Inter-networking

COS 460 & 540

CAPTAIN JAMES T. KIRK
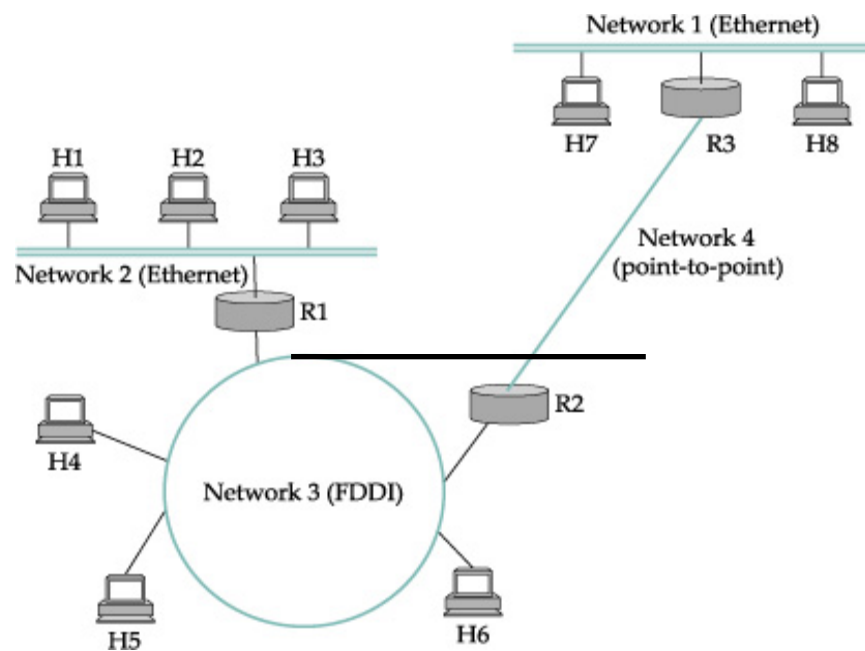
I'M SORRY, I CAN'T HEAR YOU OVER THE
SOUND OF HOW AWESOME I AM.

Problem

LAN's are great… but…

We want to **connect** them together

...across the world

Inter-networking

- Internet Protocol (IP)

- Routing

- The Internet

- Multicast*

- Multi-protocol Label Switching*

Internet Protocol (IP)

- What is an Internetwork

- Service Model

  - Datagrams, Packet Format, …

- Addressing

- Datagram Forwarding

- ARP, DHCP, ICMP, ...

What?

- "internet" vs "Internet"

- "network" vs "subnetwork"

- physical vs logical networks

Bridge, Router, Switch

- repeater & hub - physical

- bridge - link

- switch - network

- router - internet

- gateway - application

Alternatives?

- Novell's IPX

Internet Protocol (IP)

✓ What is an Internetwork

- Service Model

  - Datagrams, Packet Format, …

- Addressing

- Datagram Forwarding

- ARP, DHCP, ICMP, …

Service Model

- Datagram Delivery

- Packet Format

- Fragmentation and Reassembly
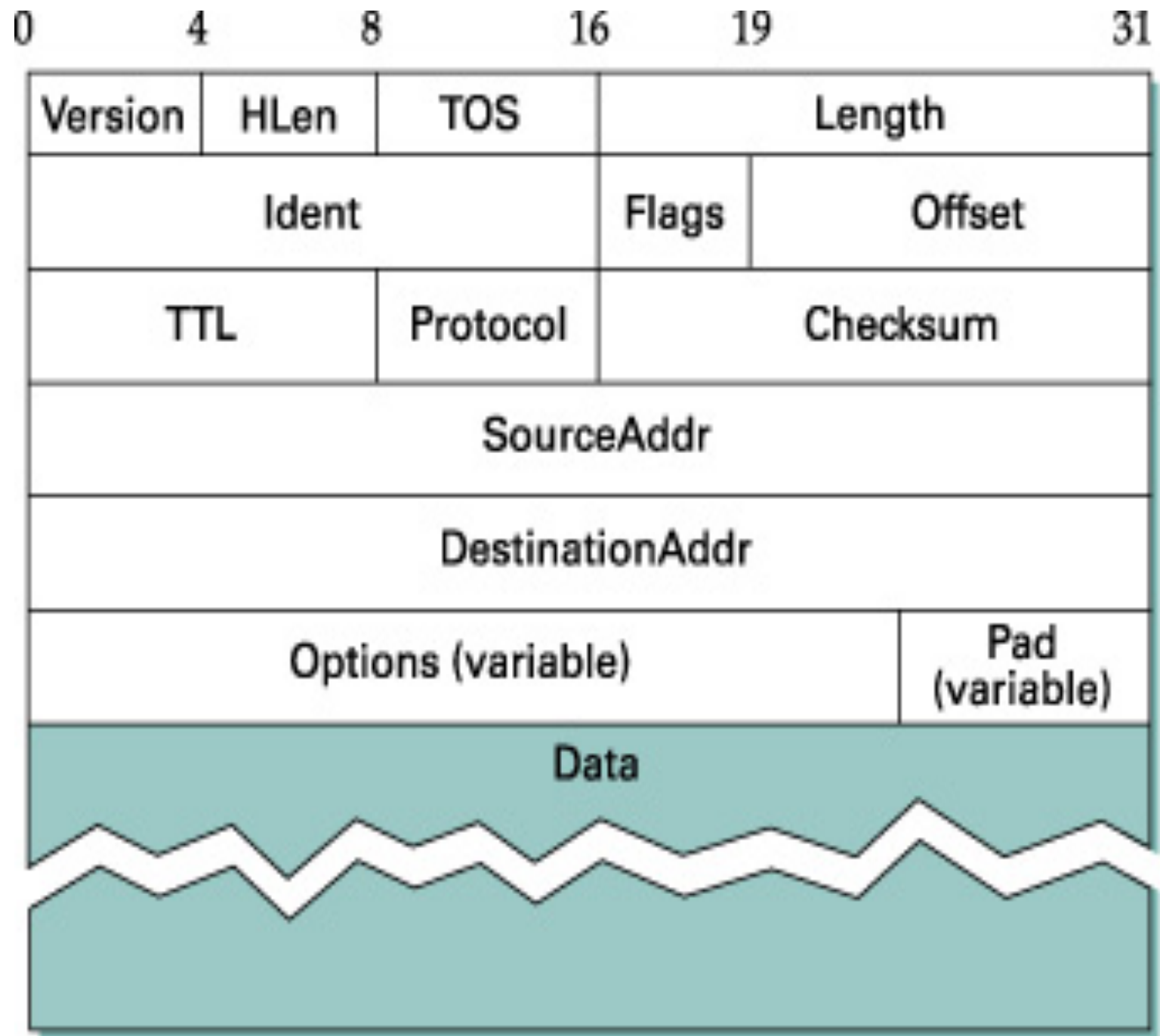
Datagram Delivery

Underlying network is..

- reliable or unreliable

- connection-oriented or connectionless

- small or large packet/frame sized

- physical, logical, wireless, ...

