of the IP Protocol

- IP – Internet Protocol
  - Most famous internet protocol developed for ARPANET
  - RFC 791, Internet STD number 5
- IP layer entity resides on each host and router
- Provides connectionless service (i.e. datagram mechanism)



The Internet Network layer

- Host, router network layer functions



Internet Addressing

- Addressing level
- Addressing scope
- Addressing mode

Addressing Level

- Physical network address
  - Used to route PDU within single physical network
- Inter-network address
  - IP address or internet address, used to route PDU across networks
  - Unique address for each end system (host) and each intermediate system (router)
- Application address
  - Process identifier assigned at destination host
  - i.e. TCP/IP port

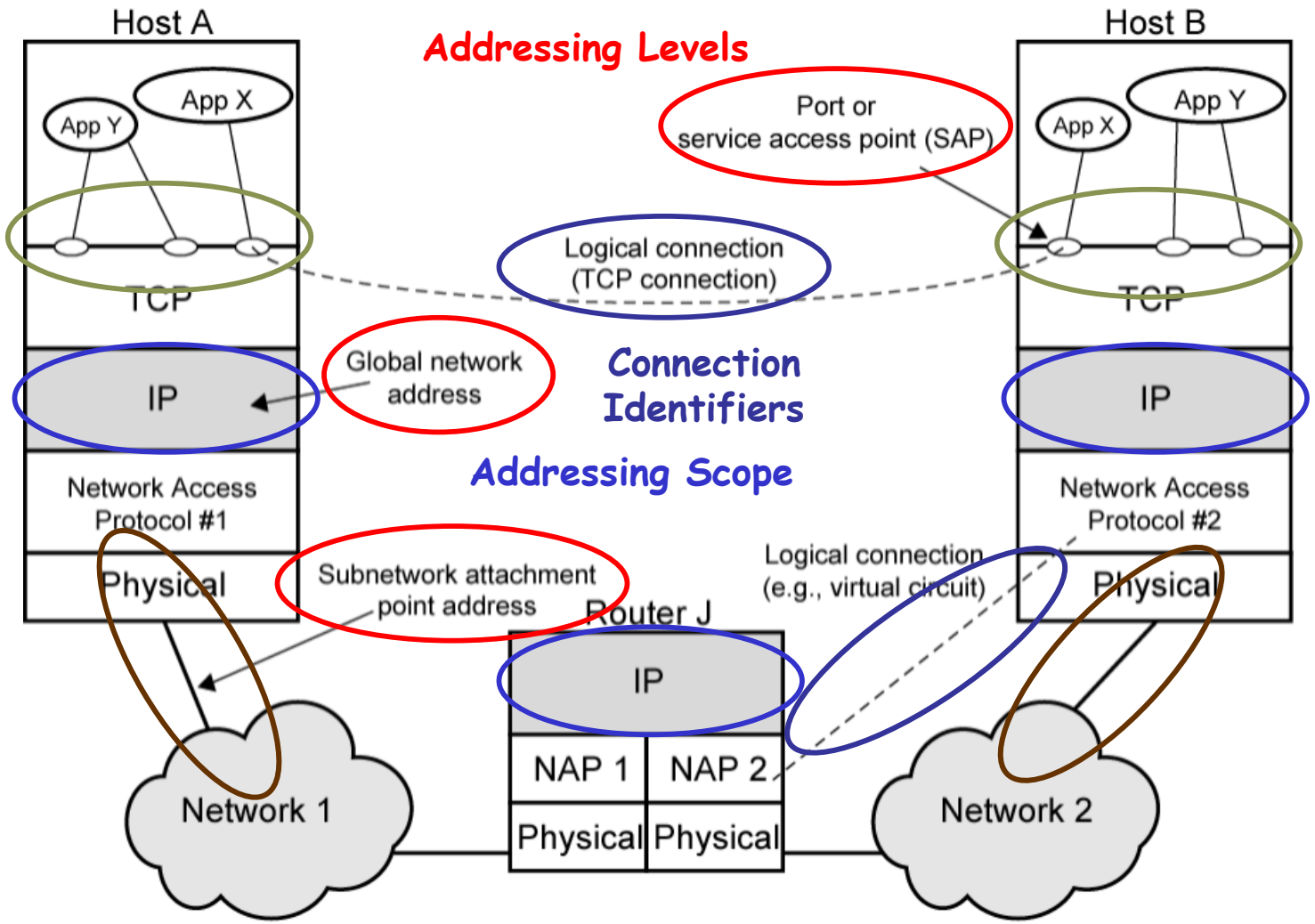
Addressing Scope

- Global address
  - Identifies host or router with global non-ambiguity
  - Synonyms permitted, i.e. a router may have more than one global address
- Network attachment address
  - Unique address for each device interface on specific network
  - e.g. *MAC* address on IEEE 802 network or *ATM* host address
- Port address
  - Above network level and unique within a system (router or host)
  - e.g. port 80 – web server listening port on TCP/IP
  - Need not be unique outside the single system

Addressing Mode

- Individual or Unicast address
  - Address referring to a single system or port
- Broadcast address
  - For all entities within a domain
- Multicast address
  - For specific subset of entities
- Anycast address
  - Any (suitable) entity within a subset

Level of Addresses



IP Operations

IP Operations

- Routing
- Datagram lifetime
- Fragmentation and re-assembly
- Error control
- Flow control

Routing

- Hosts and routers maintain routing tables
  - Indicate next router to which datagram should be sent
  - Static – may contain alternative routes
  - Dynamic – flexible response to congestion and errors
- Routing policy
  - Distance vector, Link state, Path vector
- Source routing
  - Source specifies route as sequential list of routers to be followed
- Route recording

Datagram Lifetime

- Datagrams may loop indefinitely
  - Routing based on obsolete networks information
  - TCP needs upper bound on datagram life
- Datagram marked with lifetime
  - Time To Live (TTL) field in IP
  - Once lifetime expires, datagram is discarded instead of forwarded
- Types of lifetime
  - Hop count – Decrement TTL on passing through each router

Fragmentation and Re-assembly

- Length of a packet exceeds the coming network's MTU (maximum transmission unit)
- When to fragment
  - Host – determine min of *MTUs* along the path
  - Router – fragment if the next *MTU* is exceeded
- When to re-assemble
  - Host – Packets getting smaller as data traverses internet
  - Router – infeasible since fragments may take different routes

Dealing with Failure

- Re-assembly may fail if some fragments get lost
- Re-assembly time out
  - Assigned when first fragment arrived
  - If timeout expires before all fragments arrive, discard partial data
- Use packet lifetime (TTL in IP)
  - Decrement with each fragment
  - If TTL runs out, kill partial data

Error Control

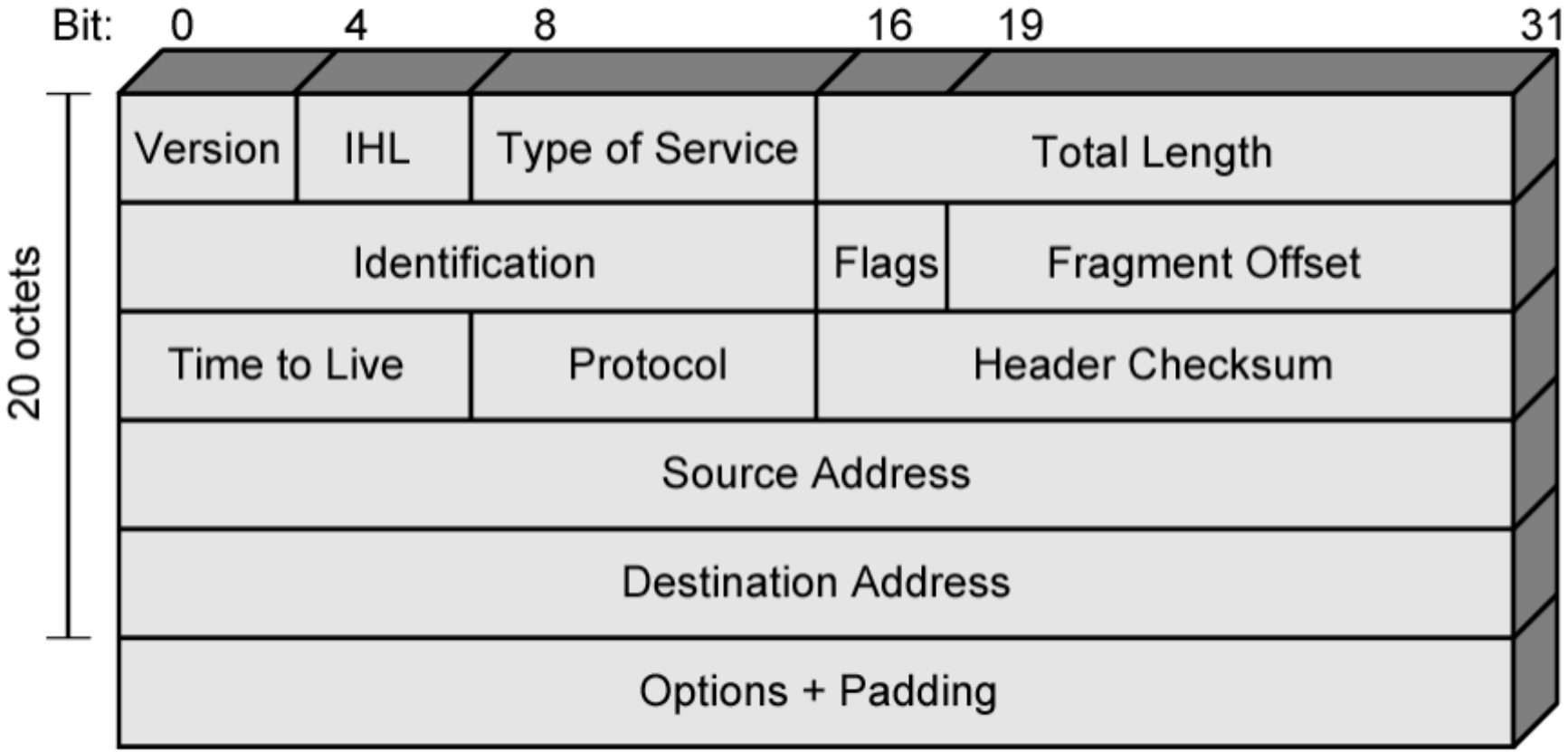
- Not guaranteed delivery
- Router should attempt to inform source if packet discarded
  - e.g. for checksum failure, TTL expiring
  - Datagram identification needed
- ICMP used to send error message
- Source may inform higher layer protocol

