

---

# Internet Protocol (IP)

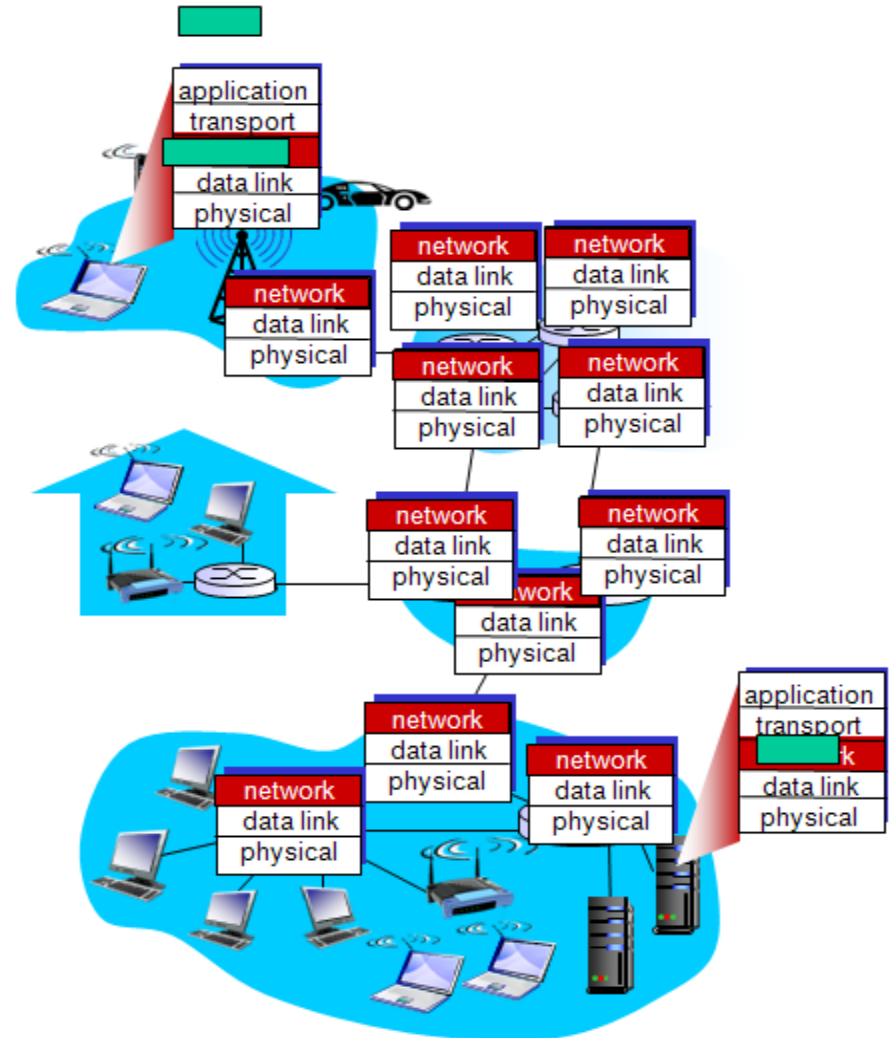
## "The" Network Layer

EE450: Introduction to Computer Networks

Professor A. Zahid

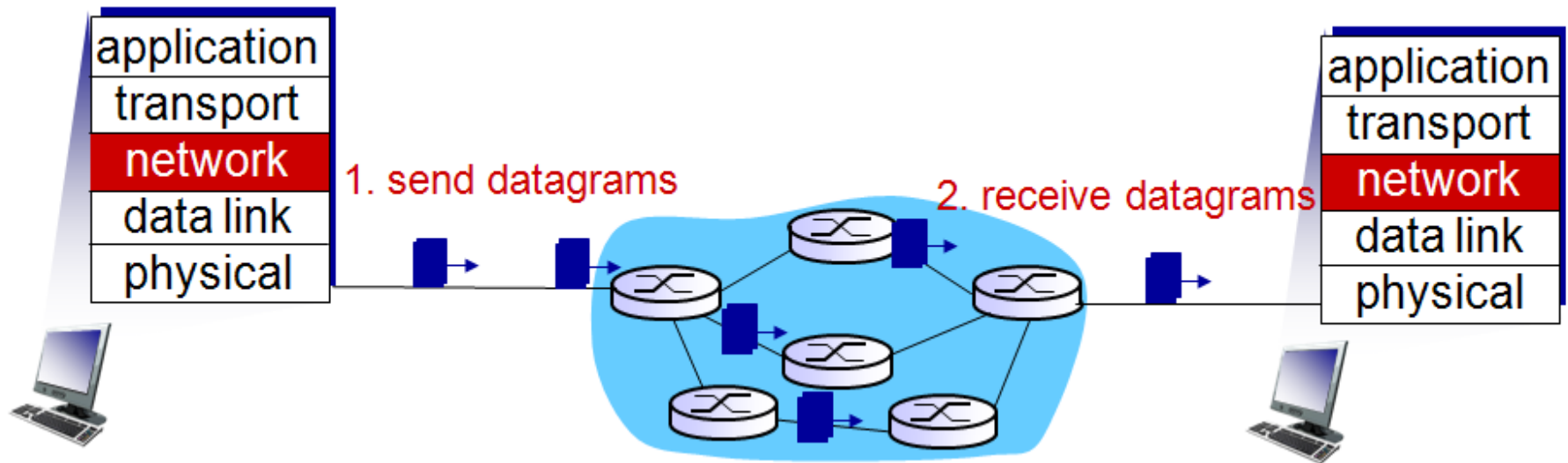
# Network Layer

- ❖ On sending side, it encapsulates TCP Segments (or UDP Datagrams) into **Packets**
- ❖ On receiving side, it delivers TCP Segments (or UDP Datagrams) to corresponding Transport layer
- ❖ Network layer protocols are implemented in **every** host and router across the Network
- ❖ Router examines header fields in all IP packets passing through it and forward according to routing tables

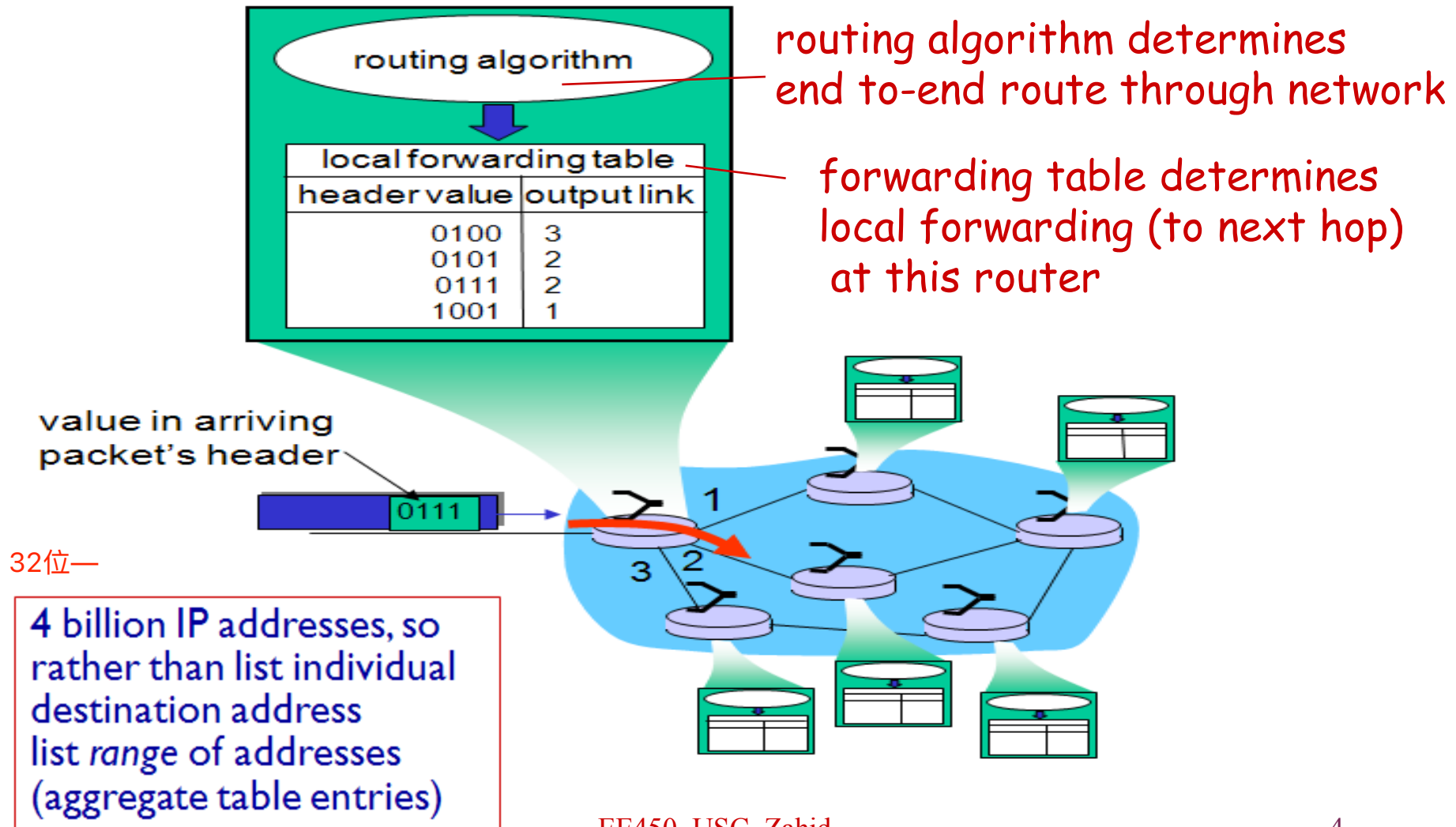


# Packet Switched Network (Review)

- No call setup at Network layer
  - Routers: no state about end-to-end connections
    - Packets may arrive out-of-order
  - Packets forwarded using destination IP address
- no establish
- router cannot remember the same Source/Destination address