Datagram Delivery

- IP **Datagram** is basis of protocol

- Provides a "**best-effort**" or unreliable service

- May be **out of order**

- **Connectionless**

Packet Format

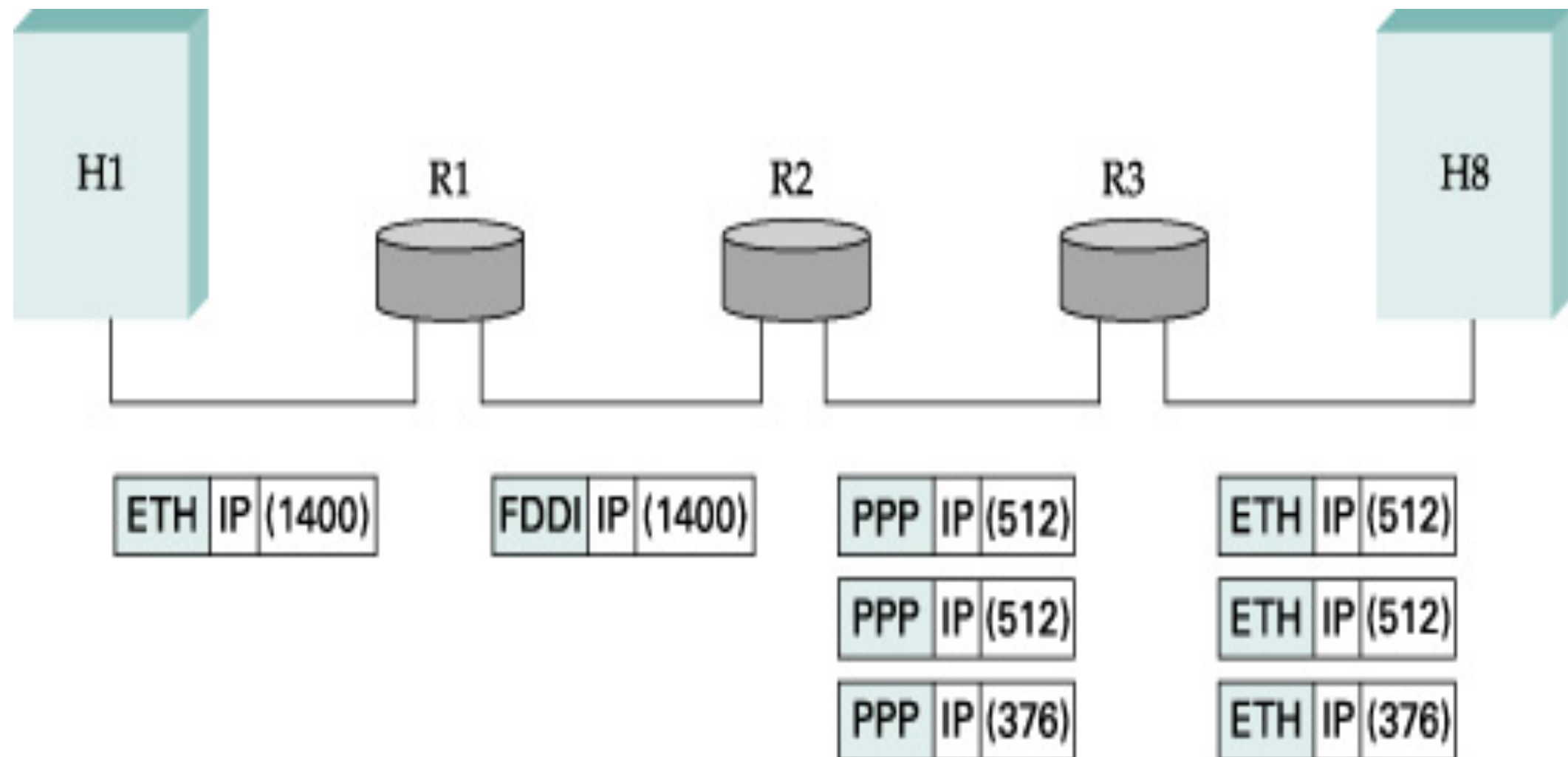| 0	4	8	16	19	31
Version	HLen	TOS	Length		
Ident			Flags	Offset	
TTL		Protocol	Checksum		
SourceAddr					
DestinationAddr					
Options (variable)				Pad (variable)	
Data					

Fragmentation

Problem:

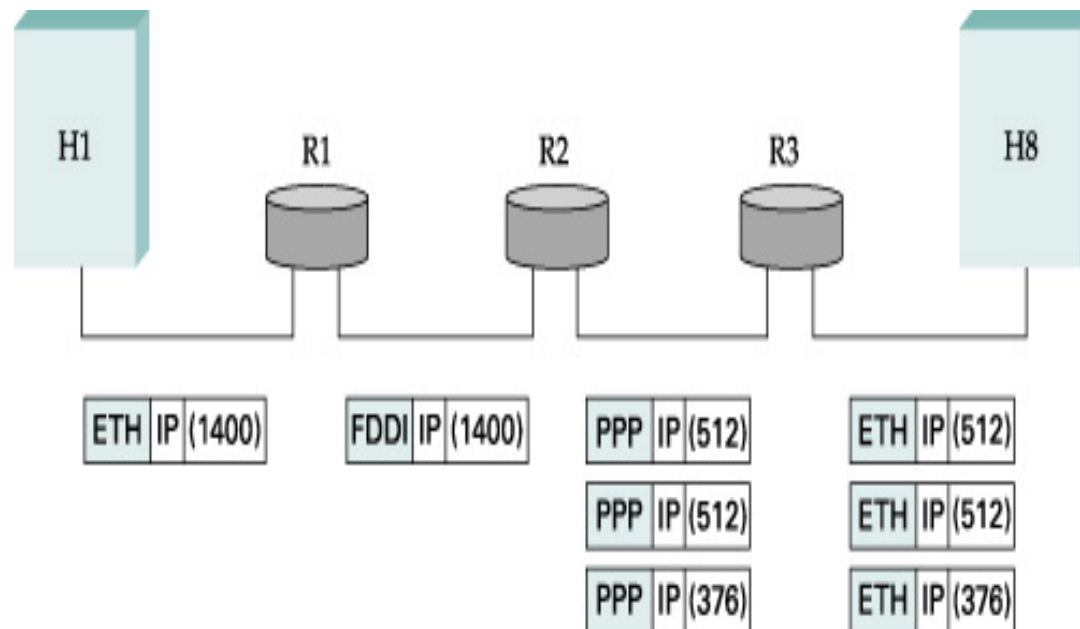- IP datagrams can be 64 kB

- The underlying network may have 512 Byte MTU

Solution:

- Each fragment is an IP datagram

- Each IP datagram is rebuilt for each network

Fragmentation

Fragmentation

H1   R1   R2   R3   H8

| ETH | IP | (1400) |
| FDDI | IP | (1400) |

PPP	IP	(512)		ETH	IP	(512)
PPP	IP	(512)		ETH	IP	(512)
PPP	IP	(376)		ETH	IP	(376)

(a)

| Start of header |
| Ident = x | | | | 0 | Offset = 0 |
| Rest of header |
| 1400 data bytes |

(b)

| Start of header |
| Ident = x | | | | 1 | Offset = 0 |
| Rest of header |
| 512 data bytes |

| Start of header |
| Ident = x | | | | 1 | Offset = 64 |
| Rest of header |
| 512 data bytes |

| Start of header |
| Ident = x | | | | 0 | Offset = 128 |
| Rest of header |
| 376 data bytes |

Internet Protocol (IP)