Flow Control

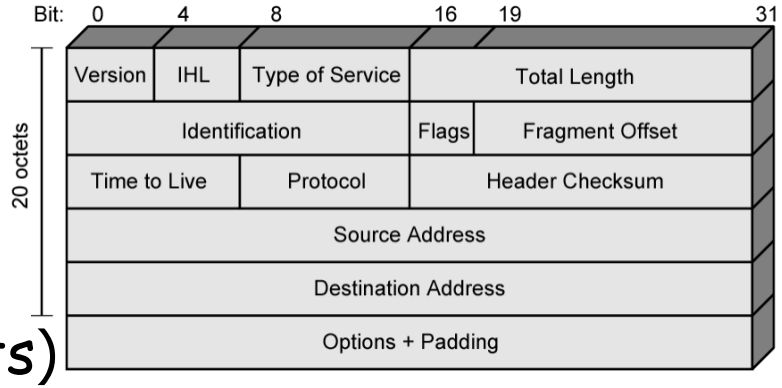
- Allows routers to limit rate of incoming data
  - Limited control functions in connectionless systems
  - New mechanisms coming soon
- Router discards incoming packets when buffer is full
  - May send source quench packets to sending host
  - Using ICMP

IP Packet Structure

IPv4 Header

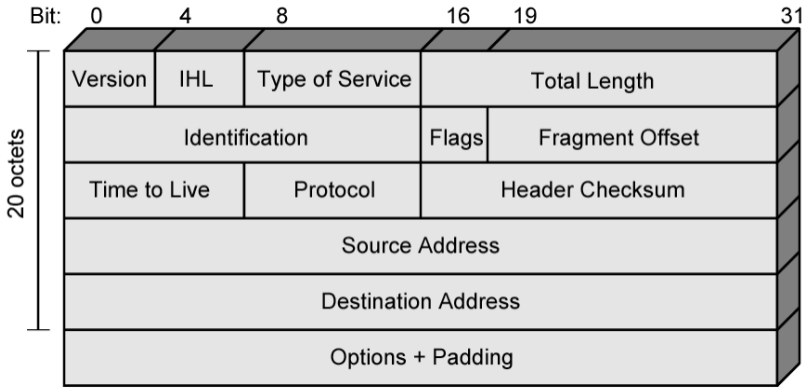

Bit:
0
4
8
16
19
31
Version
IHL
Type of Service
Total Length
Identification
Flags
Fragment Offset
Time to Live
Protocol
Header Checksum
Source Address
Destination Address
Options + Padding
20 octets

Header Fields (1)



- Version (4 bits)
  - Currently 4
  - IPv6 – see later
- Internet header length (IHL) (4 bits)
  - In 32 bit words (4 octets)
  - Minimum fixed header (20 octets) + options
- Type of service (8 bits)
  - Precedence
    - 3 bits, 8 levels defined
  - Reliability
    - 1 bit, Normal or high
  - Delay
    - 1 bit, Normal or low
  - Throughput
    - 1 bit, Normal or high

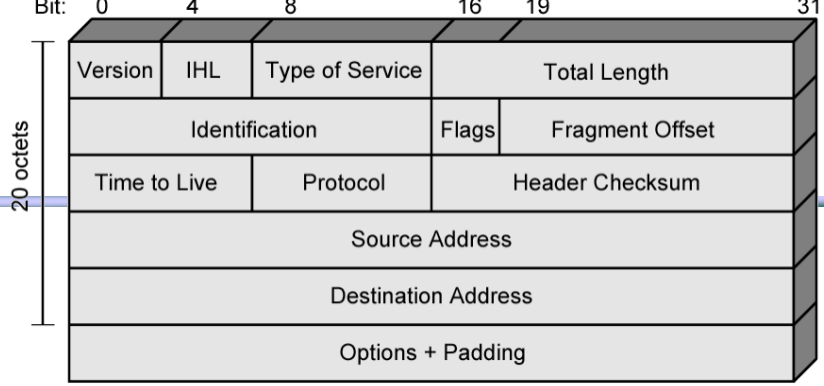
Header Fields (2)



- Total length (16 bits)
  - Of datagram, in octets
- Identification (16 bits)
  - Sequence number
  - Used with addresses and user protocol to identify datagram uniquely
- Flags (3 bits)
  - More flag, Don't fragment
- Fragmentation offset (13 bits)
- Time to live (8 bits)
- Protocol (8 bits)
  - Next higher layer to receive data field at destination

指示传输层的协议类型（TCP或UDP）

Header Fields (3)



- Header checksum (16 bits)
  - Complement sum of all 16 bit words in header
  - If not correct, router discards packets
  - Reverified and recomputed at each router, set to 0 during calculation. (Why?)
- Source address (32 bits)
- Destination address (32 bits)
- Options (variable $\leq$ 40 octets)
- Padding (variable)
  - To fill to multiple of 32 bits long

Data Field

- Carries user data from next layer up
- Multiple of 8 bits long (i.e. octet)

- Max length of datagram (header + data) 65,535 octets

IP Primitives

2 primitives

- Send (called by upper layer)
  - Request transmission of data unit
- Deliver (notify upper layer)
  - Notify user of arrival of data unit
- Parameters
  - Used to pass data and control info

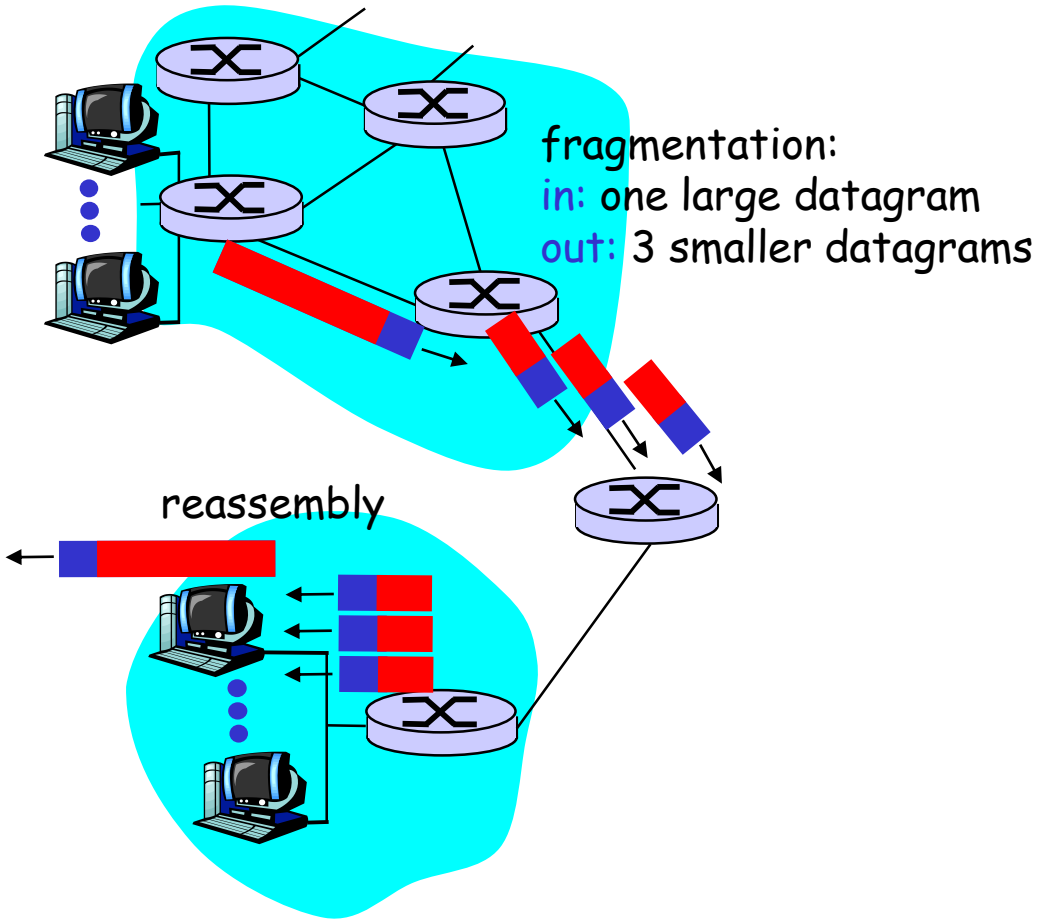
Dealing with Fragmentation

A closer look at fragmentation

- Every link has a "Maximum Transmission Unit" (MTU)
  - Largest number of bits it can carry as one unit
- A router can split a packet into multiple "fragments" if the packet size exceeds the link's MTU
- Must reassemble to recover original packet

IP Fragmentation

- IPv4 fragments at router
  - One datagram becomes several datagrams
  - IP header bits used to identify, order related fragments
- IP re-assembles at destination only