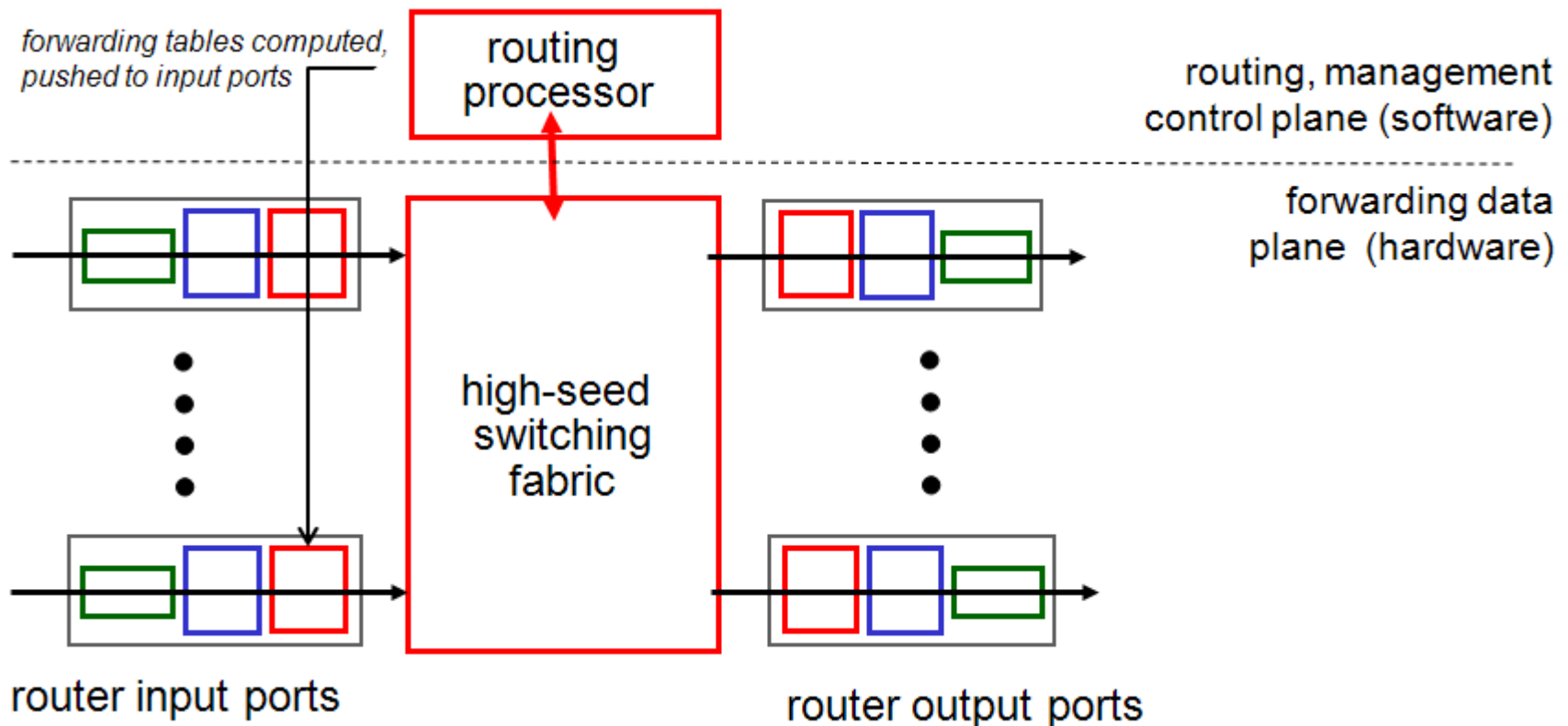
# Routing vs. Forwarding



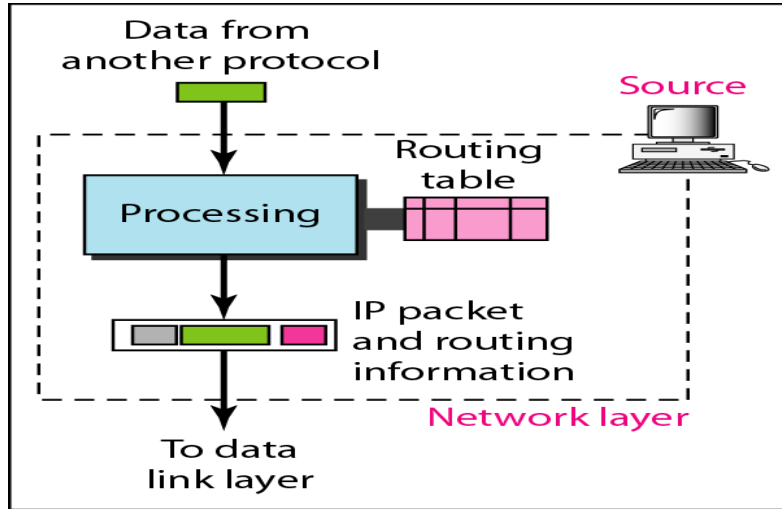
# Router Architecture (EE555)

two key router functions:

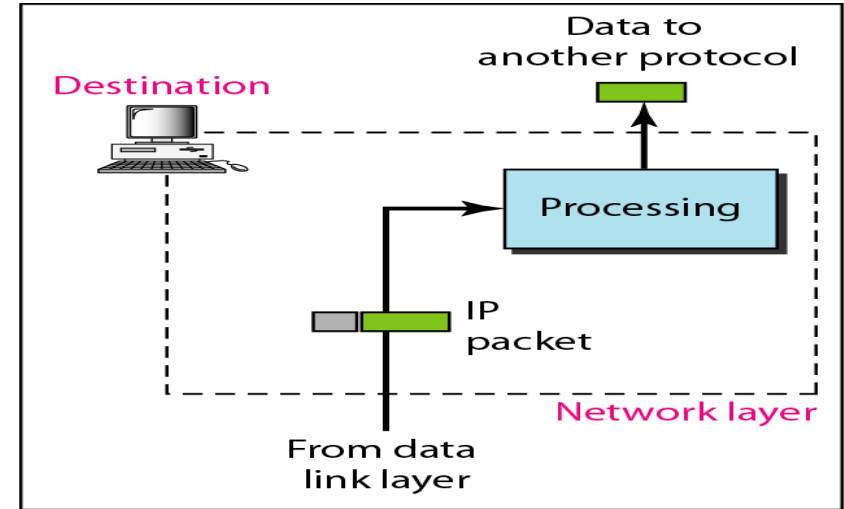
- ❖ Run Routing algorithms/protocols (RIP, OSPF, BGP)
- ❖ forwarding packets from incoming to outgoing link



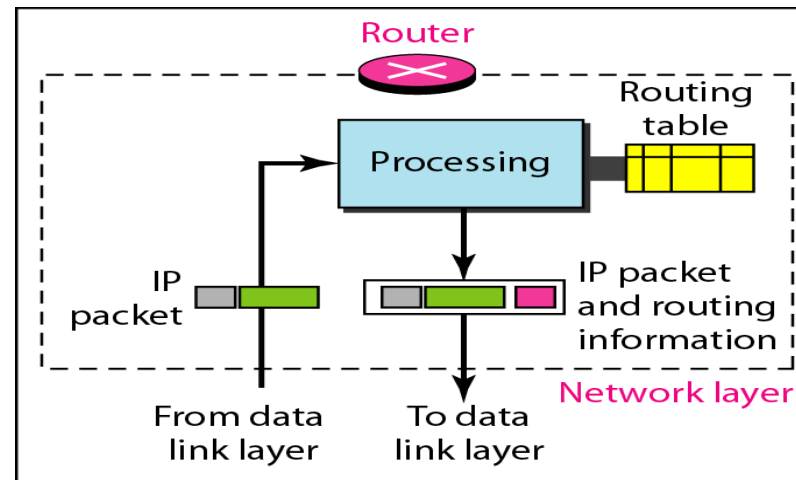
# Network Layer at Source/Router/Destination