✓ What is an Internetwork

✓ Service Model

   ✓ Datagrams, Packet Format, …

- Addressing

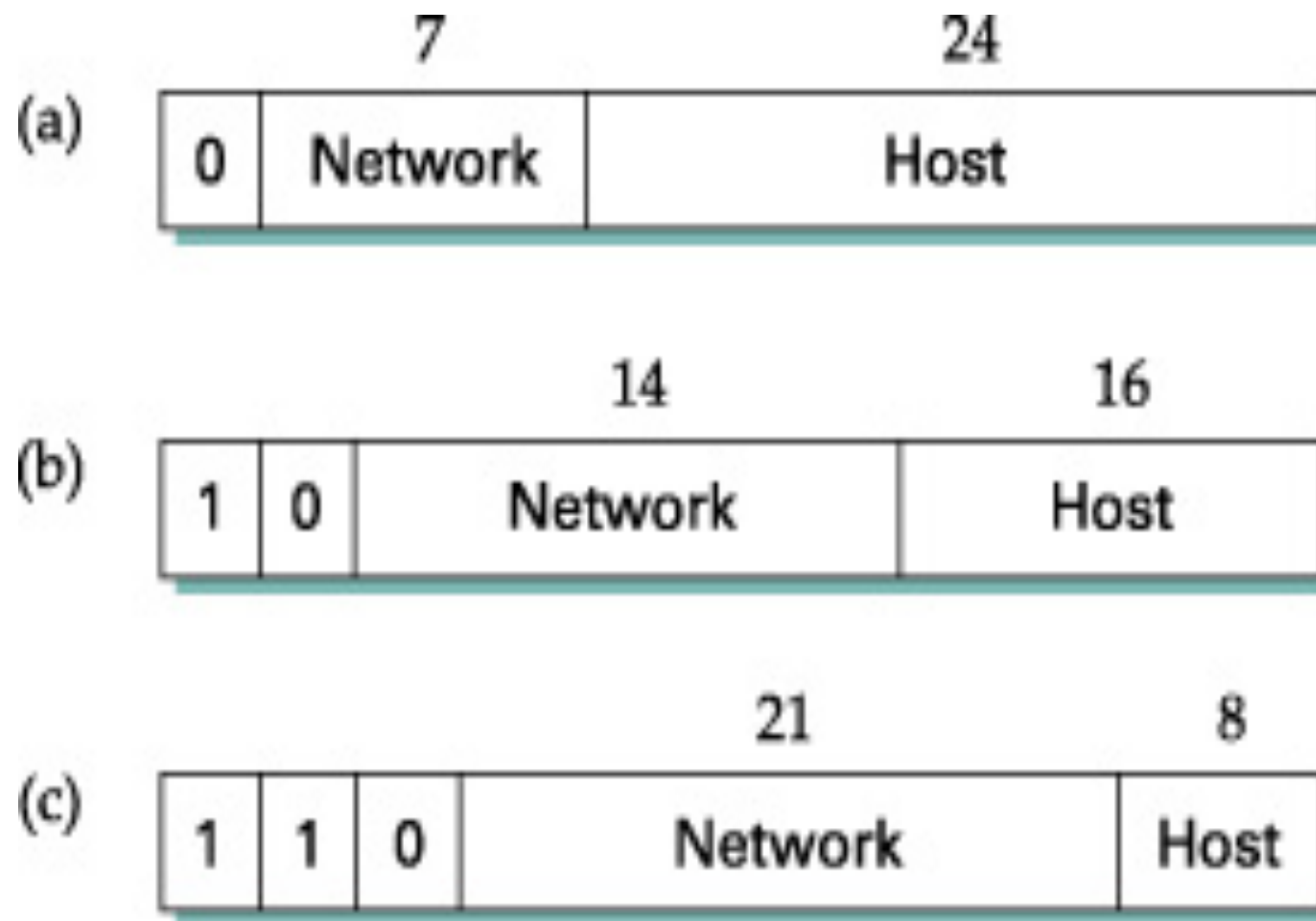- Datagram Forwarding

- ARP, DHCP, ICMP, ...

IP Addressing

- Need to identify any host on the network

- Globally **unique**

- **Hierarchal** (not flat like Ethernet)

  - for easier routing

  - represents a network of networks

IP Addressing

- 32 bit addresses

- "Dotted Decimal" format

  - 130.111.135.26

- Contains two parts

  - **Network**

  - **Host**

IP Addressing

IP Addressing

- Network portion

  - specifies a unique physical* network

  - used for routing

- Host portion

  - specifies a unique host on the network

  - local delivery

Internet Protocol (IP)

✓ What is an Internetwork

✓ Service Model

   ✓ Datagrams, Packet Format, …

✓ Addressing

• Datagram Forwarding

• ARP, DHCP, ICMP, ...

Datagram Forwarding

How do datagrams traverse the internetwork

Forwarding vs Routing

- ***forwarding*** is taking an input packet and sending it out the appropriate port

- ***routing*** is the process of building forwarding tables.

Datagram Forwarding

- Every datagram has **destination** IP

- Network part **uniquely identifies** a physical network

  - All **hosts/routers** on network can **communicate** with all others
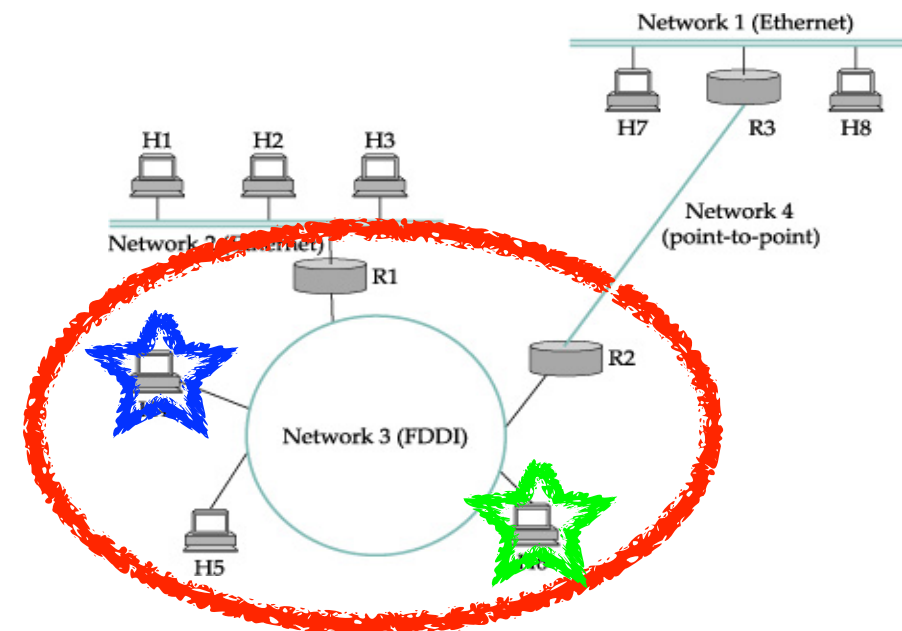
- Every network has a **router** on the net

Datagram Forwarding

Each host and router maintains a forwarding table

Network	Next-hop
1 130.111	R1 130.111.32.1
2 141.114	R2 141.114.1.1
default	R2 76.5.4.3

Datagram Forwarding
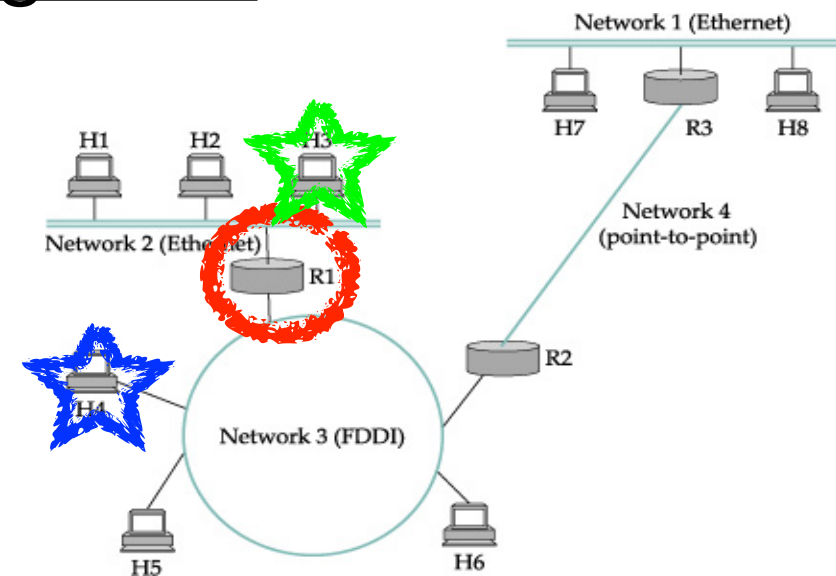
If destination network == source network