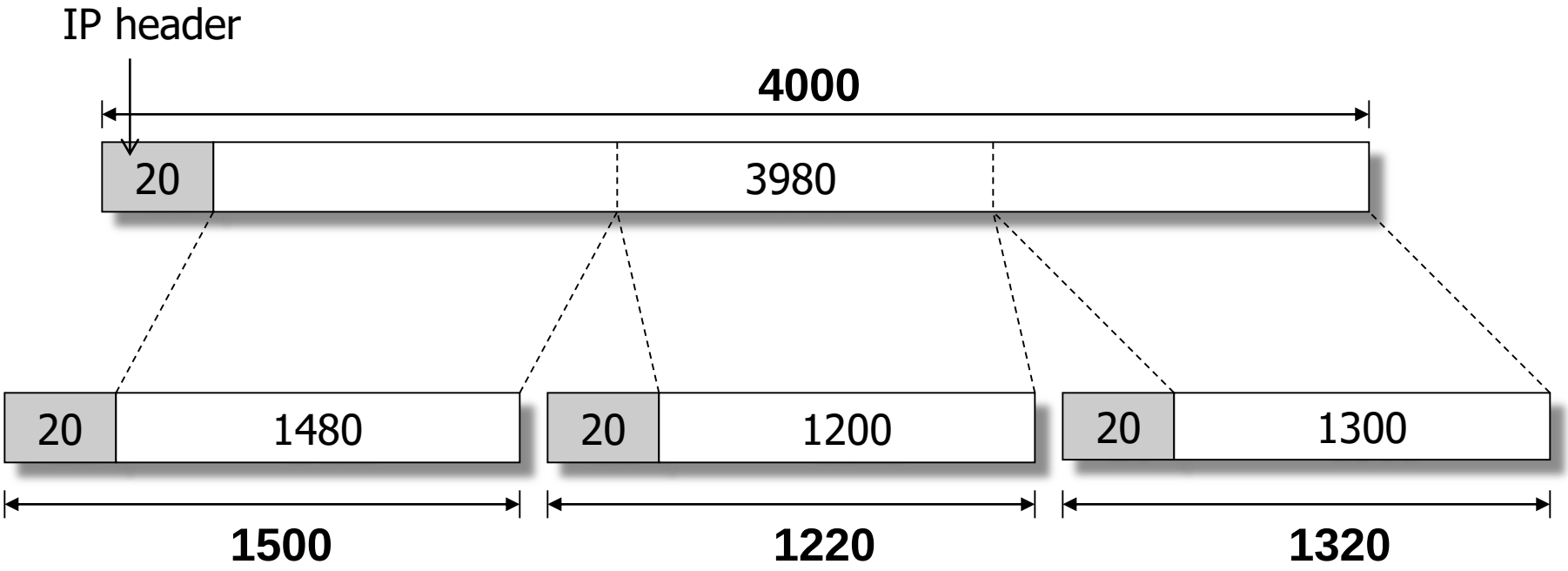


Example of fragmentation

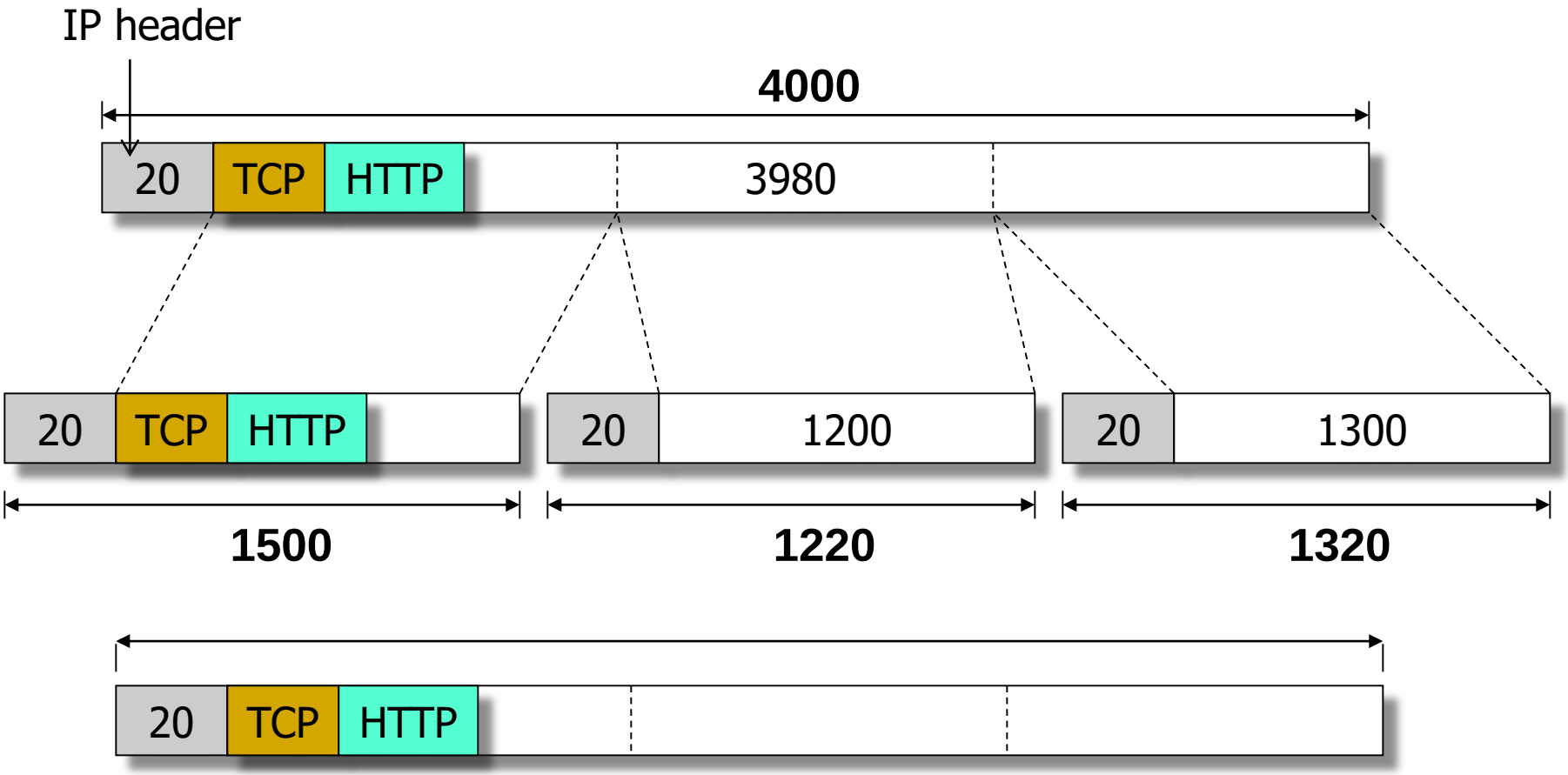
- A 4000 byte packet crosses a link w/ MTU=1500B

Example of fragmentation

- A 4000 byte packet crosses a link w/ MTU=1500B



Why reassemble?



Must reassemble before sending the packet to the higher layers!

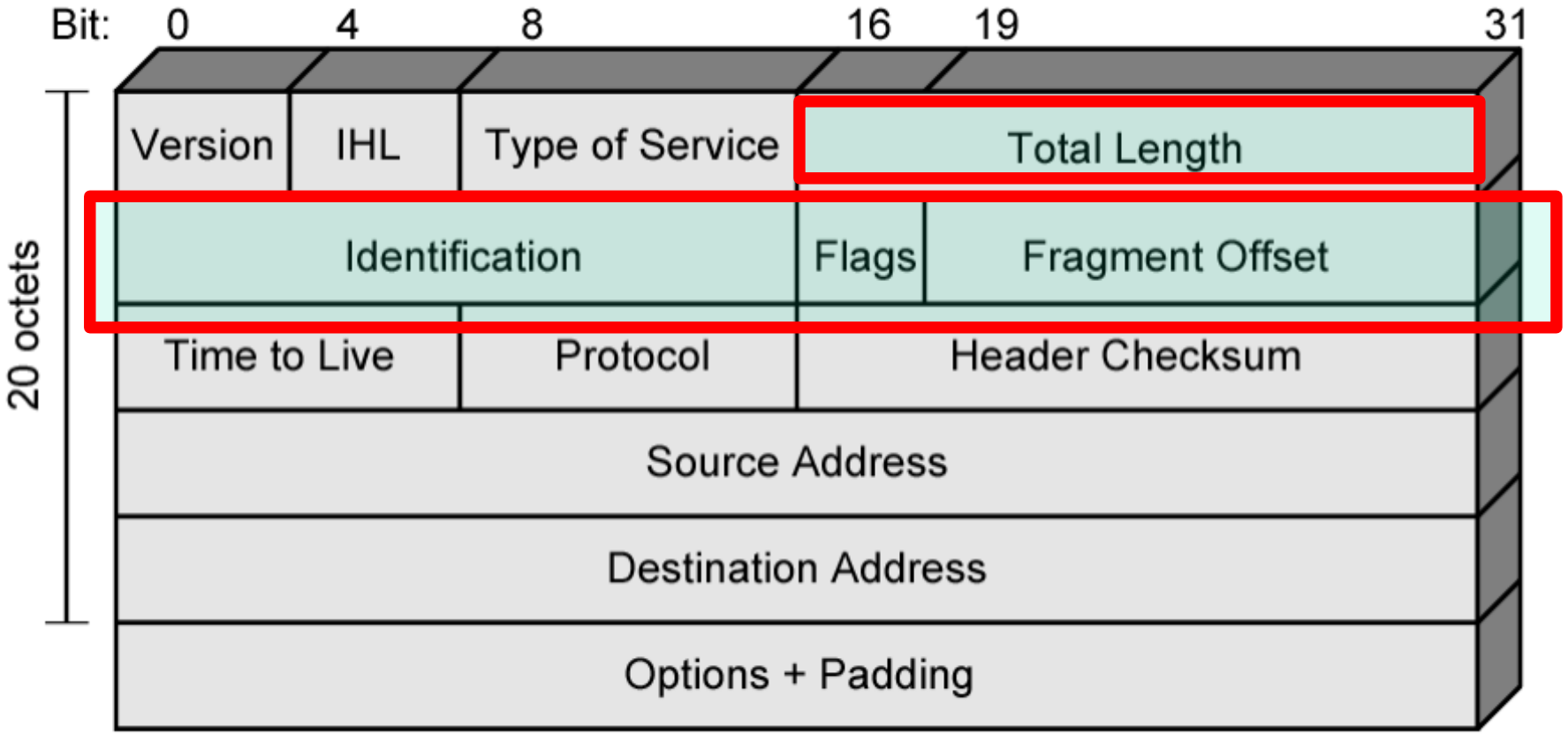
Reassembly: What fields?

- Need a way to identify fragments of the packet
  - Introduce an identifier
- Fragments can get lost
  - Need some form of sequence number or offset
- Sequence numbers / offset
  - How do I know when I have them all? (need max seq# / flag)
  - What if a fragment gets re-fragmented?

IP Fragmentation Fields

- Data Unit Identifier (ID)
  - Identifies end system originated datagram, also needs:
  - Source and destination address, Upper layer (e.g. TCP)
- Data length
  - Length of user data in octets including header
- Offset
  - Position of fragment of user data in original datagram
  - In multiples of 64 bits (i.e. 8 octets)
- More flag
  - Indicates that this is not the last fragment