a. Network layer at source

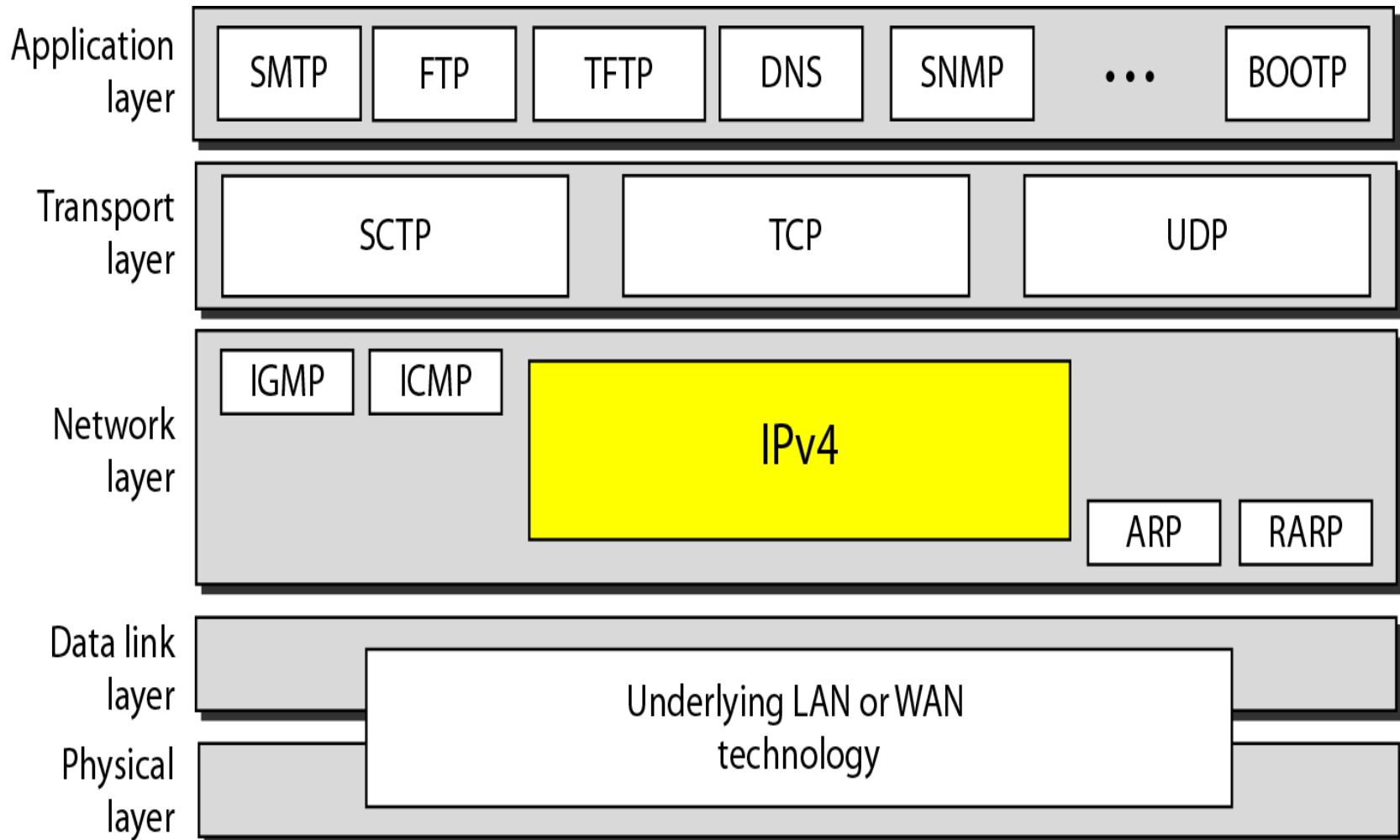


b. Network layer at destination



c. Network layer at a router

# The IP Protocol



# The Internet Protocol

---

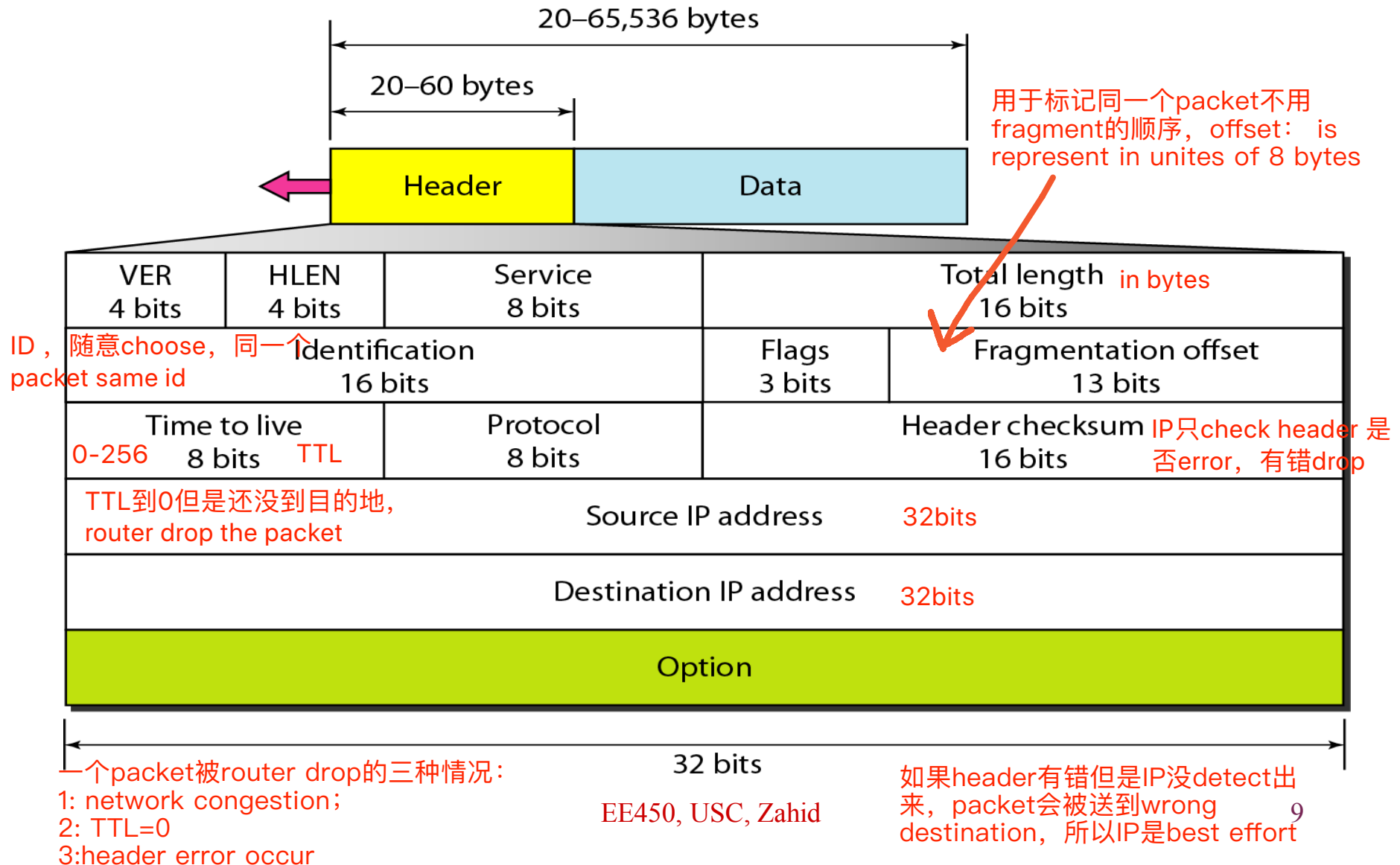
- IP is a connection-less, unreliable network layer protocol designed to be used in a connection-less packet switched network such as the Internet.
- IP provides best effort services in the sense
  - There is no guarantee of delivery of error-free packets
  - There is no guarantee of ordered delivery of packets
  - There is no guarantee of delivery of packets, i.e. some packets may be lost, some packets may be duplicated
- IP relies on upper layer transport protocols (TCP) to take care of these problems



max size:  $2^{16}=65536$  bytes

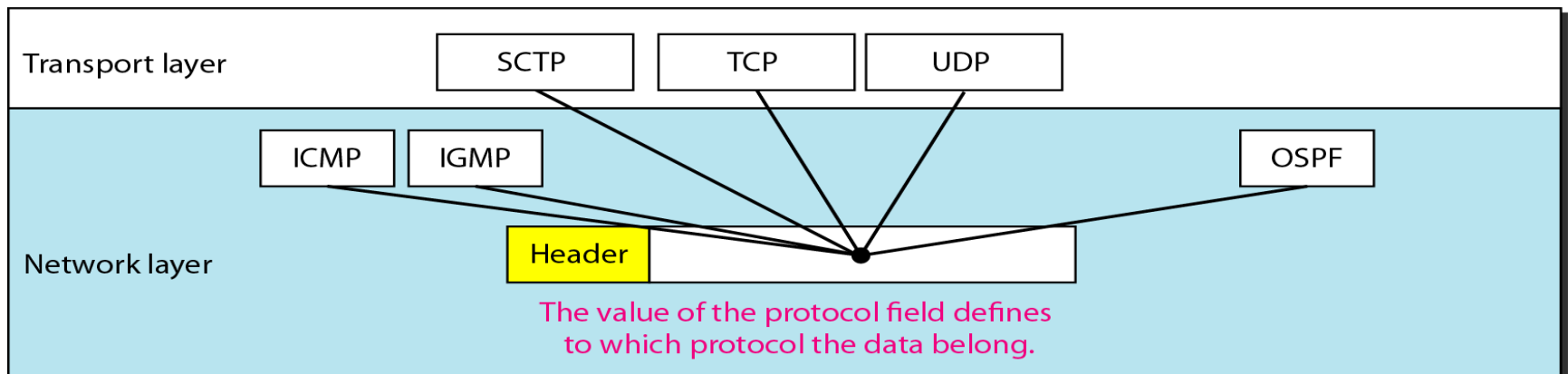
# IP Packet Format

不用记



# The Protocol Field

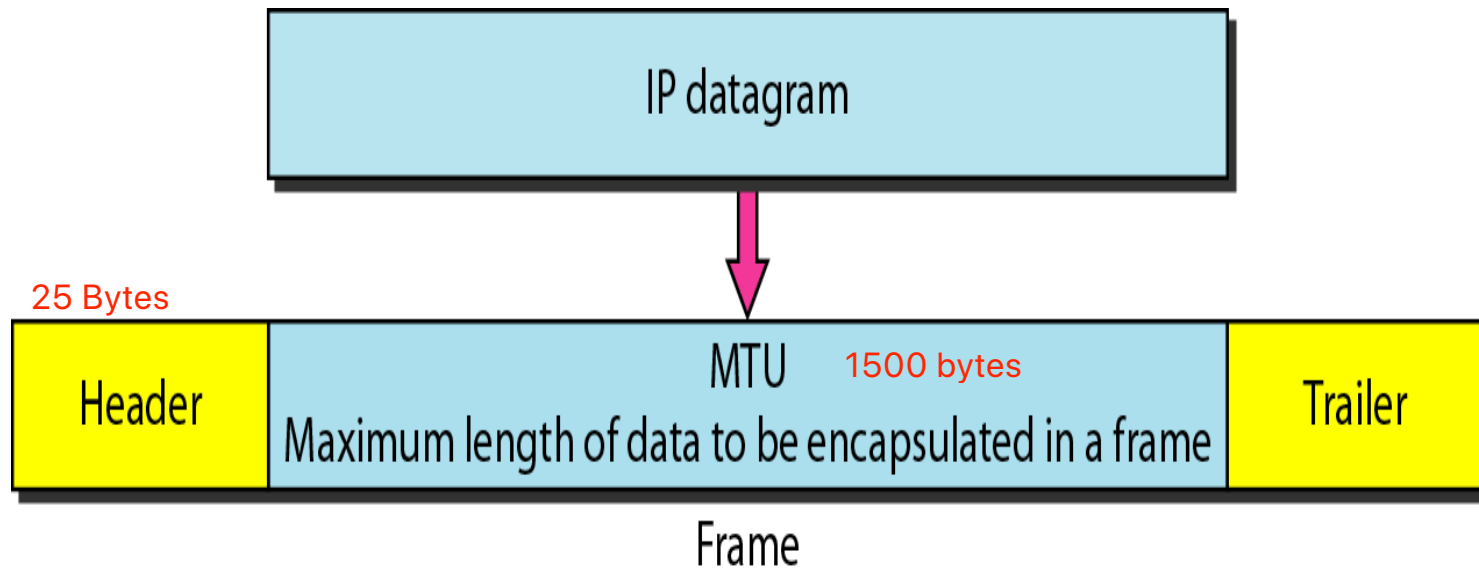
- The protocol field (8-bits) defines the protocol that is using the services of IP. It defines the final destination protocol the packet should be delivered to. This is important since several protocols could be multiplexed over IP
  - ICMP : 1, IGMP: 2, TCP: 6, EGP: 8, UDP: 17, IPv6: 41, OSPF: 89, etc...