deliver locally

Datagram Forwarding

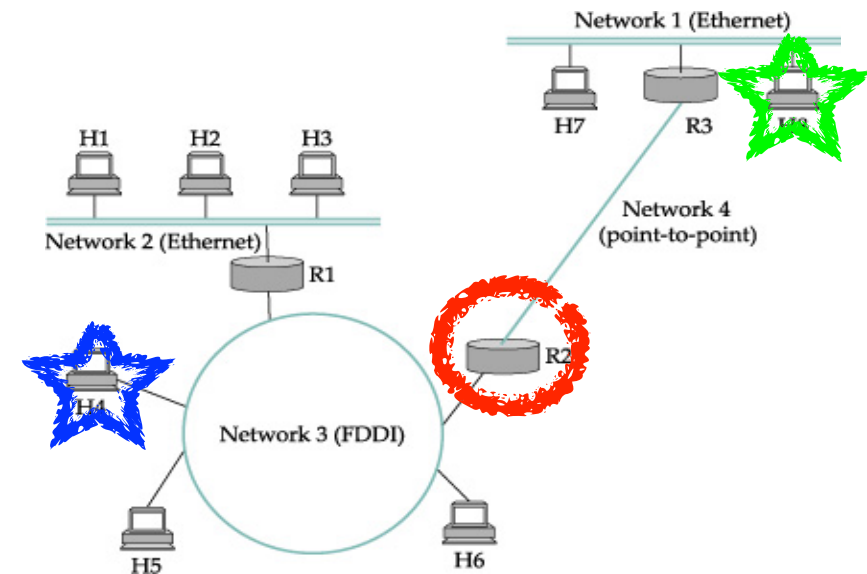if destination network in <u>forwarding table</u>

deliver to next-hop

Datagram Forwarding

if destination network <u>unknown</u>

deliver to default

Datagram Forwarding

- Local delivery for local datagrams

- Routers forward datagrams towards the physical network

Internet Protocol (IP)

✓ What is an Internetwork

✓ Service Model

   ✓ Datagrams, Packet Format, …

✓ Addressing

✓ Datagram Forwarding

• ARP, DHCP, ICMP, …

Address Resolution Protocol (ARP)

Local delivery

it's not that simple

IP Address != Ethernet Address

Address Resolution

- Finds **local** or **link-level** address for an **IP** address

    - Both hosts on the same IP Network

    - Discover / Dynamic

- Uses *broadcast* feature of link-level

Address Resolution

1. Do we have the IP-MAC addresses cached?

2. Send out broadcast query
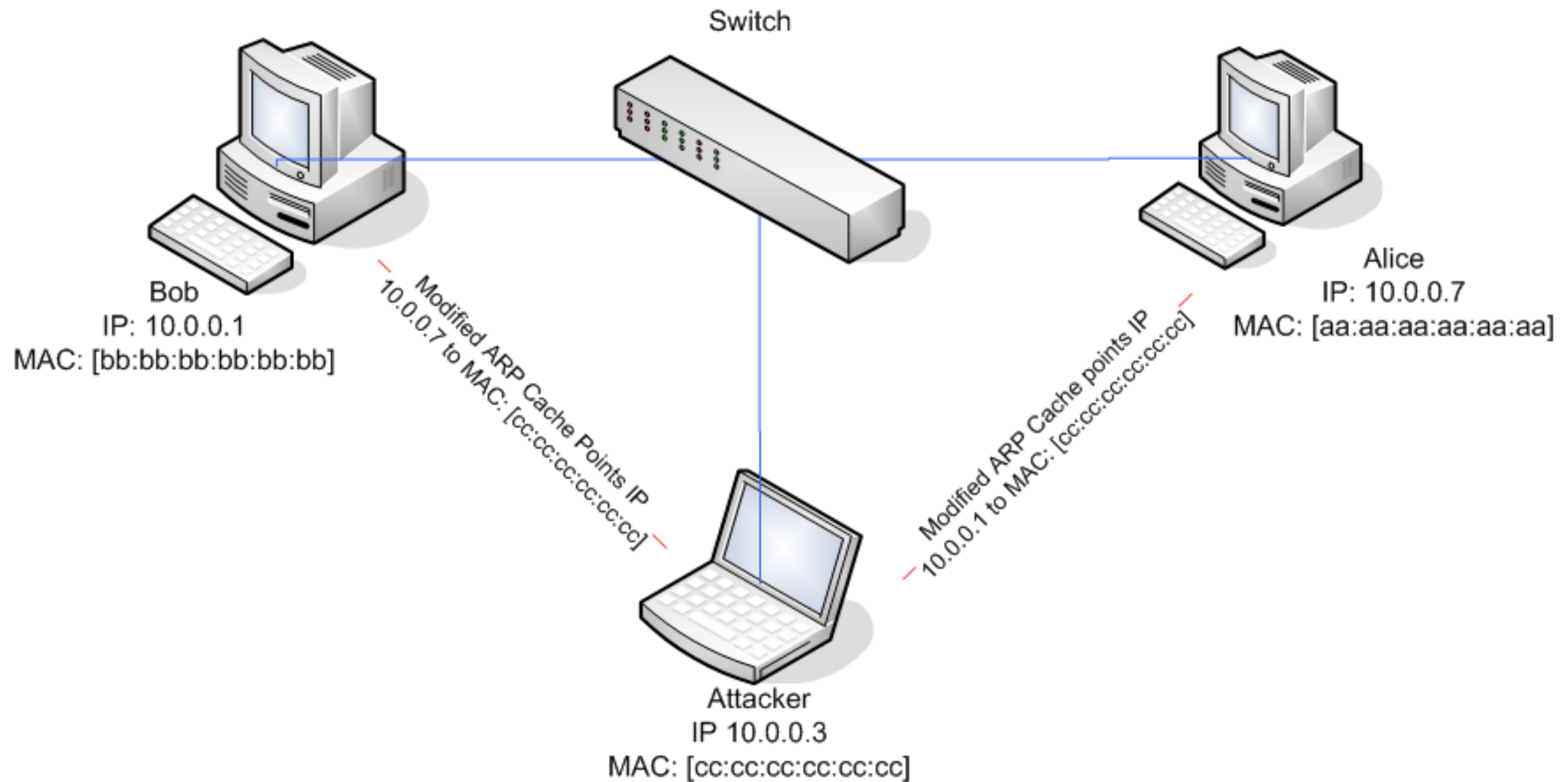
3. Look for response, and fill in cache

ARP Packet

0	8	16	31
Hardware type = 1		ProtocolType = 0x0800	
HLen = 48	PLen = 32	Operation	
SourceHardwareAddr (bytes 0–3)			
SourceHardwareAddr (bytes 4–5)		SourceProtocolAddr (bytes 0–1)	
SourceProtocolAddr (bytes 2–3)		TargetHardwareAddr (bytes 0–1)	
TargetHardwareAddr (bytes 2–5)			
TargetProtocolAddr (bytes 0–3)			

Request or Response

ARP

- Problems?