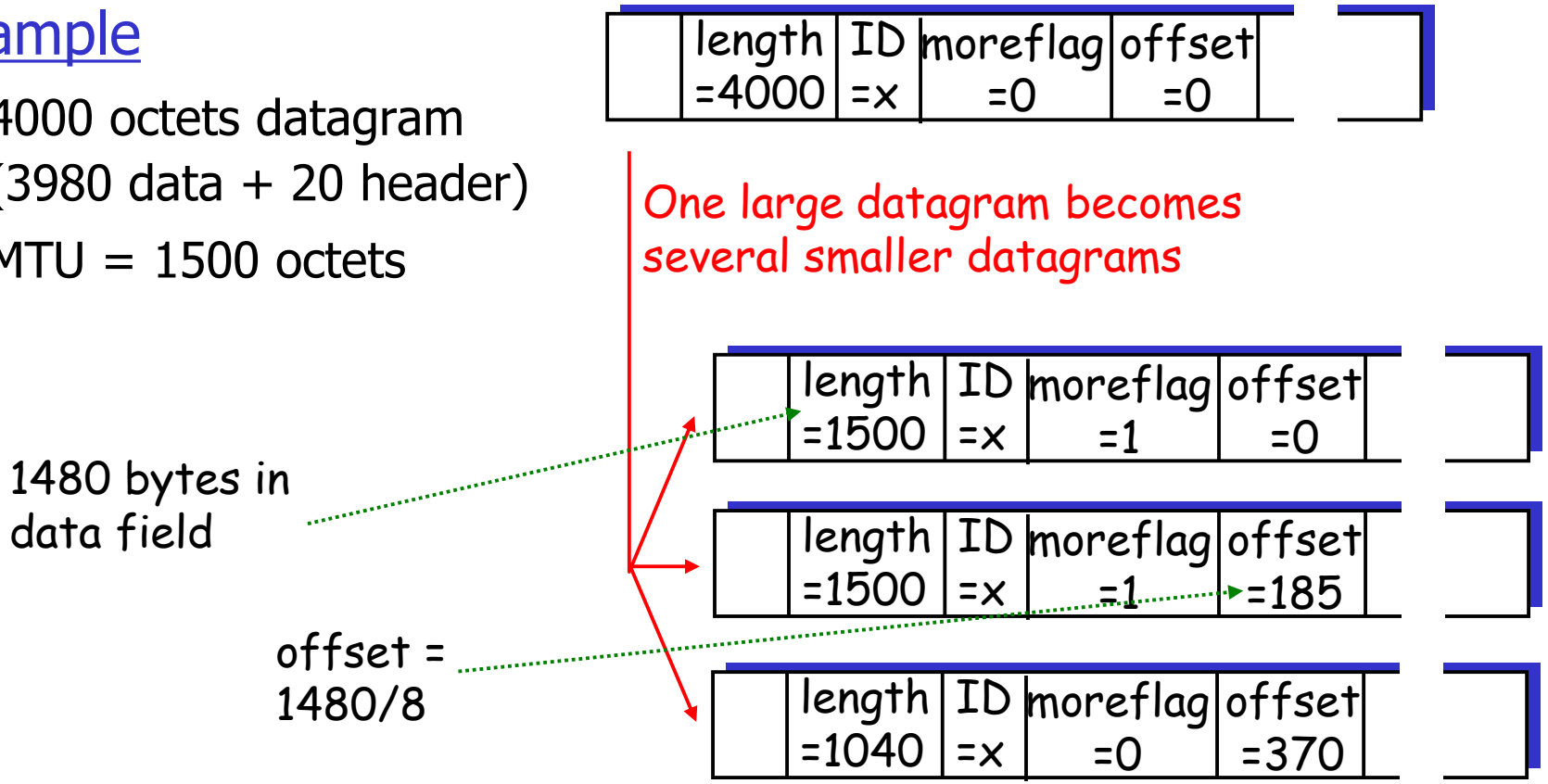
IP Field for Fragmentation



Fragmentation Example

Example

- 4000 octets datagram (3980 data + 20 header)
- MTU = 1500 octets



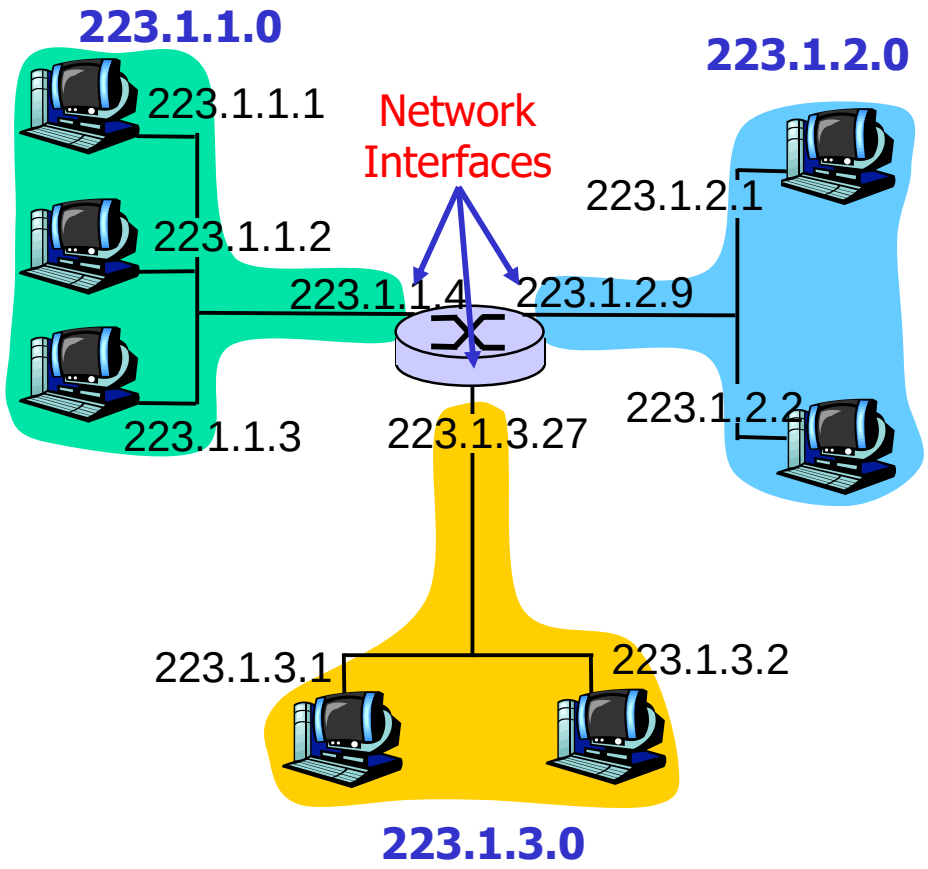
Datagram Re-assembly

- Must prepare enough buffer space at reassembly point
- As fragments with the same ID arrive, data are inserted in proper position in the buffer
  - Use *Length* and *Offset* header fields
  - Use *More flag* to determine if end fragment arrived
- Until entire data field is reassembled
  - Starting with an *Offset* of 0 and ending with a false *More flag*

IP Address

IP Addressing

- IP address
  - 32 bit global internet address for each interface
  - Network part (high order bits)
  - Host part (low order bits)
- Physical network (from IP perspective)
  - Can reach each other without intervening router

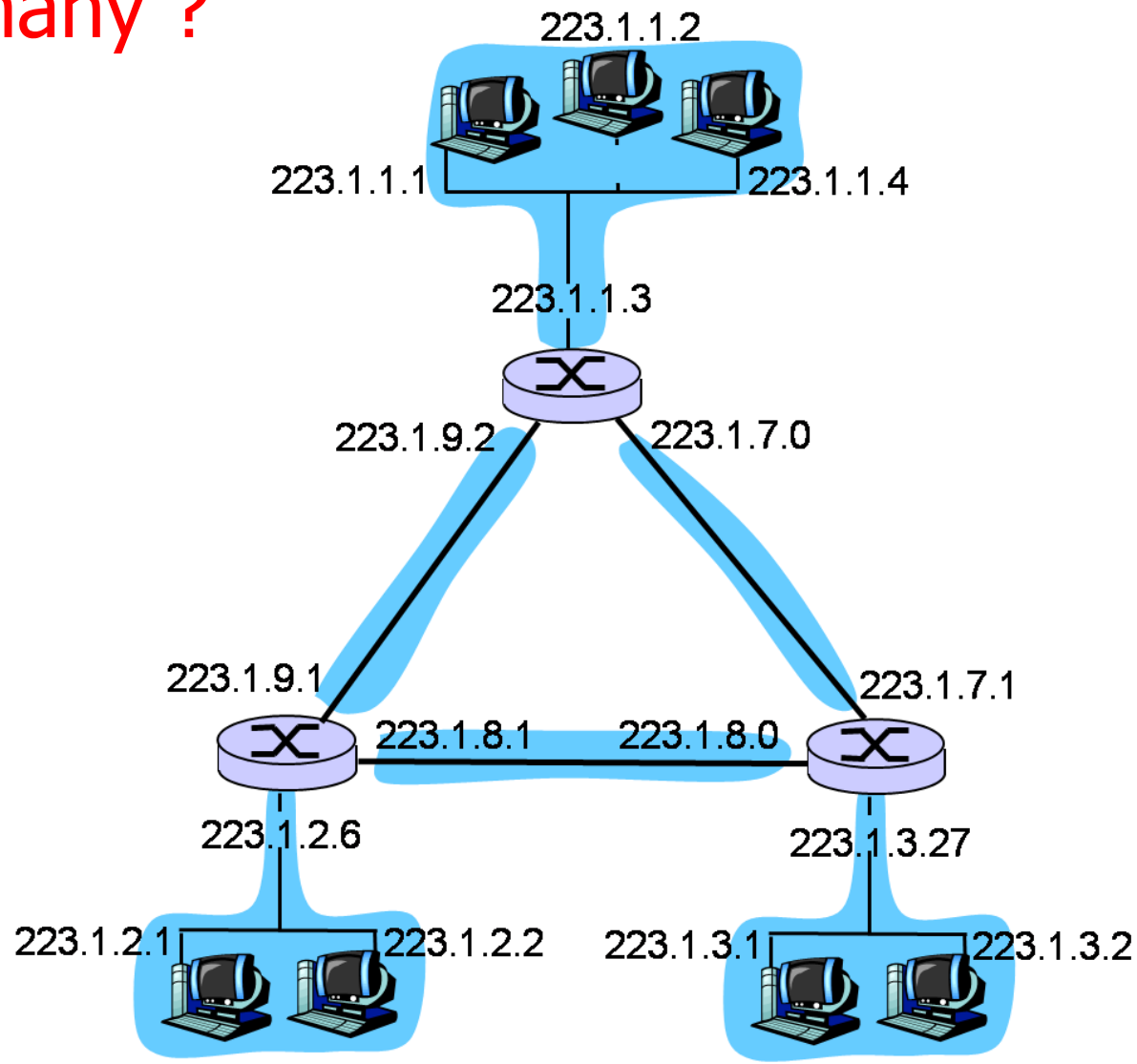


223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Count the Physical Networks

- How many ?