# Fragmentation

- IP packet may travel over different networks (LANs and WANs)
- A router de-capsulate an IP packet from the frame it receives, process it, and encapsulate it in another frame
- Frame size and format varies depend on the data link protocol used by the physical network through which the frame is traveling
- MTU (Maximum Transmission Unit) is the maximum size of the data field (payload) in the frame
- If Packet size > MTU, Need for Fragmentation

# MTU: Maximum Transfer Unit



<u>Protocol</u>	<u>MTU (Octets)</u>
Ethernet	1500
Token Ring (4 Mbps)	4464
Token Ring (16 Mbps)	17914
FDDI	4352
X.25	576
PPP	296

the payload of each fragment (in bytes) must be a multiple of 8  
(except for the last fragmentation)

# Fragmentation (Continued)

---

- Each fragment has its own header (most of fields are copied, some will change, including the total length, the Flags and the fragmentation offset fields)
- A fragmented datagram may itself be fragmented if it encounters a network with smaller MTU
- A packet can be fragmented by a source host or by any router in the path. Re-assembly of the packet must be done at the destination host because those fragments become independent packets and may travel different routes

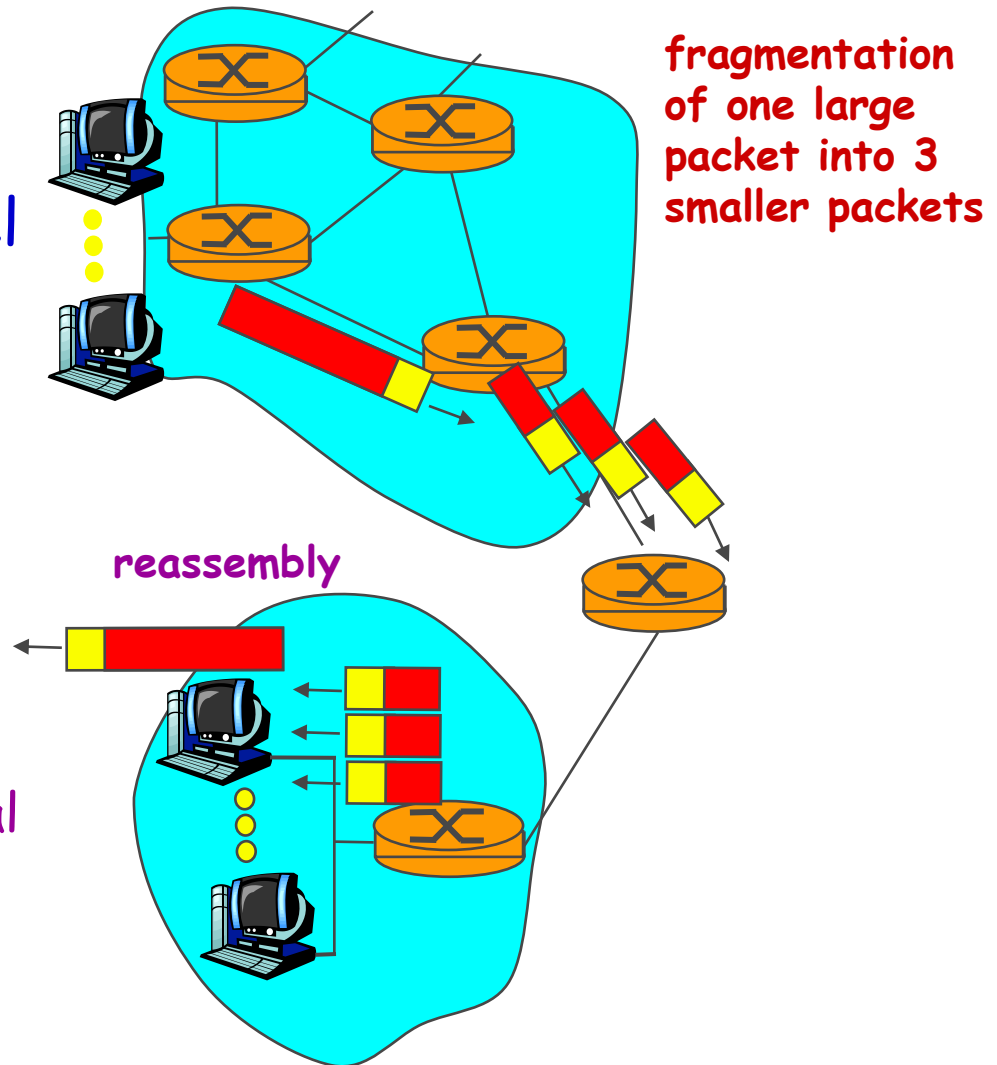
# Fragmentation and Reassembly

- Network links have MTU (max.transfer size) - largest possible link-level frame.

- Different link types, Different MTUs

- large IP datagram "fragmented" within net

- One packet becomes several packets (fragments)
  - "Reassembled" only at final destination
  - IP header bits used to identify, order related fragments



# Fields related to Fragmentation

用于让receiver判断一个packet是不是在一起

- **Identification (16 bits)**: All fragments of a packet has the same ID number which is the same as that of the original packet. The R/x knows that all fragments having the same ID should be assembled into one packet

如果D标志但是下一个router无法送这么大的packet, router会丢掉, 然后发送ICMP

- **Flags (3 bits)**: 

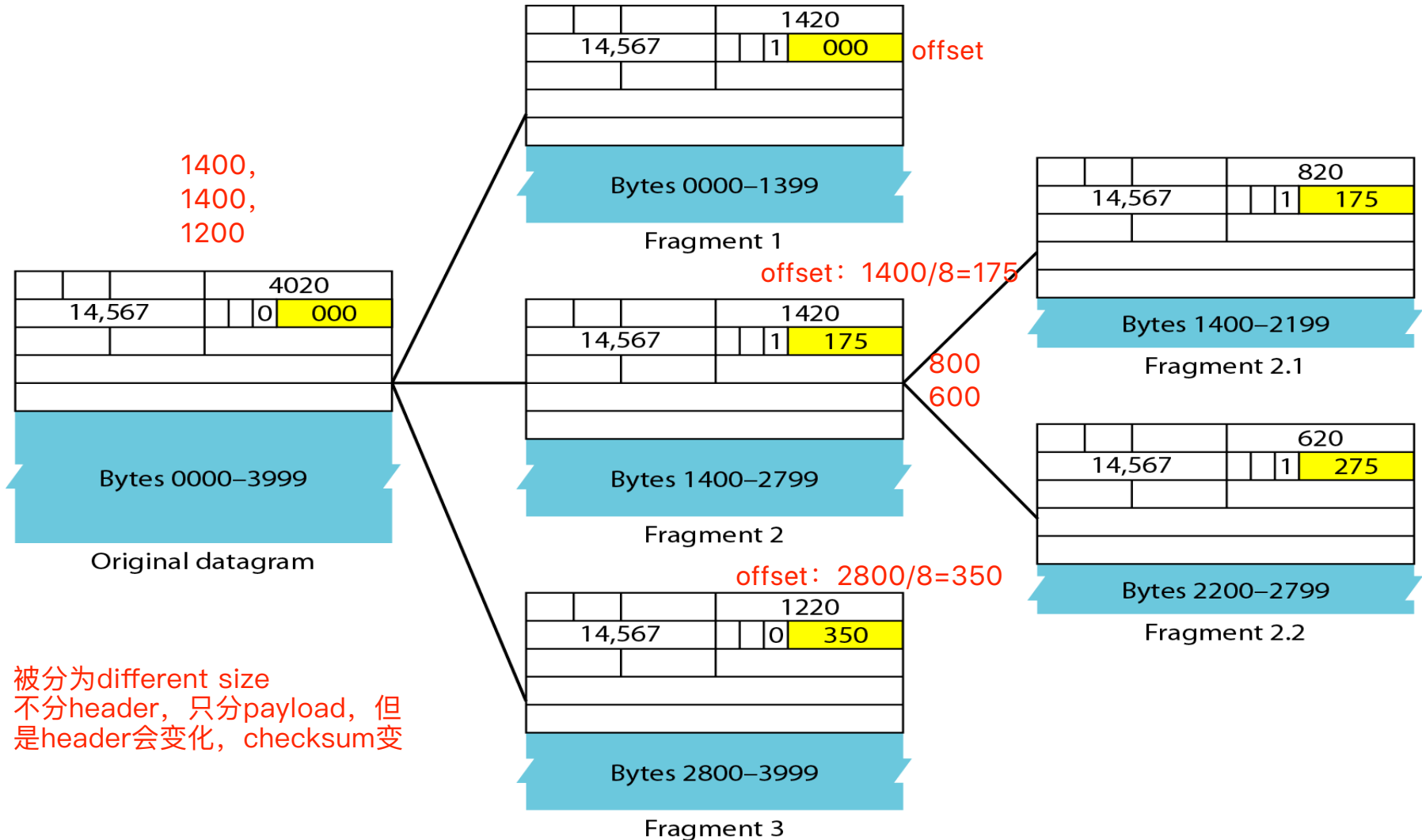
	D	M
--	---	---

  
D: Do not Fragment  
M: More Fragment

- **Fragmentation Offset (13 bits)**: Relative position of the fragment to the whole packet measured in units of 8 Bytes

被分为different size  
不分header, 只分payload, 但是  
header会变化, checksum变

# Fragmentation Example (cont.)



被分为different size  
不分header, 只分payload, 但是header会变化, checksum变



# IP v4. Addressing

---

- The Internet is made of combination of LANs and WANs connected via routers
- A host needs to be able to communicate with another host without worrying about which physical network must be passed through
- Hosts must therefore be identified **uniquely** and **globally** at the network layer
- For efficient and optimum routing, routers must also be identified uniquely and globally at the network layer

# IP v4. Addressing (Cont.)

- IPv4 address is a 32-bit address, implemented in software, is used to uniquely and globally identify a host or a router on the Internet
- A device can have more than one IP address if it is connected to more than one network  
(multi-homed)
- An IP address have two parts, the **netid** and the **hostid**. They have variable lengths depending on the class of the address
- All devices on the same network have the same netid