  - ARP Spoofing
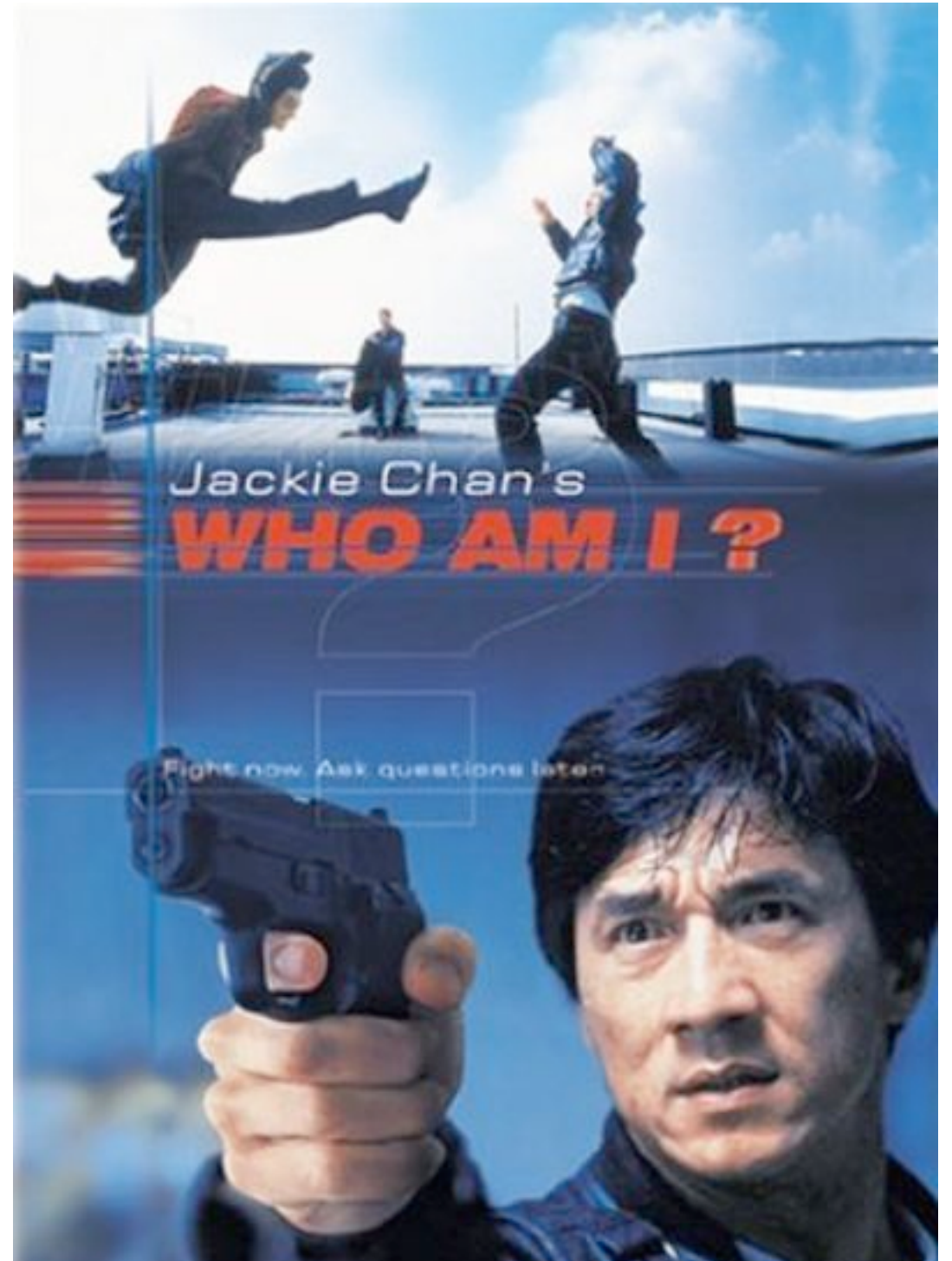   (http://en.wikipedia.org/wiki/ARP_spoofing)

ARP Spoofing

Switch

Bob
IP: 10.0.0.1
MAC: [bb:bb:bb:bb:bb:bb]

Modified ARP Cache Points IP
10.0.0.7 to MAC: [cc:cc:cc:cc:cc:cc]

Alice
IP: 10.0.0.7
MAC: [aa:aa:aa:aa:aa:aa]

Modified ARP Cache points IP
10.0.0.1 to MAC: [cc:cc:cc:cc:cc:cc]

Attacker
IP 10.0.0.3
MAC: [cc:cc:cc:cc:cc:cc]

DHCP

Dynamic Host Configuration Protocol

Who Am I?

DHCP

- Static configuration of hosts

  - not flexible or adaptable to changes

  - cumbersome

- Dynamic configuration

  - provide host with an IP address

  - additional information: router, boot info...

DHCP

1. Broadcast request: DHCPDISCOVER

2. Look for responses: DHCPOFFERs

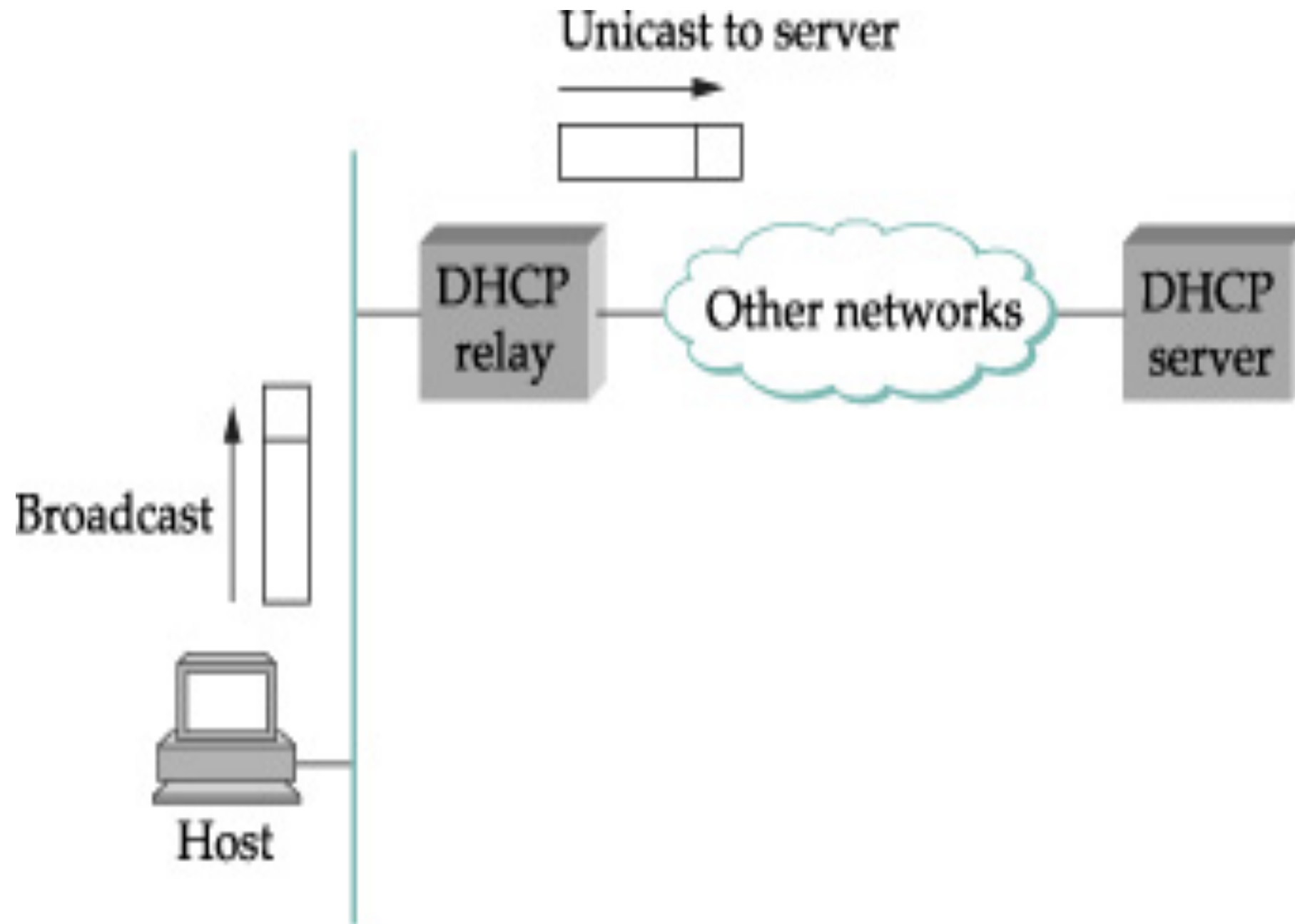3. Pick one and DHCPREQUEST

4. Wait for DHCPACK

DHCP

Problem:
Broadcast means you need a DHCP server on every network.

Solution:

DHCP forwarding

DHCP Forwarding

DHCP

- IP Address is **leased** for a given amount of time.

- Host must **renew** the lease with the server.