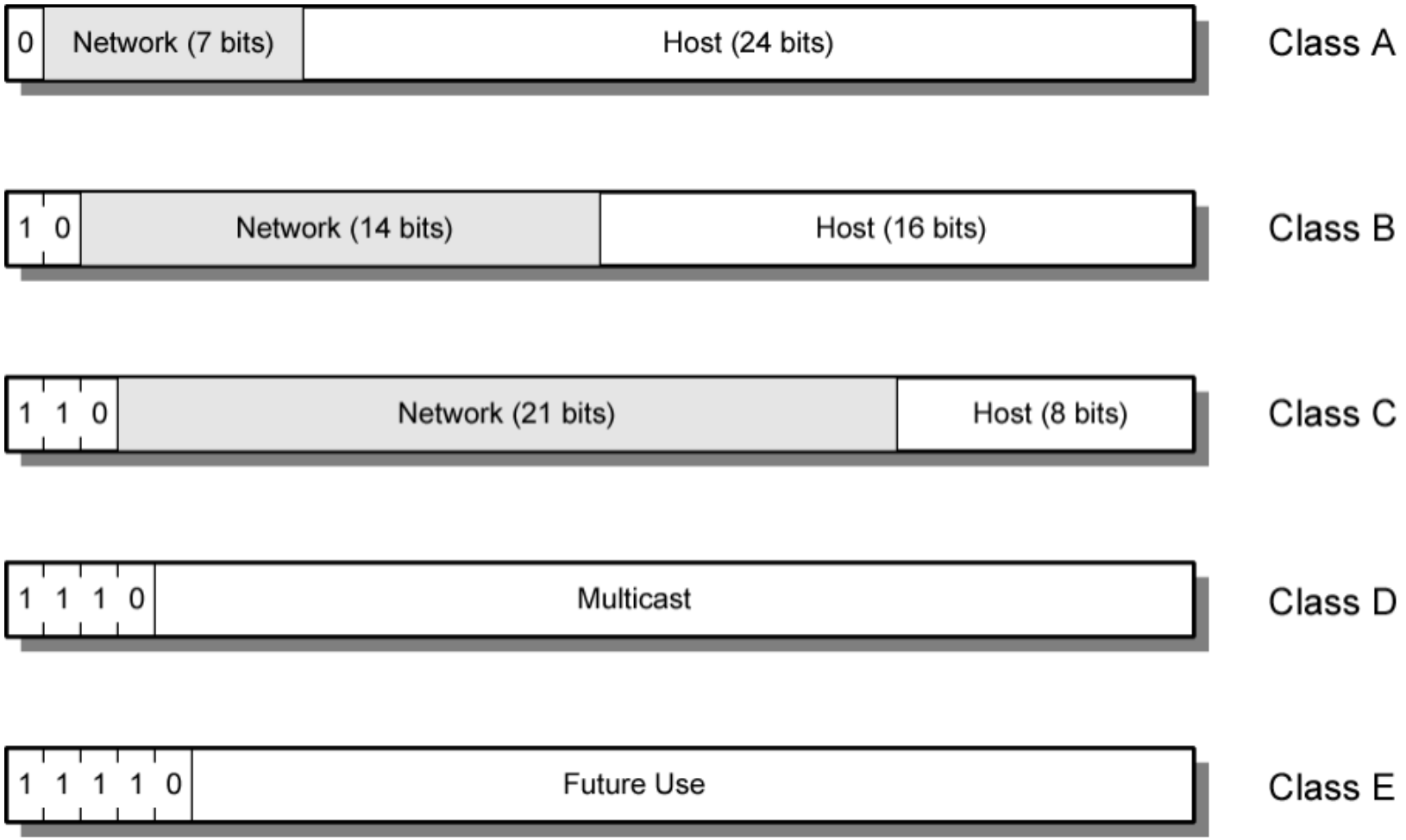


IP Address

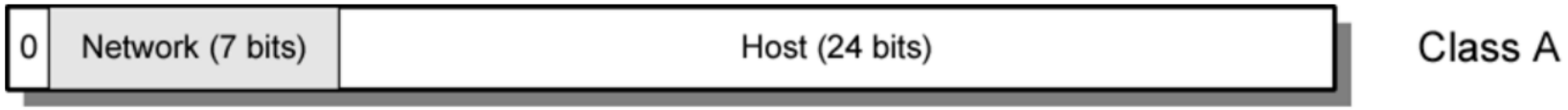
- A separate address is required for each physical interface of a host/router to a network
  - Facilitates routing
- Use Dotted-Decimal Notation

- netid unique & administered by
  - American Registry for Internet Numbers (ARIN)
  - Reseaux IP Europeens (RIPE)
  - Asia Pacific Network Information Centre (APNIC)
- hostid assigned within designated organization

IPv4 Address Formats



IP Addresses – Class A



- Start with binary 0
- Reserved *netid*
  - All 0 reserved
  - 01111111 (127) reserved for loopback
- Range 1.x.x.x to 126.x.x.x
- Up to 16 million hosts
- All allocated

A类地址：
首位为0；
支持2^7-2=126个网段；
每个网段支持主机数为2^{24}-2
=16777214（全0和全1的地址要扣除，全0是网络号，全1是广播号）

127.*.*.*: 回环测试，用于测试本地网卡。127.0.0.1 “localhost”

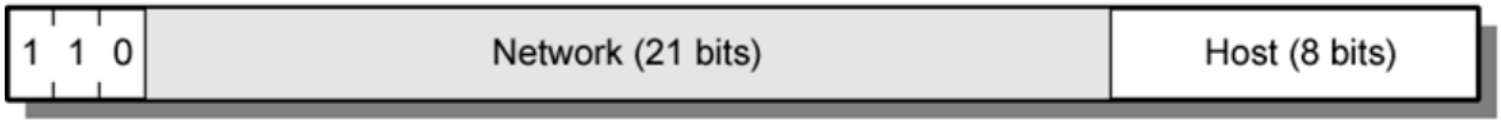
IP Addresses – Class B

- Start with 10
- Range 128.0.x.x to 191.255.x.x
- Second Octet also included in network address

- 2^{14} = 16,384 class B networks
- Up to 65,000 (=2^{16}-2) hosts

- All allocated

IP Addresses – Class C

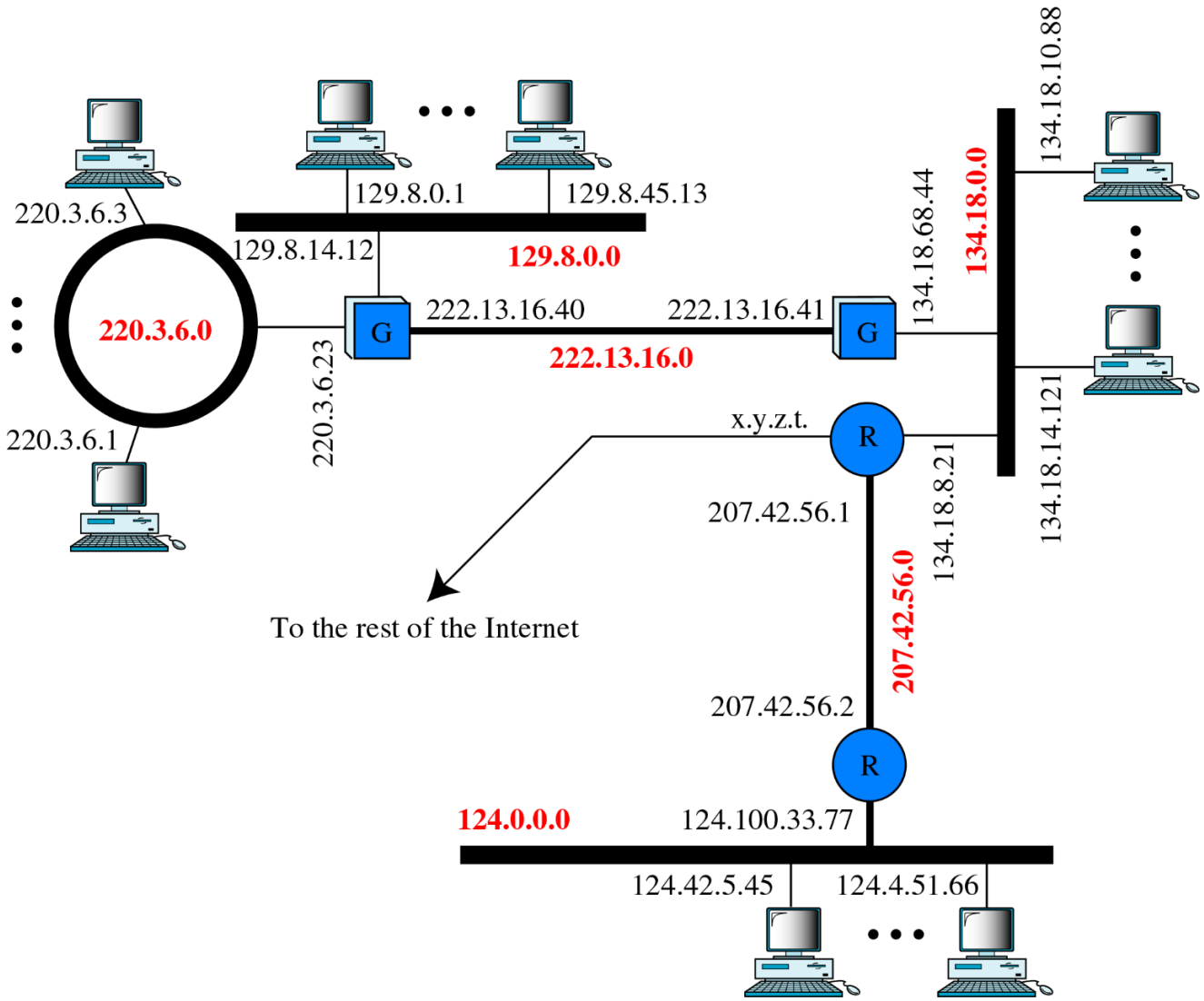


- Start with 110
- Range 192.0.0.x to 223.255.255.x
- Second and third octet also part of network address

- 2^{21} = 2,097,152 networks
- Up to 254 (=2^8-2) hosts

- Nearly all allocated

Inter-Networks with Addresses



IP Address Classes Exercise

Address	Class	Network	Host
10.2.1.1			
128.63.2.100			
201.222.5.64			
192.6.141.2			
130.113.64.16			
256.241.201.10			