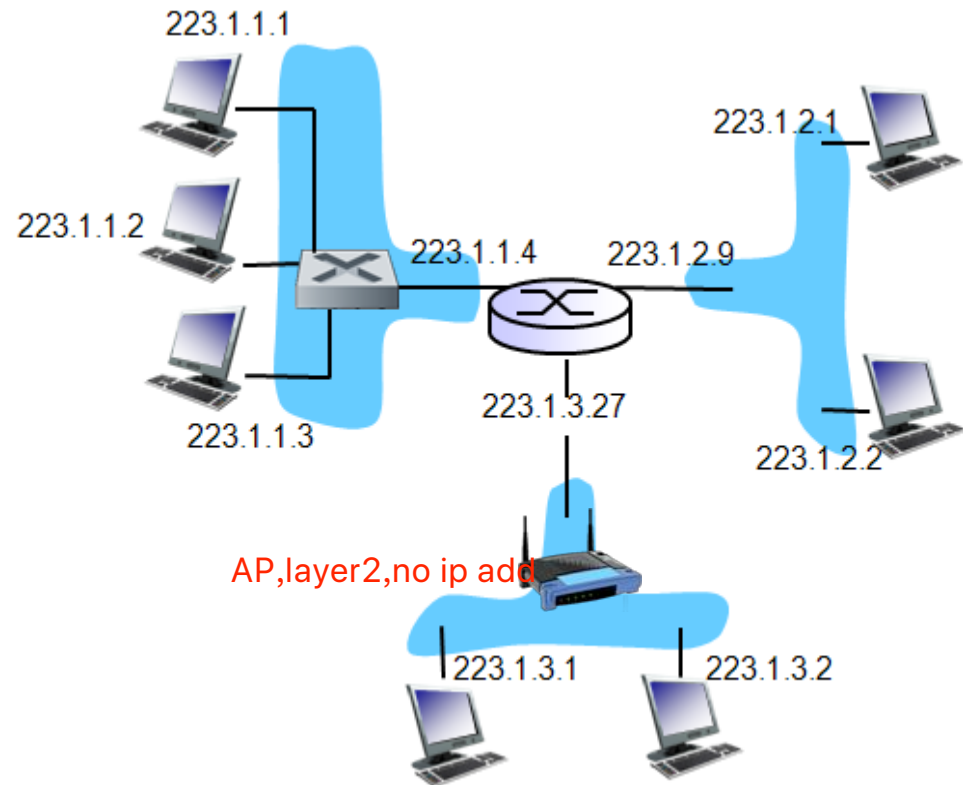
identify network of host you attached, globe  
unique assigned by ICANN

assigned by local adminiton

# IP Addresses (Example)

**A:** wired Ethernet interfaces connected by Ethernet switches

- **IP address:** 32-bit identifier for host, router interface
- **Interface:** connection between host/router and physical link
  - router's typically have multiple interfaces
  - host typically has one or two interfaces (e.g., wired Ethernet, wireless 802.11)
- IP addresses associated with each interface



**B:** wireless WiFi interfaces connected by WiFi base station

# Classful IP Addressing

need to  
memerize

	First byte	Second byte	Third byte	Fourth byte
Class A	0 net-id (7) host: $2^{24}$ 个			
Class B	10 net-id (14) host: $2^{16}$ 个			
Class C	110 net-ID (21) host: $2^8$			
Class D	1110			
Class E	1111			

a. Binary notation

	First byte	Second byte	Third byte	Fourth byte
Class A	0-127 0和127是保留的			
Class B	128-191			
Class C	192-223			
Class D	224-239			
Class E	240-255			

b. Dotted-decimal notation

# Number of Blocks/Size of Blocks in Classful IP Addressing

<i>Class</i>	<i>Number of Blocks</i>	<i>Block Size</i>	<i>Application</i>
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

**In classful addressing, a large part of the available addresses are wasted. It is hence being replaced by "Classless IP Addressing"**

# Special IP Addresses

<u>Special Address</u>	<u>Netid</u>	<u>Hostid</u>	<u>Source/Destination</u>
network itself ■ <b>Network Address</b>	75.0.0.0—>75.3.1.128 <b>Specific</b>	<b>All 0's</b>	<b>None</b>
■ <b>Direct Broadcast Address</b> when you know the address of network, 75.255.255.255	<b>Specific</b>	<b>All 1's</b>	你知道network add 的时候广播 <b>Destination</b>
■ <b>Limited Broadcast Address</b> 255.255.255.255 broadcast on this network	<b>All 1's</b>	<b>All 1's</b>	<b>Destination</b> 广播
■ <b>This host on this network</b> 0.0.0.0	<b>All 0's</b>	<b>All 0's</b>	<b>Source</b> 找DHCP的时候用
■ <b>Specific host on this network</b> 75.3.1.128—>0.3.1.128 host-id	<b>All 0's</b>	<b>Specific</b>	<b>Destination</b>
■ <b>Loopback address</b> use to test protocol in your machine, 常用: 127.0.0.1	<b>127</b>	<b>Any</b>	<b>Destination</b>

do not leave your machine, not leave your NIC

# Private IP Addressing

- One of the problems in IP network address allocation is that many hosts do not require access to hosts in other networks  $\Rightarrow$  Assigning Globally unique public IP addresses for such hosts may be wasteful
- IETF proposed the use of **Private IP addresses** that are not advertised outside the private network.

netid	Range		net broadcast	Total
10.0.0.0	to	10.255.255.255		$2^{24}$
172.16.0.0	to	172.31.255.255		$2^{20}$
192.168.0.0	to	192.168.255.255		$2^{16}$

在同一个网络里private ip是unique的, 在不同网络可以相同

Private IP Addresses  
are non-routable

only 1 network

16个network,  $2^{16}$ 个host

$2^8$ 个network,  $2^8$ 个host

# NAT: Network Address Translation

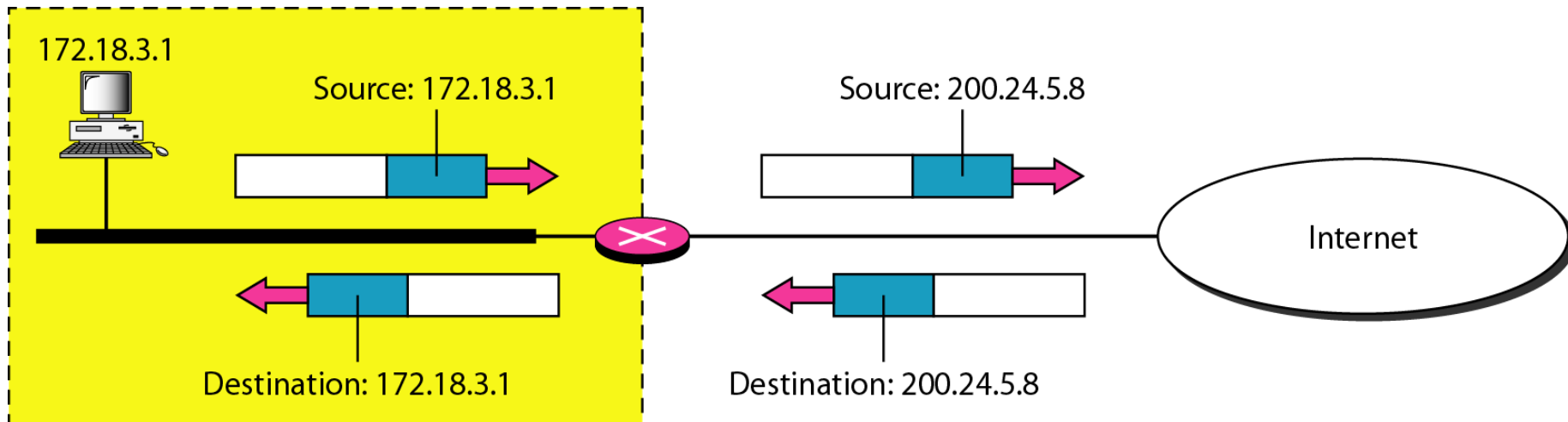
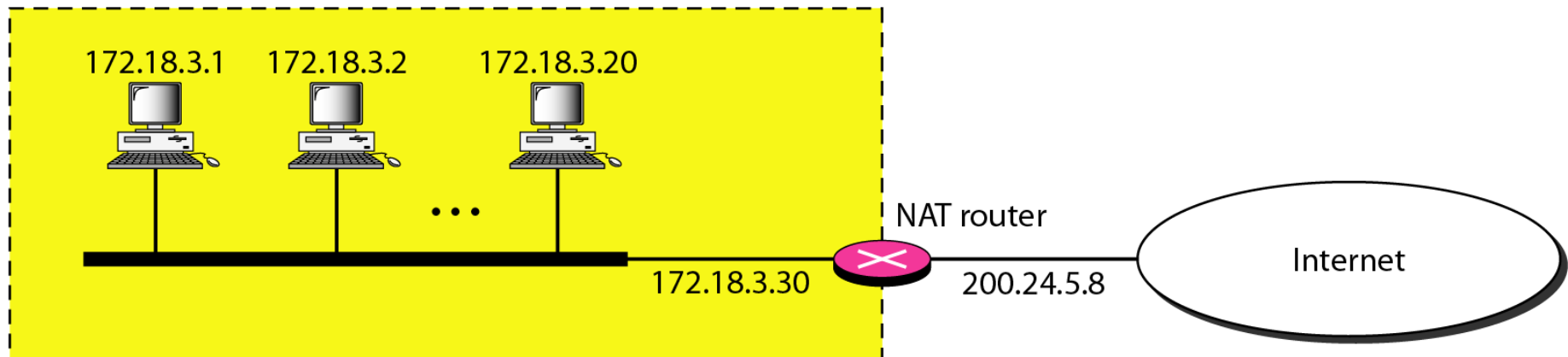
---

- NAT is a protocol that maintains a translation table for mapping an internal private IP address to a globally unique IP addresses and vice versa
- May be **static** (one-to-one) or **dynamic** (from a pool of global IP addresses)
- Implemented in the border (access) router separating the private network and the public Internet or as a stand-alone **multi-homed Server**
- NAT is a special type of Proxy Server. It was introduced with Windows 2000