- Server can deny renewal

  - Can request a new lease

ICMP

When things don't go as planned.

ICMP

Remember: Best-effort service

- Diagnostic purposes (ping)

- TTL (hops) reaches 0 at a router

- Host not reachable (network error)

- Network redirection

ARP, DHCP, ICMP

These protocols are at the edge between the Network and IP layers

- Address Resolution Protocol (ARP)

- Dynamic Host Configuration Protocol (DHCP)

- Internet Control Message Protocol (ICMP)

Internet Protocol (IP)

✓ What is an Internetwork

✓ Service Model

   ✓ Datagrams, Packet Format, …

✓ Addressing

✓ Datagram Forwarding

✓ ARP, DHCP, ICMP, …

Internet Protocol (IP)

✓What is an Internetwork

✓Service Model

 ✓Datagrams, Packet Format, …

✓Addressing

✓Datagram Forwarding

✓ARP, DHCP, ICMP, …

Virtual Networks (VPN)
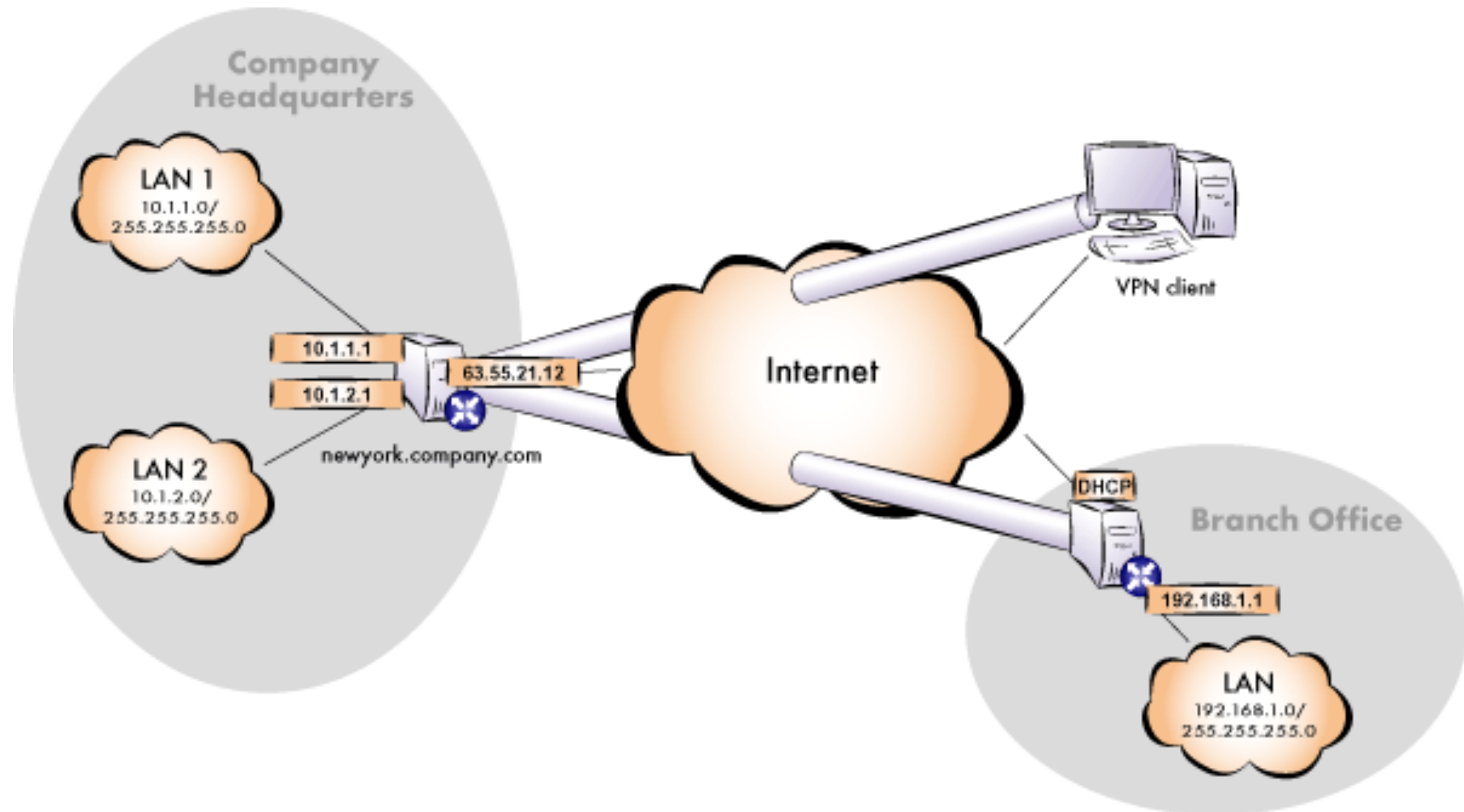
VPN
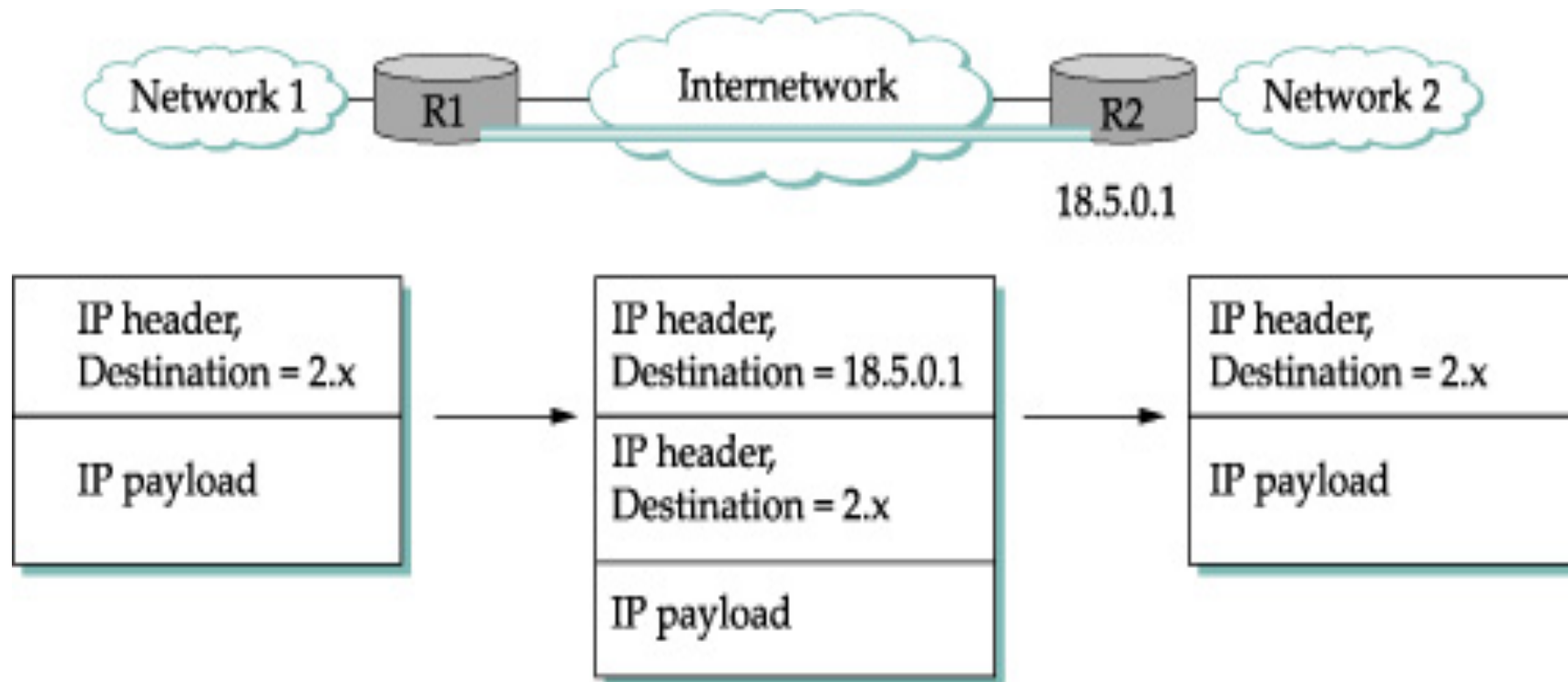
Inter-networking