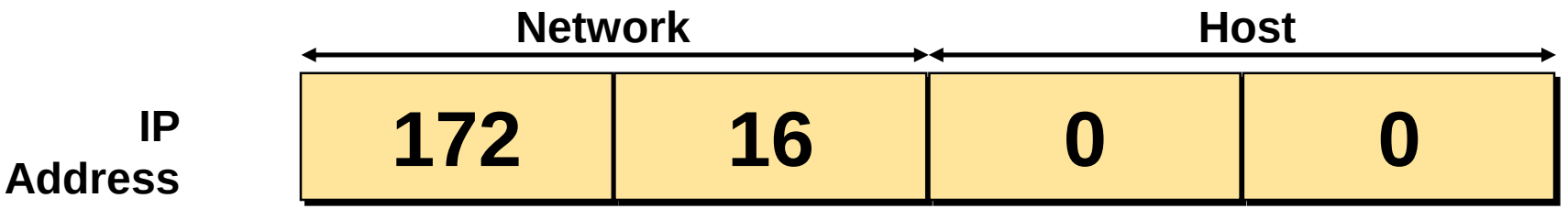
IP Address Classes Exercise

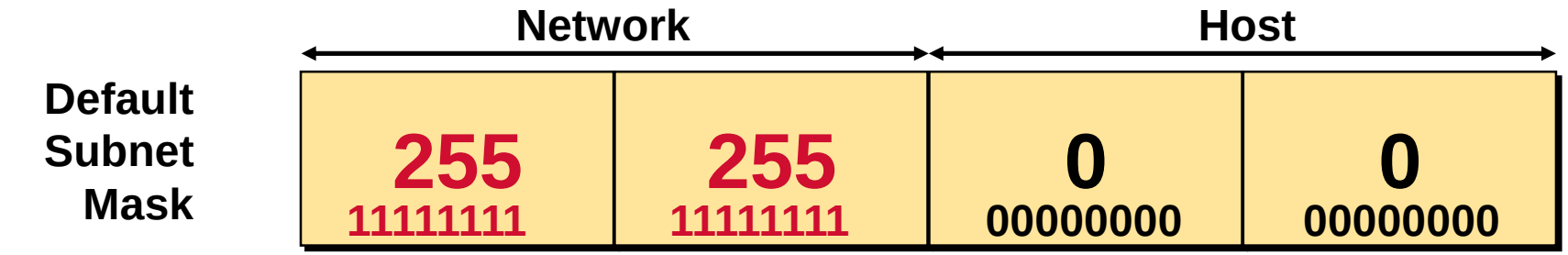
Address	Class	Network	Host
10.2.1.1	A	10.0.0.0	0.2.1.1
128.63.2.100	B	128.63.0.0	0.0.2.100
201.222.5.64	C	201.222.5.0	0.0.0.64
192.6.141.2	C	192.6.141.0	0.0.0.2
130.113.64.16	B	130.113.0.0	0.0.64.16
256.241.201.10	Nonexistent		

Subnets and Subnet Masks

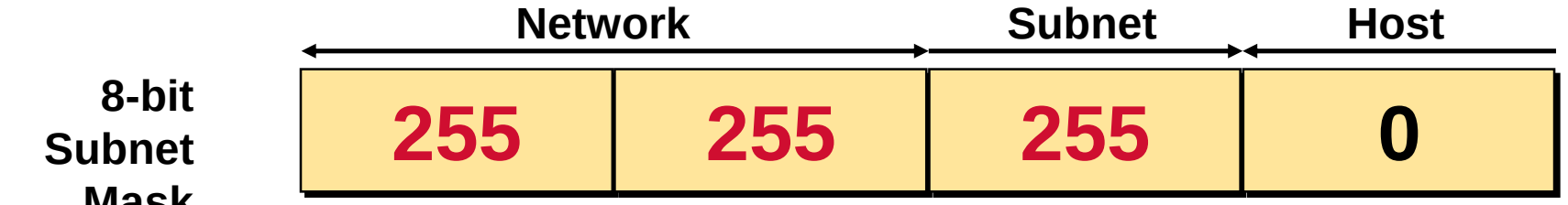
- Handle problem of network address inadequacy
- Host portion of address partitioned into subnet number and host number
  - Subnet mask indicates which bits are subnet number and which are host number
  - Each LAN assigned a subnet number, more flexibility
  - Local routers route within subnetted network
- Subnets looks to rest of internet like a single network
  - Insulate overall Internet from growth of network numbers and routing complexity

Subnets and Subnet Masks



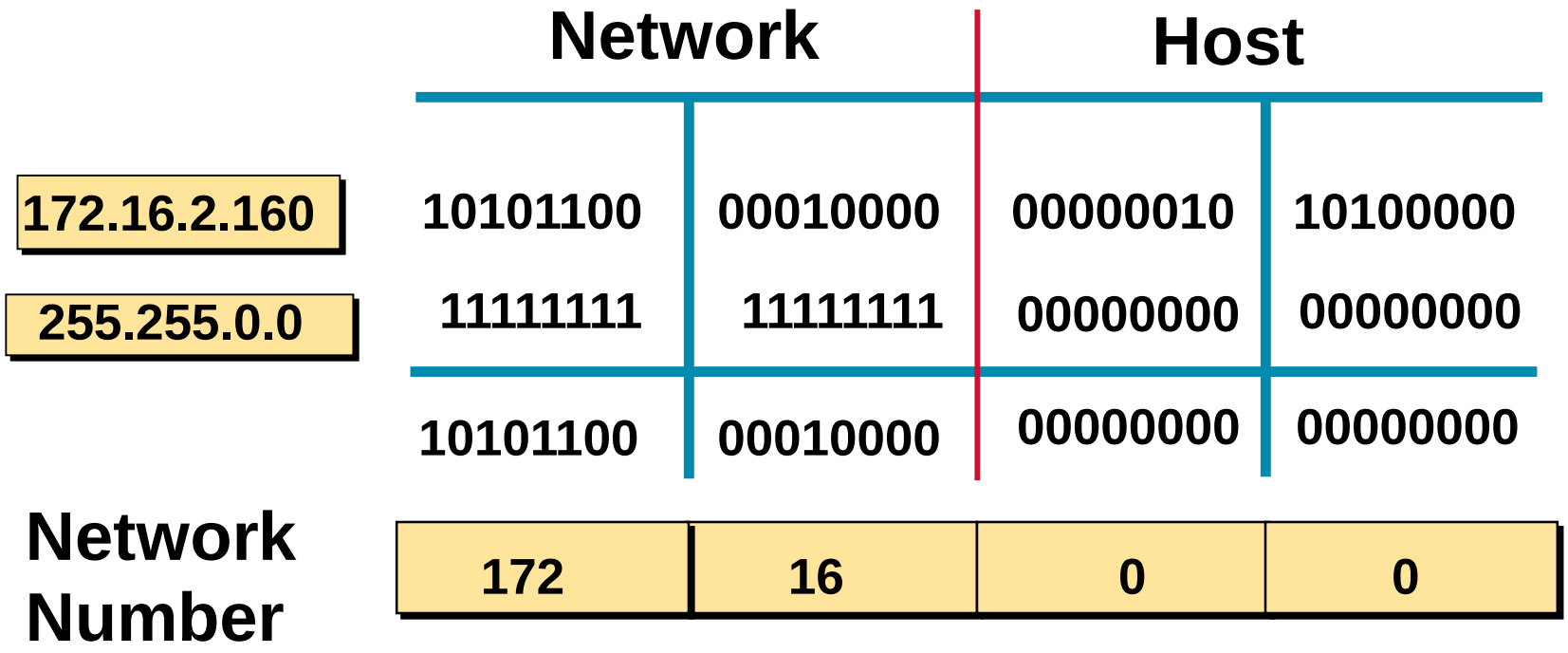


Also written as "/16" where 16 represents the number of 1s in the mask.



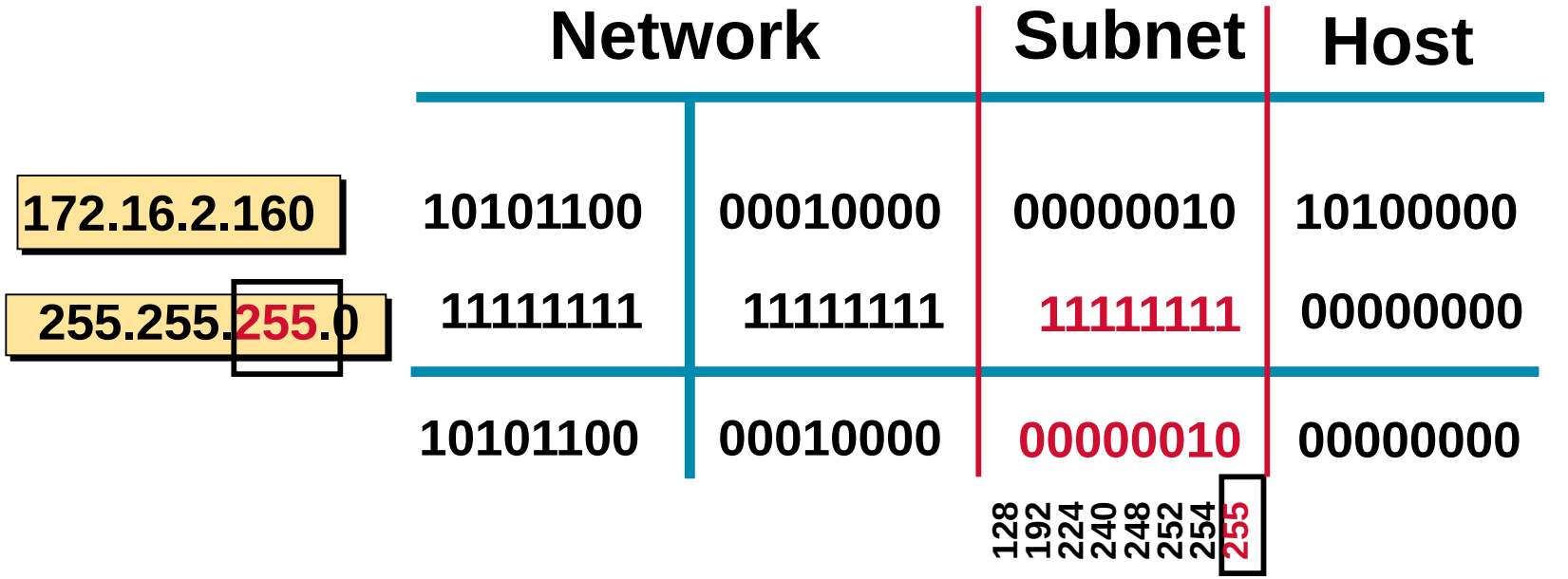
Also written as "/24" where 24 represents the number of 1s in the mask.