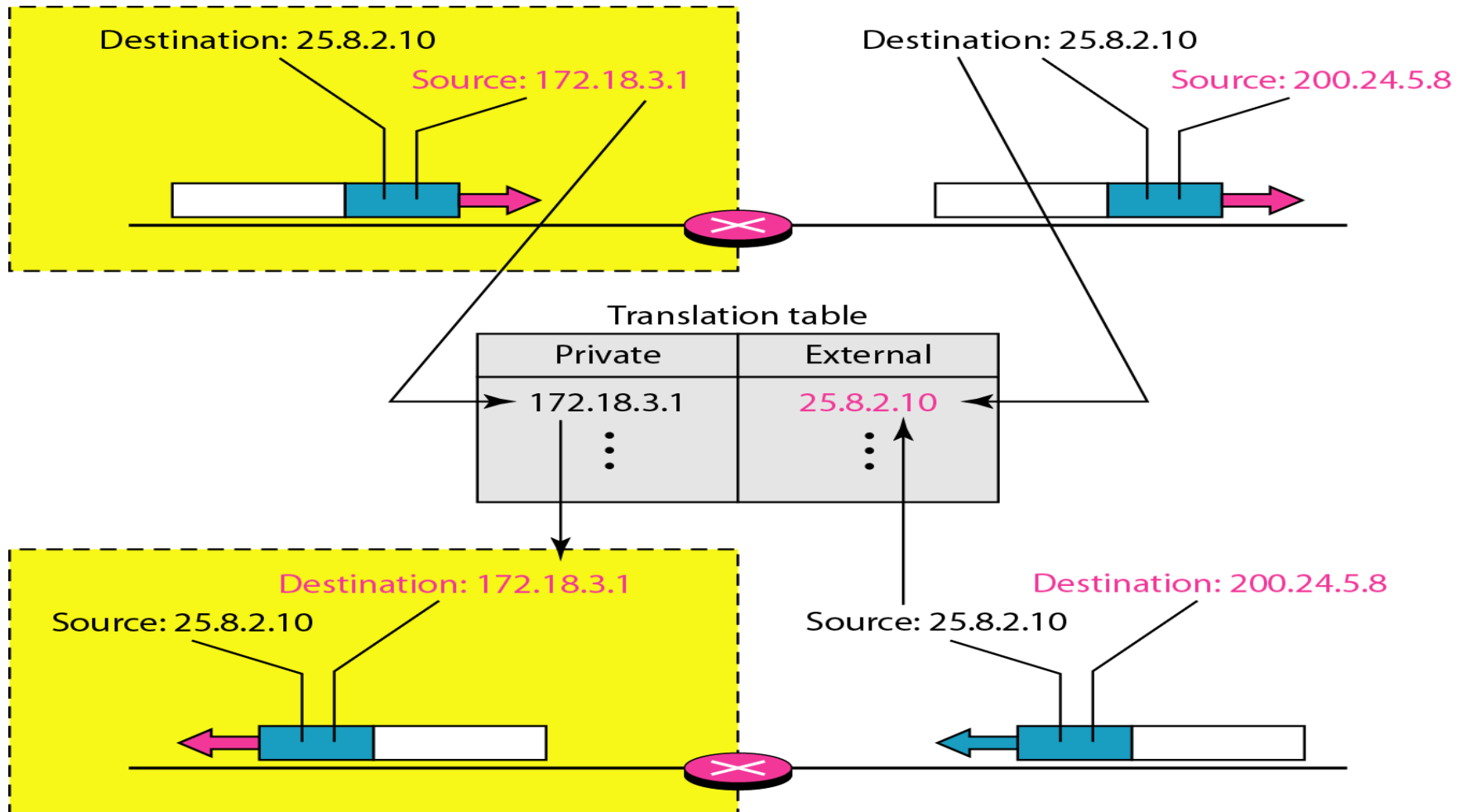


# NAT Implementation

Site using private addresses

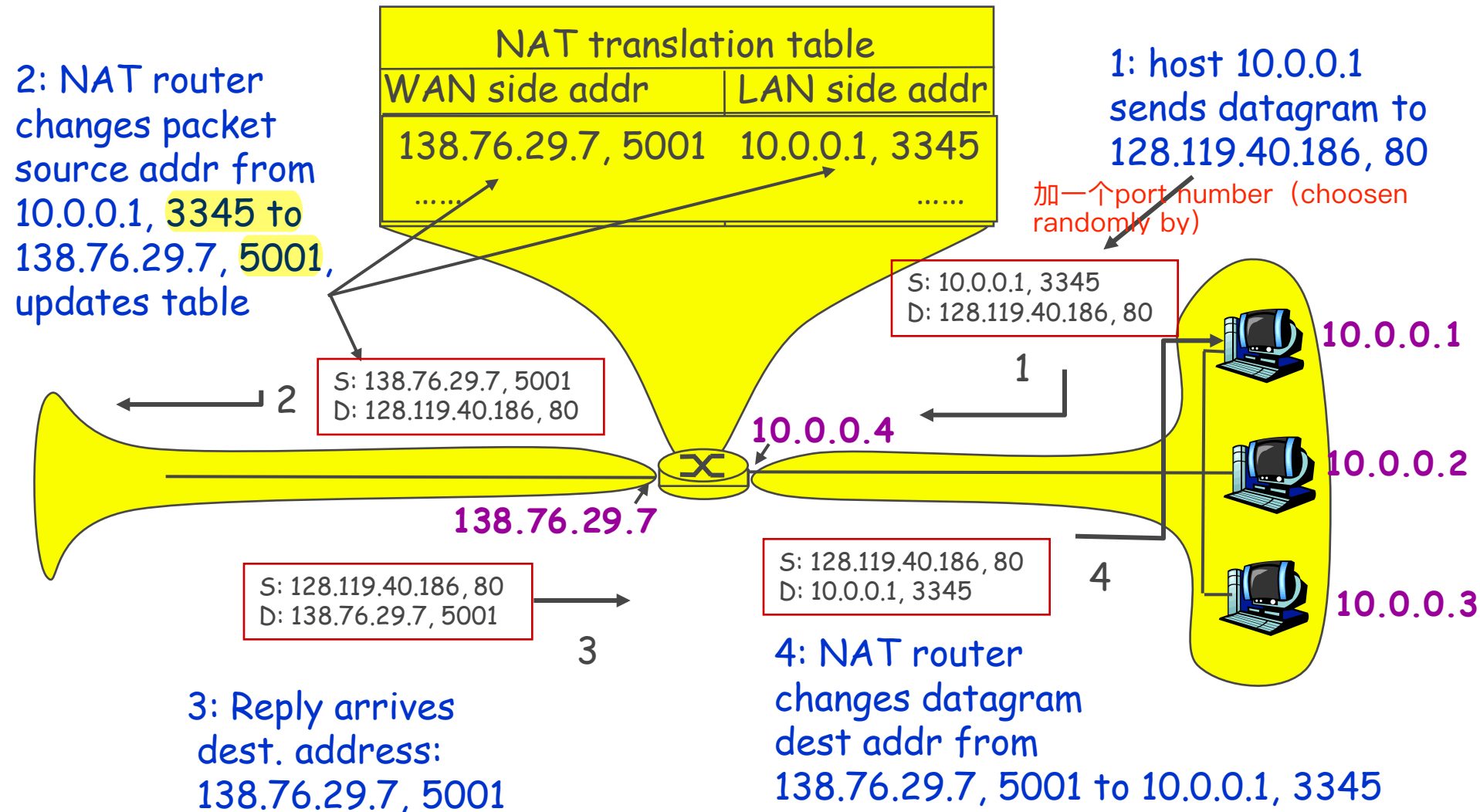


# NAT Address Translation



# NAT & PAT

NAT only 只看IP header, 不看payload  
N&P: 会看payload的port number



# Motivation for NAT

---

- 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 or visible by outside world (a security plus)

# Subnetting

- Subnetting is the process of creating multiple segments within a single IP network address space

dividing a network into smaller size network called subnet

a subnet is a network(just smaller size),  
两个subnet需要一个router

subnetting is transparent to the rest of the world
- From the perspective of a node outside the network, all nodes on any of the subnetworks appear to be on the original single network

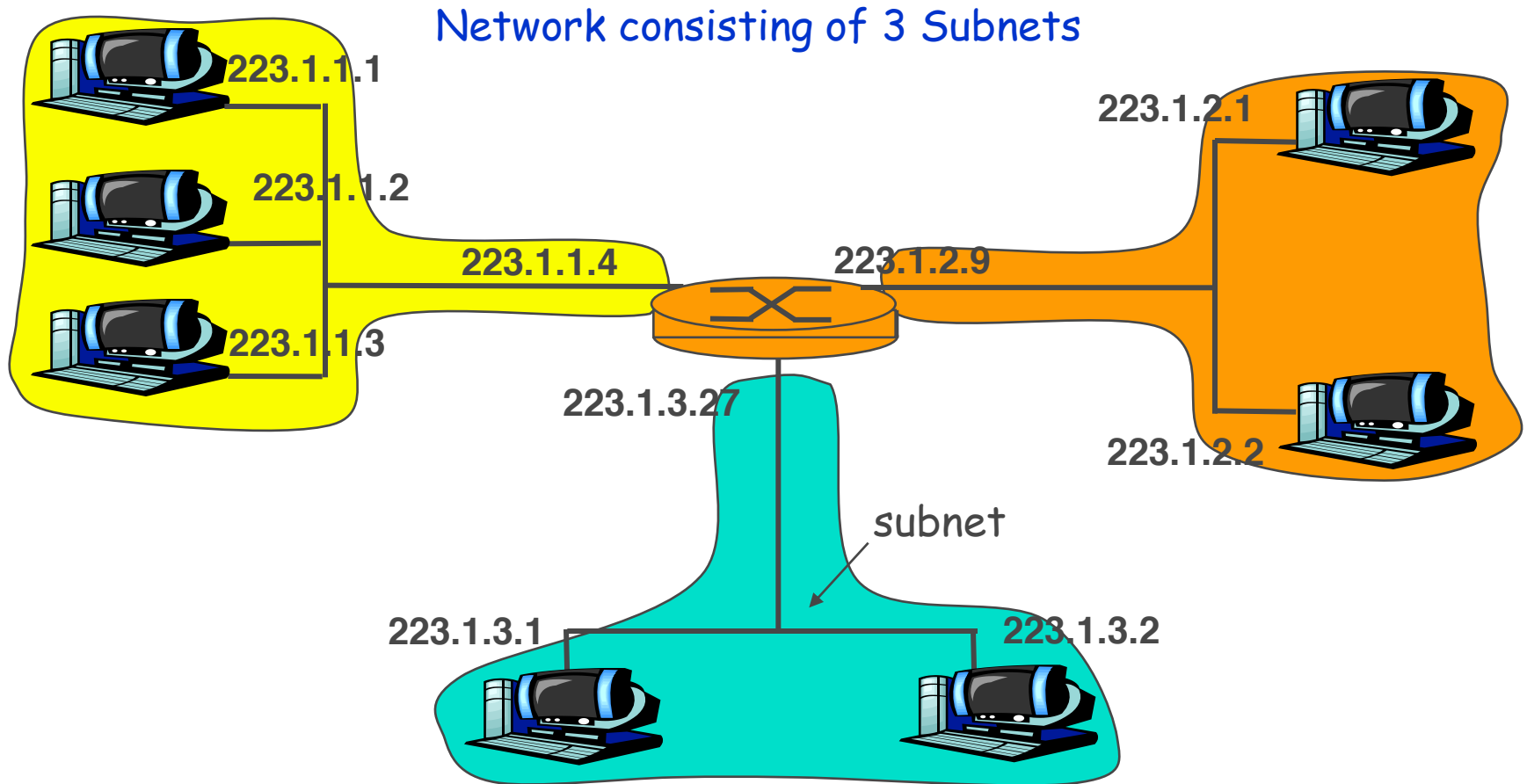
从外界看  
subnet会被作为一个整体
- Internet routing tables are not affected by subnetting, I.e. routing tables need not be overloaded with information about routes to all internal subnets, just information to the access router/gateway

in the routing table: we see the following:

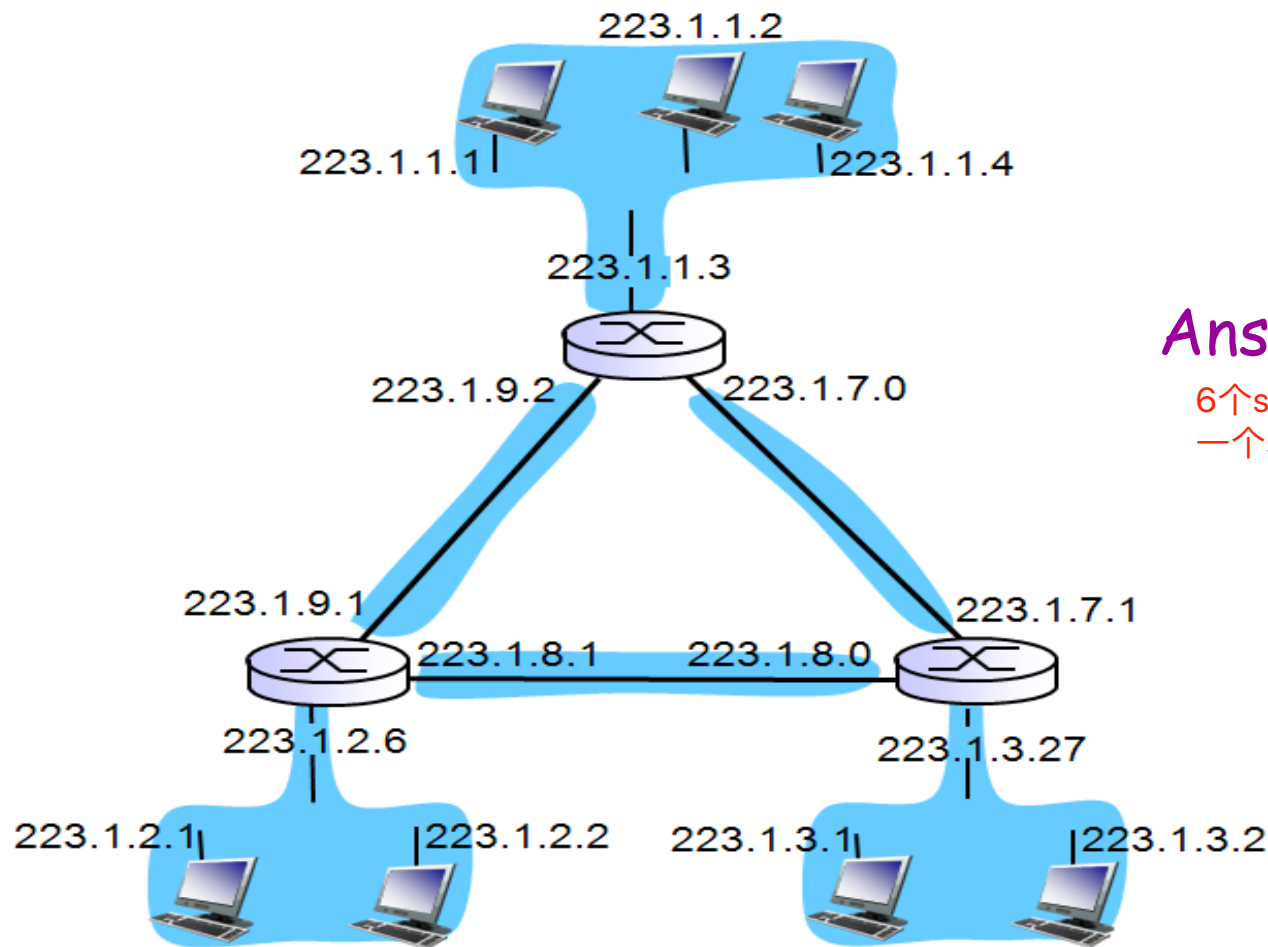
# Subnetting (Continued)

- Classes A, B and C in IP addressing are designed with two levels of hierarchy (netid & hostid)
- Problem: An organization with a class B address can not have more than network and all  $2^{16}$  hosts are attached to that network  $\Rightarrow$  A nightmare in managing this network, single broadcast domain, security issues, etc...
- Subnetting create another level of hierarchy (netid, subnetid and hostid). Delivery of IP packets involves three steps; delivery to the site router, delivery to the subnet router, delivery to the host

# Example of Subnetting



# Example of Subnetting (How many?)



**Answer:**

6个subnet, 每个蓝色的区域是一个subnet



all devices in the same subnet must have same subnet mask;  
devices on different subnet may have same subnet mask;  
if no subnet:  
A(255.0.0.0)  
B(255.255.0.0)

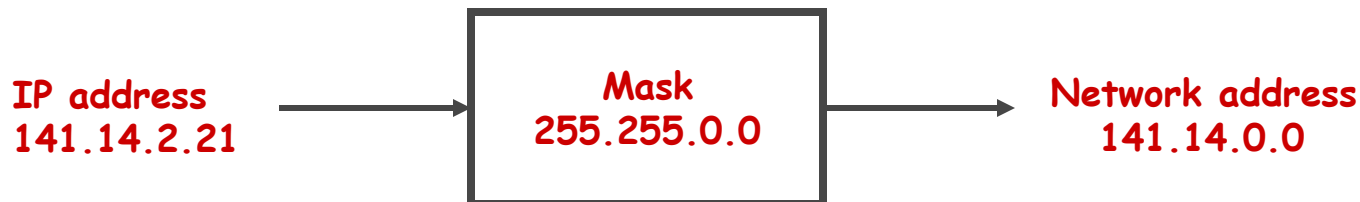
# Subnet Masking

- Subnetting is achieved by "stealing" some bits from the hostid field to represent the subnet portion of the address
- Those bits that are used for the subnetid are identified through the use of a **subnet mask**
- Masking is the process of extracting the address of the physical network (if subnetting is not used) or the subnet address (if subnetting is used) from an IP address  
Subnet mask相同不一定在同一个subnet  
not an address
- A subnet mask is a 32-bit pattern having a "1" in every netid and subnetid locations and a "0" in every hostid location

# Subnet Masking (Continued)

- Subnet masking is performed (both at the host and at the router) by applying "bit-wise-AND" operation between the IP address and the subnet mask  
最后结果: subnet address

- Example 1: Class B network without subnetting
  - 141.14.2.21      10001101.00001110.00000010.00010101
  - 255.255.0.0      11111111.11111111.00000000.00000000
  - "Bit-wise and"   10001101.00001110.00000000.00000000



class B default mask:  
255.255.0.0

# Example of Subnetting

3个subnet

9 subnet id

Mask: 255.255.255.128  
Subnet #: 128.96.34.0

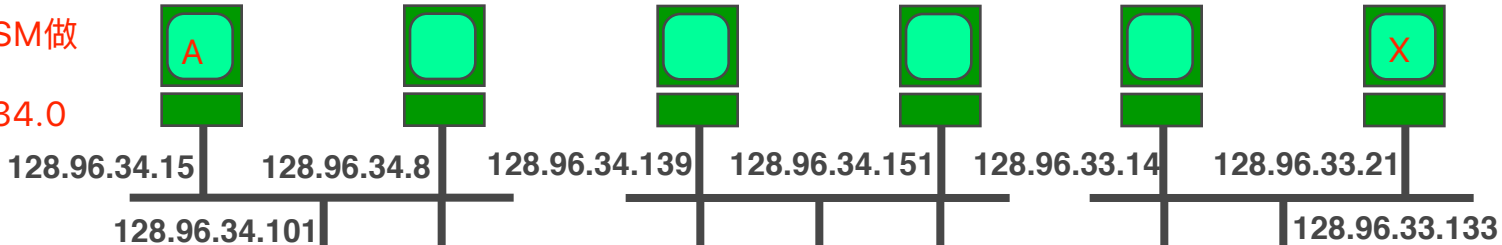
9 subnet id

Mask: 255.255.255.128  
Subnet #: 128.96.34.128

8 subnet id

Mask: 255.255.255.0  
Subnet #: 128.96.33.0

A的IP与SM做  
与:  
128.96.34.0



如果A想send something to X:  
1. X的IP AND A的SM 128.76.33.128 (represent nothing)  
2. 与本subnet id不一样, 所以是another subnet  
3. need R1 MAC address (ARP)  
4. R1 get the packet, 然后用X的IP AND 他每个interface的SM, 都不一样, 需要R2的MAC  
5. R2 get the packet, 同样运算

Internal routers  
use subnet IDs  
to route

External routers do not  
know about subnet IDs,  
only net IDs

# Subnet Router Routing Table

Partial Table in Router 128.96.33.1

Subnet Number	Subnet Mask	Next Hop
128.96.34.0	255.255.255.128	Left Router
128.96.34.128	255.255.255.128	Left Interface
128.96.33.0	255.255.255.0	Right Interface

IP address of R1

```
for each table entry do
  if (DestAddr & Subnet Mask) = SubnetNumber
    if NextHop is an Interface
      deliver packet directly to DestAddr
    else
      deliver packet to Router
```

a block can be further divided into subblocks  
it is transparent to the rest of the world,  
smaller block can be aggregated in to bigger blocks (sometimes known as supernetting)