✓Internet Protocol (IP)

★Routing

- The Internet

- Multicast*

- Multi-protocol Label Switching*

Routing

- Network as a graph

- Distance Vector (RIP)

- Link State

  - Open Shortest Path First (OSPF)

- Mobile Routing

Forwarding vs Routing

- *forwarding* is taking an input packet and sending it out the appropriate port

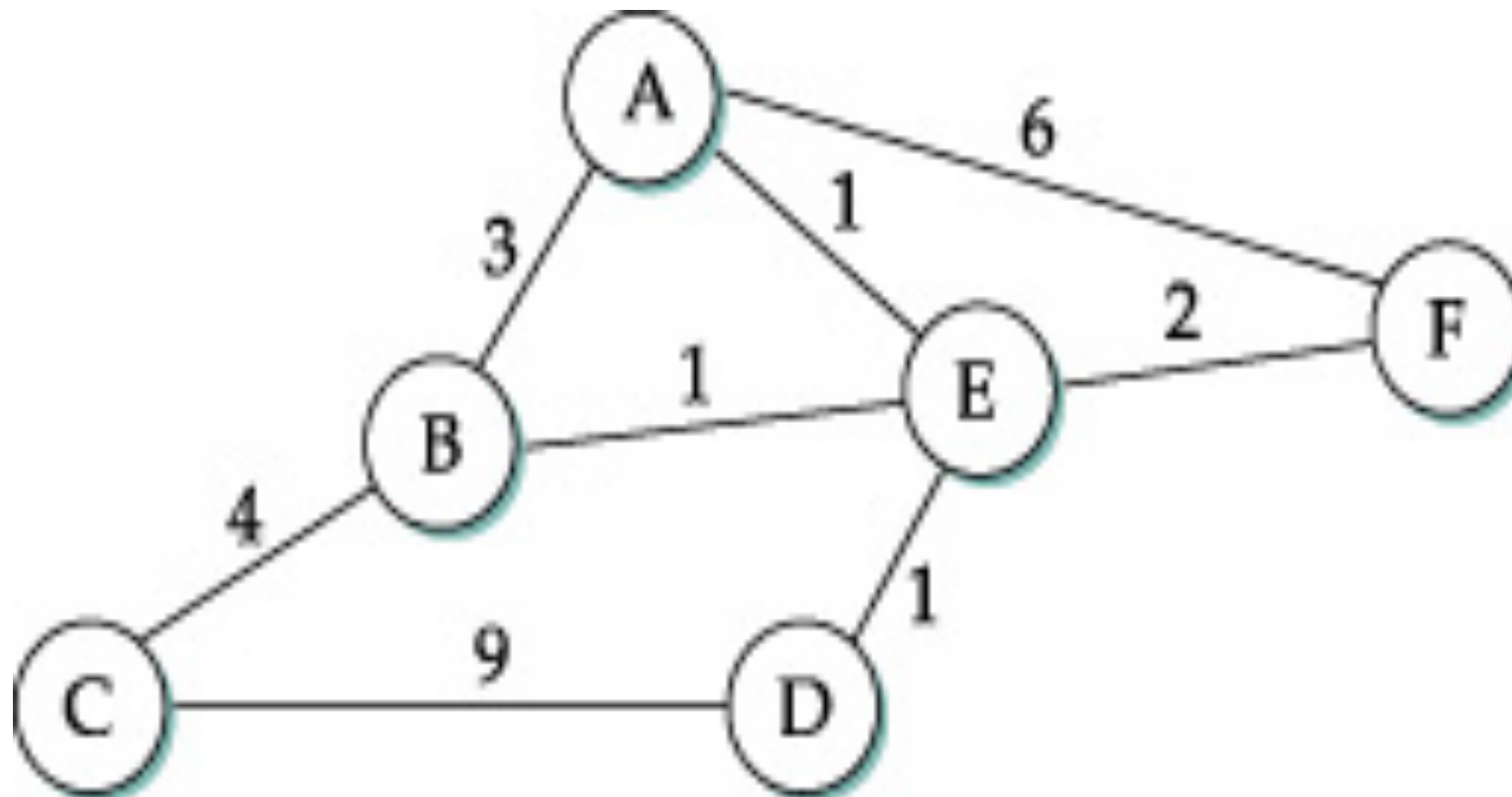- *routing* is the process of building forwarding tables.

Tables

Routing Table

Network	Next Hop
42	192.168.1.1

Forwarding Table

Network	Port	MAC
42	1	00:BC:D4:34:32:0B

Network Graph

Distance Vector (RIP)