Subnet Mask without Subnets



Subnets not in use—the default

Subnet Mask with Subnets



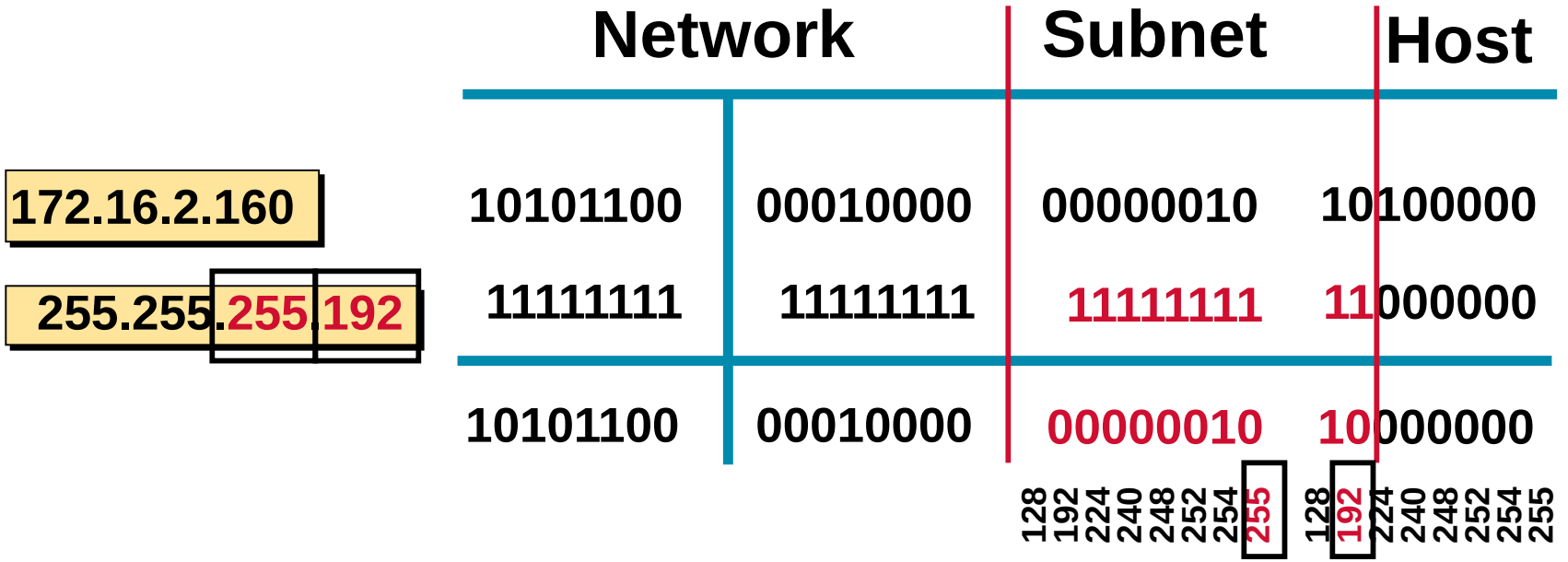
Network Number

172	16	2	0

Host No: 2-254

Network number extended by eight bits

Subnets and Subnet Masks



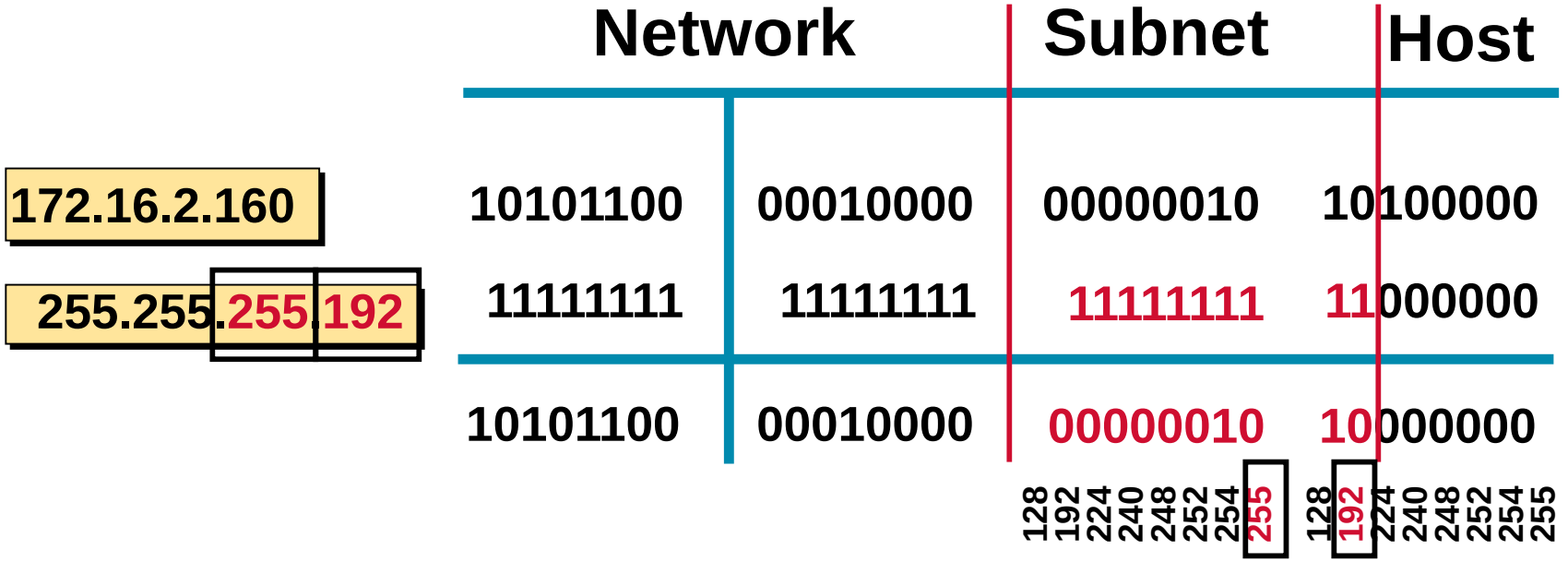
Network Number

172	16	2	128

Host No: 130-190

Network number extended by ten bits

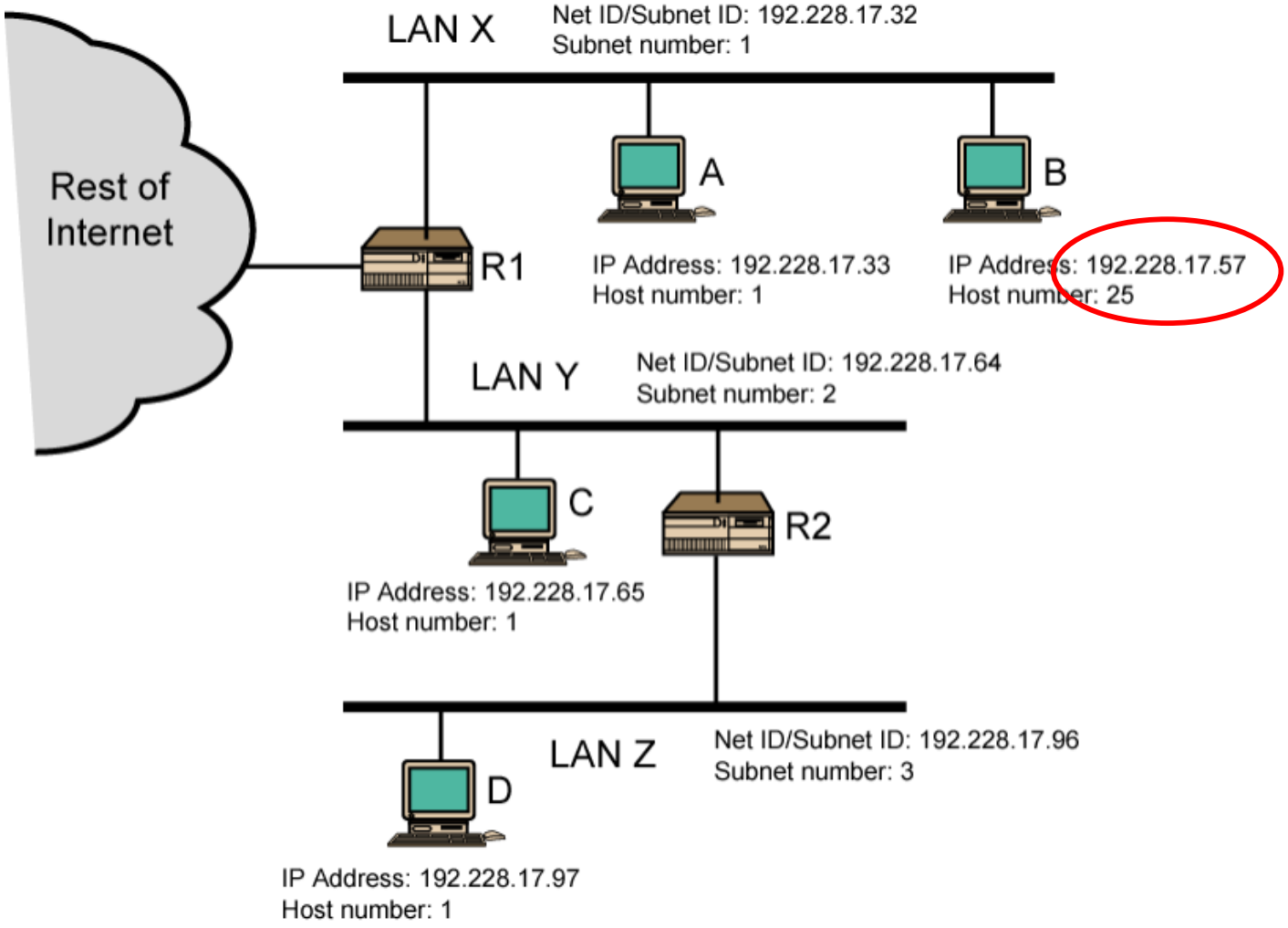
Subnets and Subnet Masks



Network Number

172	16	2	0	Host No: 2-62
172	16	2	64	Host No: 66-126
172	16	2	128	Host No: 130-190
172	16	2	192	Host No: 194-254

Routing Using Subnets (1)



Routing Using Subnets (2)

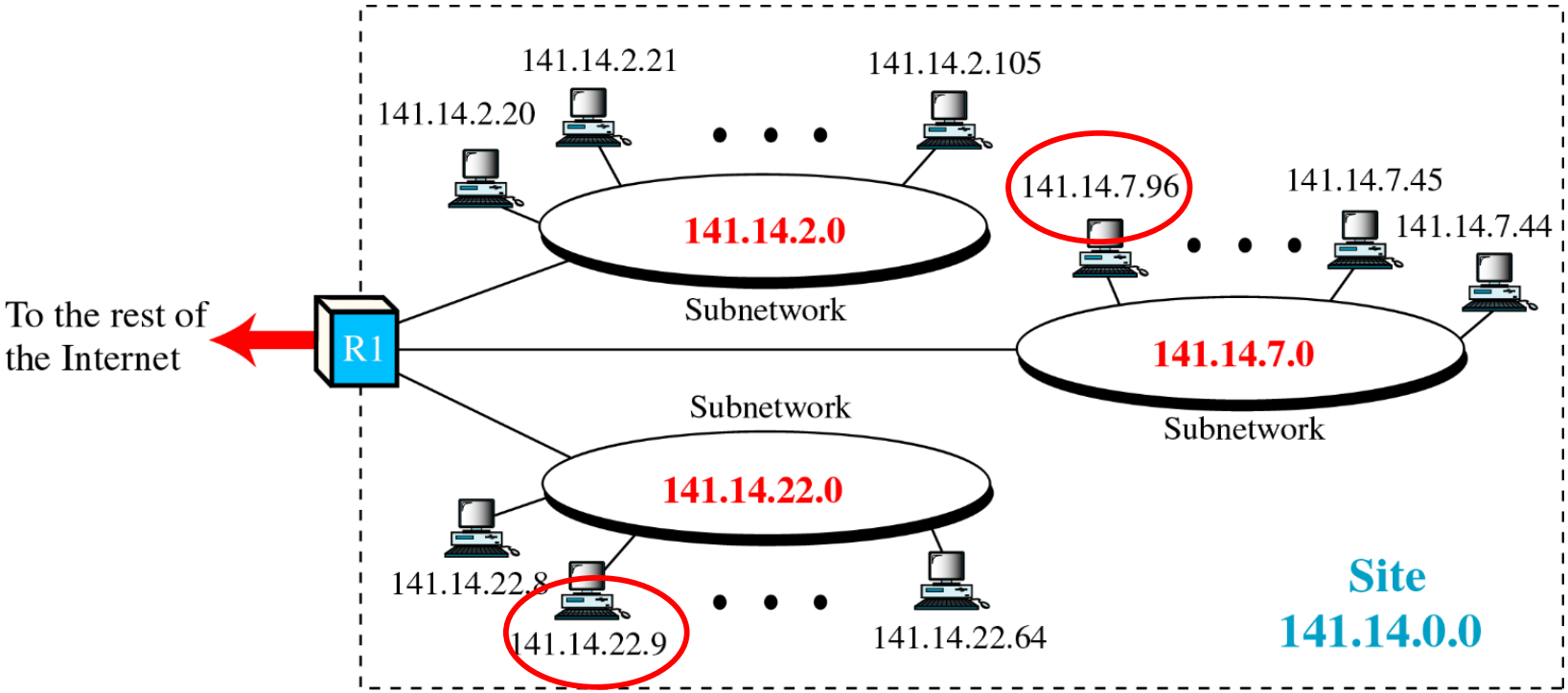
(a) Dotted decimal and binary representations of IP address and subnet masks