# Classless IP Addressing

- To overcome IP Address depletion and allow more organizations ability to access the Internet  
# bits of host:  $32-n$
- Get rid of classes and assign addresses in "Blocks". The size of the block (The # of IP addresses in the block) depends on the organization demand  
power of 2:  $2^{(32-n)}$  /18 is larger than /22
- There are some restrictions on these blocks

# Restrictions of Address Blocks

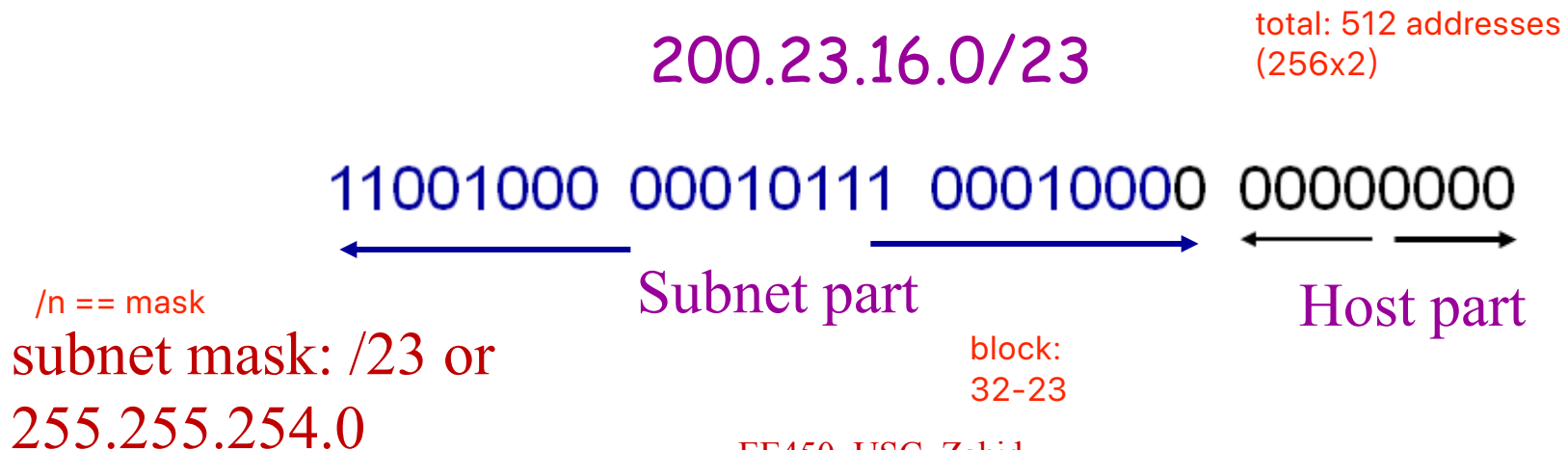
---

- The number of addresses in a block must be a power of 2 (1, 2, 4, 8, etc...)
- The addresses in the block are contiguous, i.e. one after the other<sup>no gap</sup>
- The first address in the block (in decimal) must be divisible by the size of the block

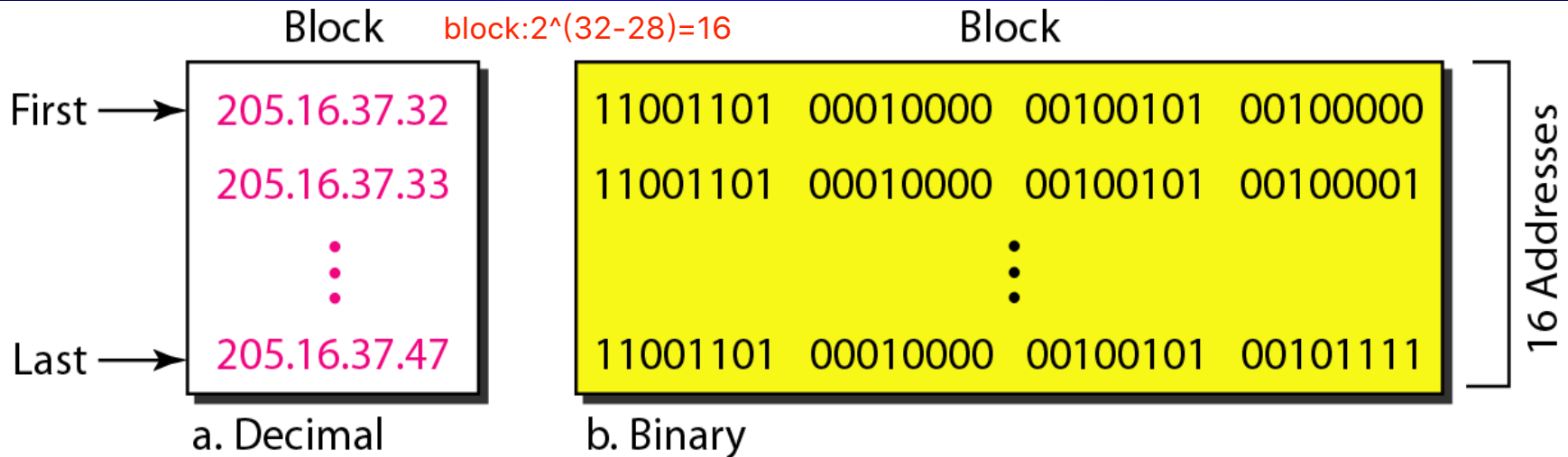
# CIDR Notation

## CIDR: Classless Inter Domain Routing

- Network and Subnet portions of address of arbitrary length
- address format:  $a.b.c.d/x$ , where  $x$  is # bits in network and subnet portion of address  
known as network prefix



# Example: A block of Size 16



In IPv4 addressing, a block of addresses can be defined as  $w.x.y.z /n$  in which  $w.x.y.z$  defines one of the addresses and the  $/n$  defines the mask. The number of address in the block is  $2^{32-n}$

The first address in the block can be found by setting the rightmost  $32 - n$  bits to 0s. Hence the first address in the block represents the Network/Subnet address of the organization

The last address in the block can be found by setting the rightmost  $32 - n$  bits to 1s. Hence the last address in the block represents the broadcast address in the organization network



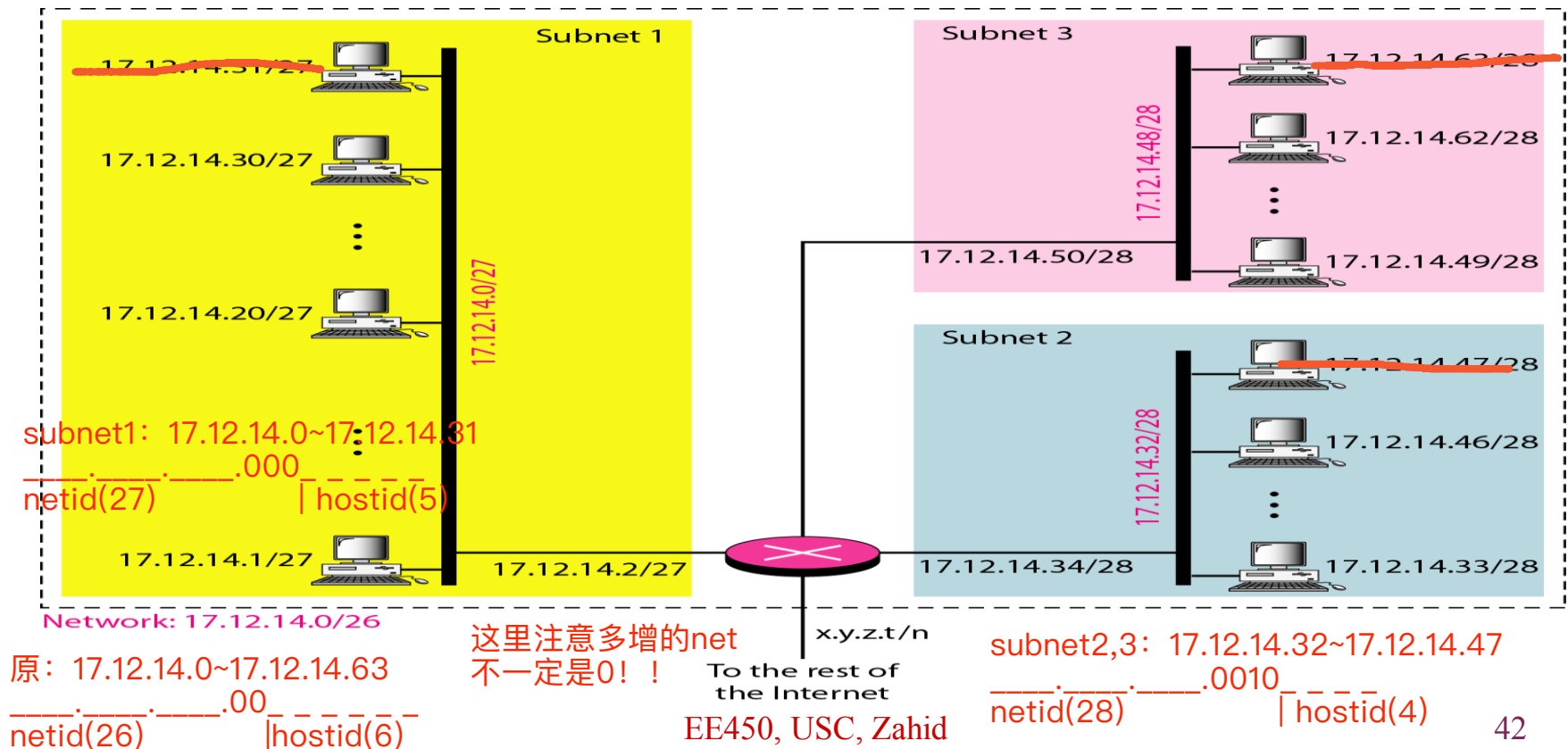
# Example of Classless IP Addressing

- An ISP gets a block of IP addresses. The Block is 200.23.16.0/20 . The size of the block is  $2^{12} = 4096$ . The ISP has 8 customers (organizations), each requiring a "block" of size 512

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

# Example of Classless IP Addressing

- Organization was assigned a block 17.12.40.0/26 classless  
(Size of block is  $2^6$  64). Organization has three departments. Three subnets of sizes 32, 16 and 16 addressable devices: 30, 14, 14



# Hierarchical Addressing

Hierarchical addressing allows efficient advertisement of routing information