	Binary Representation	**Dotted Decimal**
IP address	11000000.11100100.00010001.00111001	192.228.17.57
Subnet mask	11111111.11111111.11111111.11100000	255.255.255.224
Bitwise AND of address and mask (resultant network/subnet number)	11000000.11100100.00010001.00100000	192.228.17.32
Subnet number	11000000.11100100.00010001.001	1
Host number	00000000.00000000.00000000.00011001	25

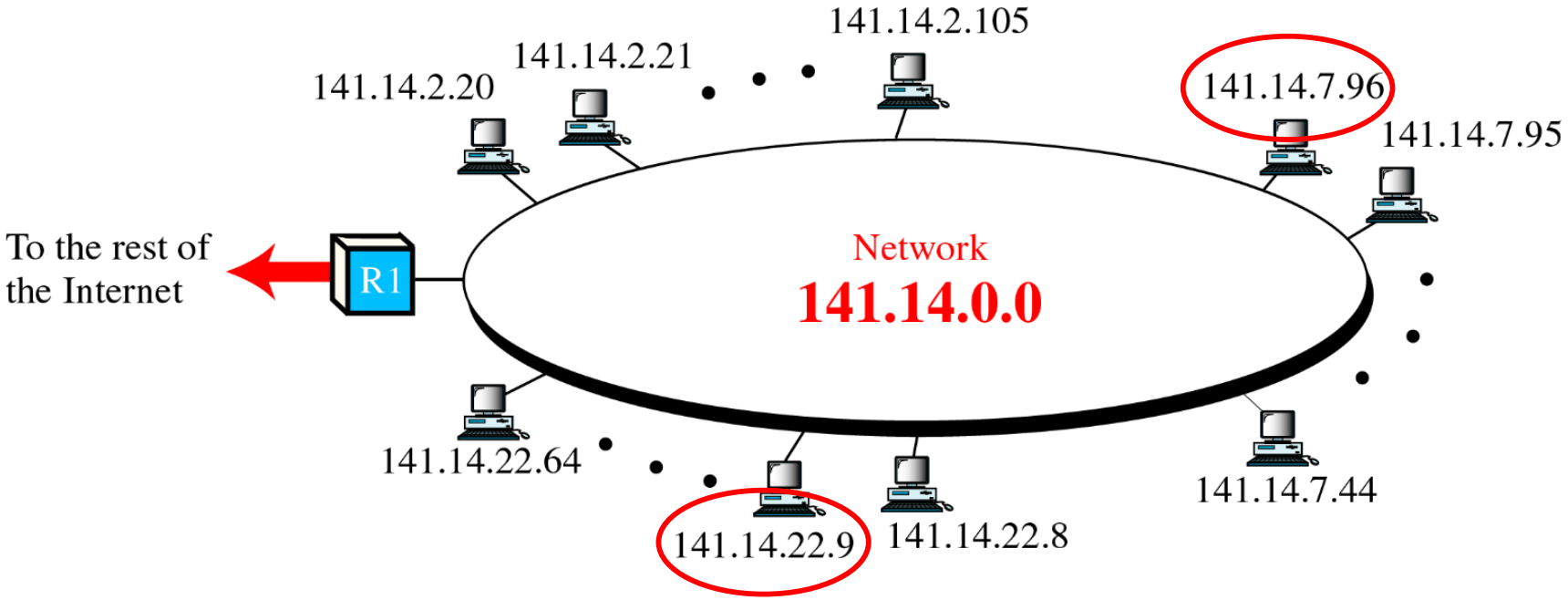
(b) Default subnet masks

	Binary Representation	**Dotted Decimal**
Class A default mask	11111111.00000000.00000000.00000000	255.0.0.0
Example Class A mask	11111111.11000000.00000000.00000000	255.192.0.0
Class B default mask	11111111.11111111.00000000.00000000	255.255.0.0
Example Class B mask	11111111.11111111.11111000.00000000	255.255.248.0
Class C default mask	11111111.11111111.11111111.00000000	255. 255. 255.0
Example Class C mask	11111111.11111111.11111111.11111100	255. 255. 255.252

Subnets Example



Subnets to the Rest



CIDR Notation

- Classless Inter Domain Routing (CIDR)
  - An IP address is represented as "*A.B.C.D/n*", where *n* is called the IP (network) prefix

IP Address	10 .	217 .	123 .	7
	00001010	11011001	01111011	00000111
Subnet	255 .	255 .	240 .	0
	11111111	11111111	11110000	00000000
Network ID	00001010	11011001	01110000	00000000
CIDR	10.217.112.0/20			

CIDR Notation

- Identifying a CIDR block requires both an address and a mask
  - Slash notation
  - 128.211.168.0/21 for addresses 128.211.168.0 – 128.211.175.255
  - Here the /21 indicates a 21 bit mask
  - All possible CIDR masks can easily be generated
  - /8, /16, /24 correspond to traditional class A, B, C categories
- IP addresses are now arbitrary integers, not classes

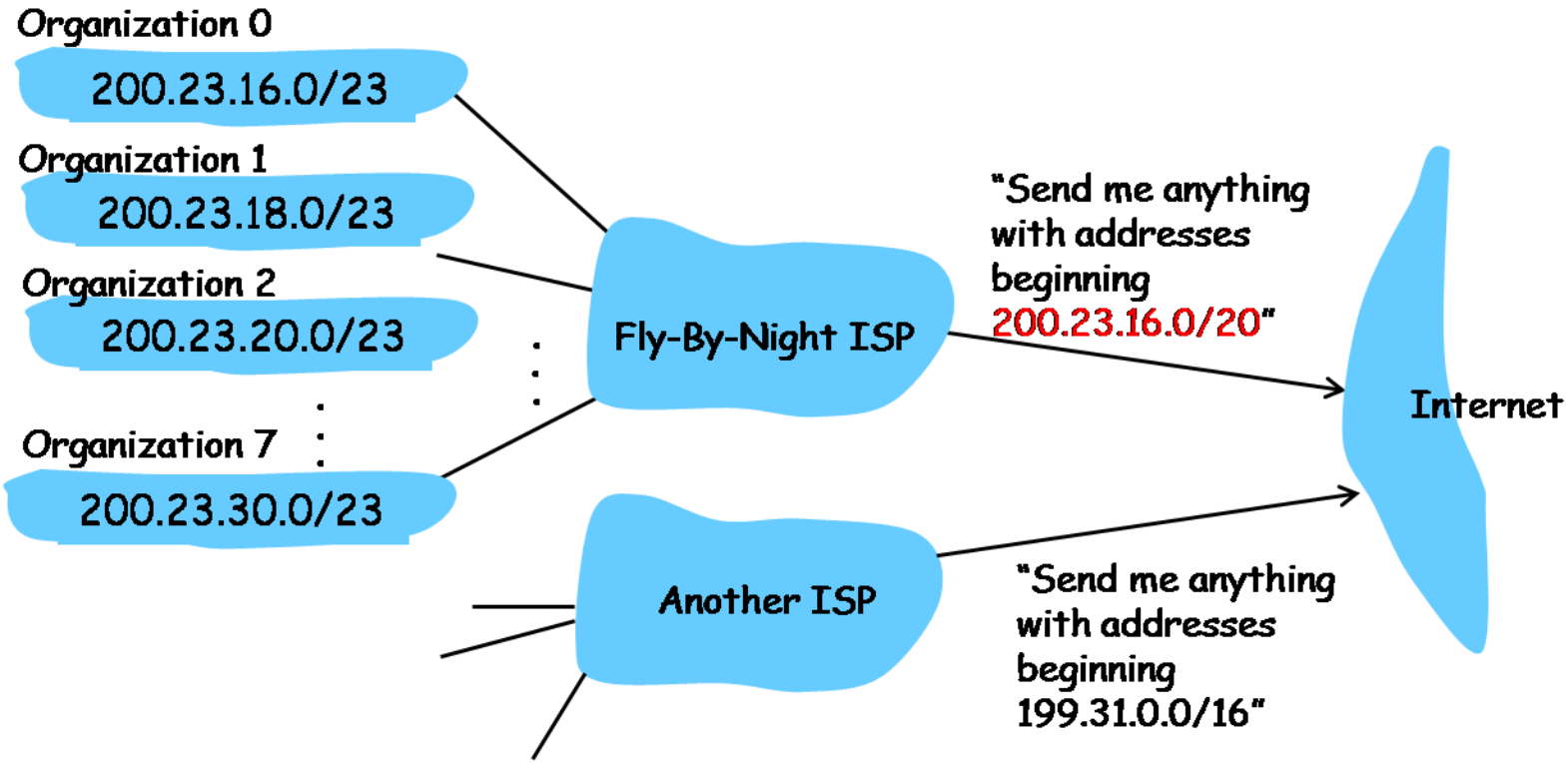
More General Case

- An ISP can be looked as a set of subnets
  - Support many organizations (Intranets)
  - Hierarchical addressing

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

Route Aggregation

- Allows efficient advertisement of routing information



Summary

- IP Operations
- IPv4包头格式
- IP地址及分配（A类，B类，C类）
  - 子网掩码
  - CIDR地址表示
  - 如何进行子网划分