- Table of distance/cost to all nodes

- Distribute to immediate neighbors

- Link Down = ∞

- periodic & triggered
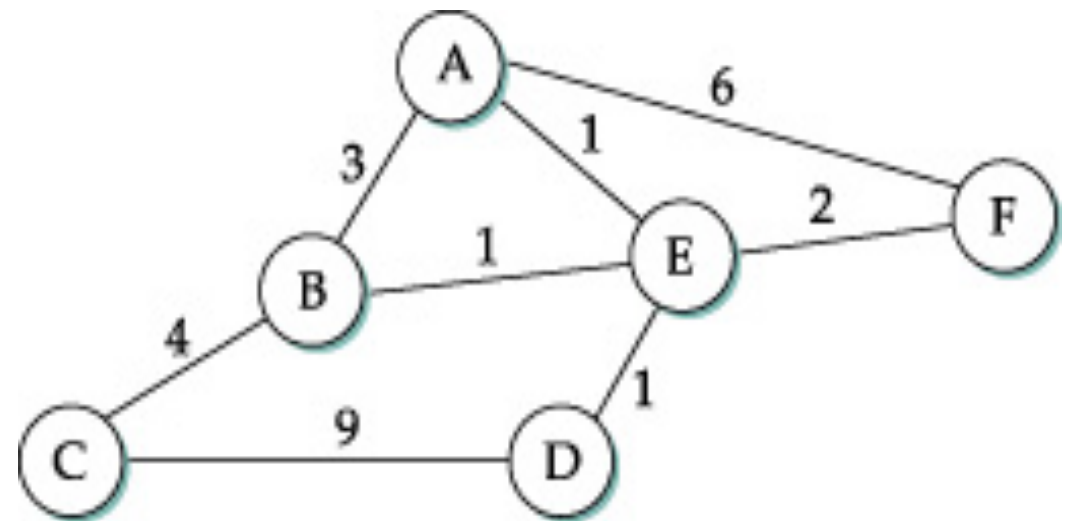
Exercise

1. Make "Routing Table" of network

   Destination   Cost   Next Hop

2. Send messages to neighbors with

   From: _____

Destination	Cost

3. Use messages from neighbors and destroy

4. Repeat from step 2

Distance Vector (RIP)
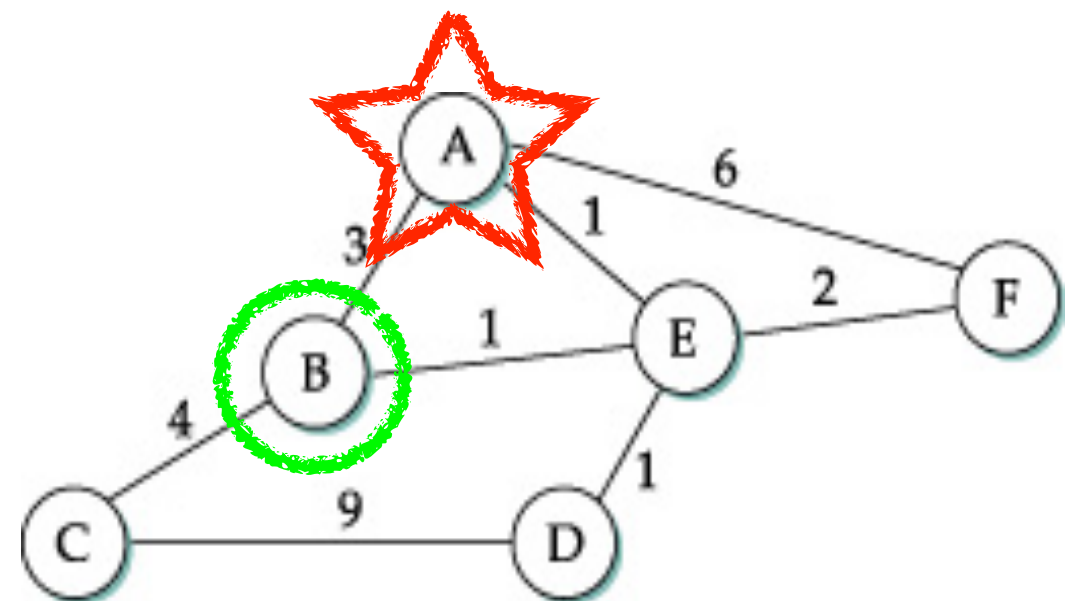
Dest	Cost	Hop
B	3	B
C	∞	-
D	∞	-
E	I	E
F	6	F
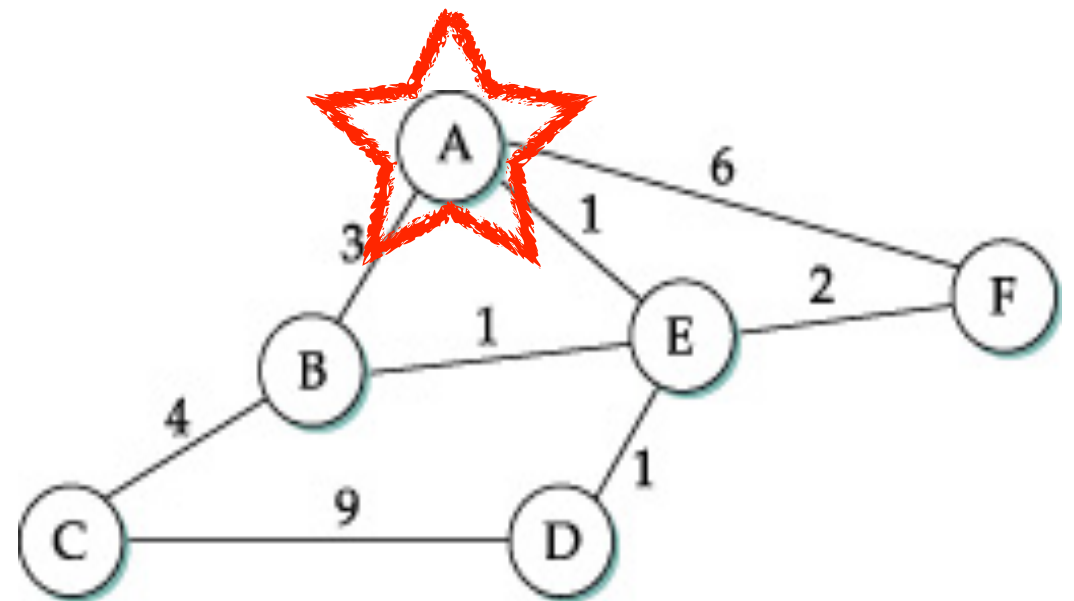
Distance Vector (RIP)

Dest	Cost	Hop
B	3	B
C	7	B
D	∞	-
E	1	E
F	6	F

Dest	Cost	Hop
A	3	A
C	4	C
D	∞	-
E	1	E
F	∞	-

Distance Vector (RIP)

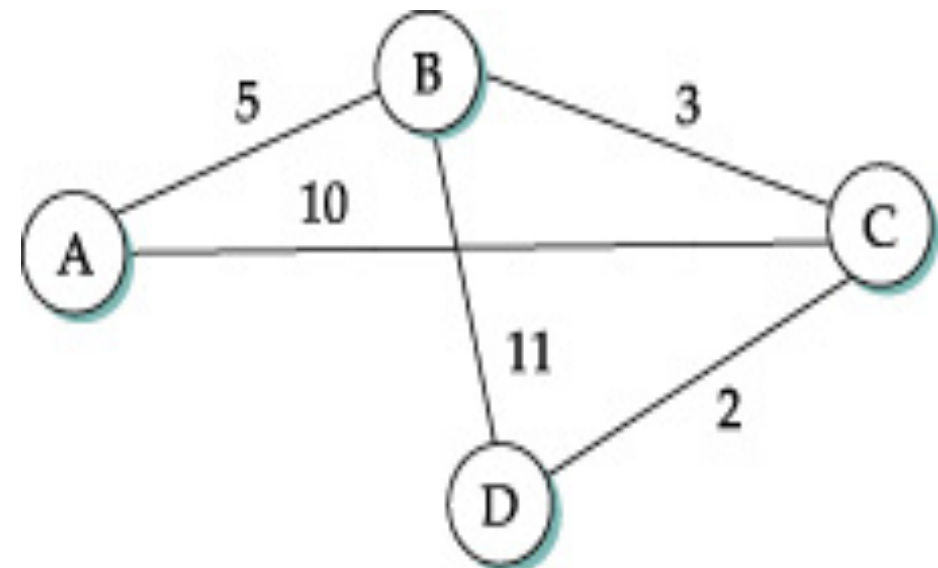| Dest	Cost	Hop
B	2	E
C	6	E
D	2	E
E	1	E
F	3	E
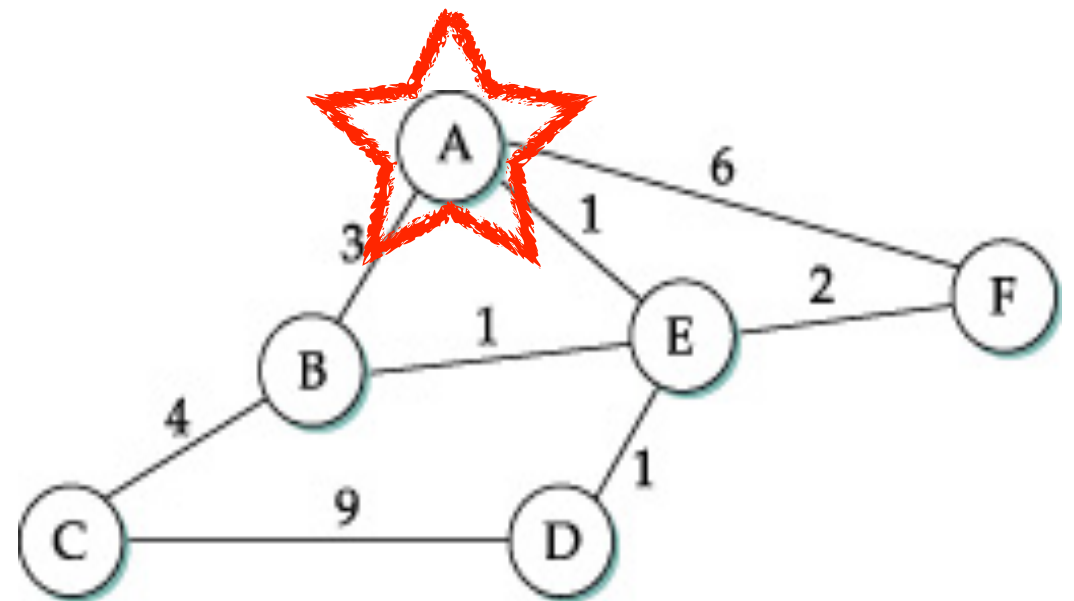
Distance Vector (RIP)

- Complete Map

- Exchange with immediate neighbors
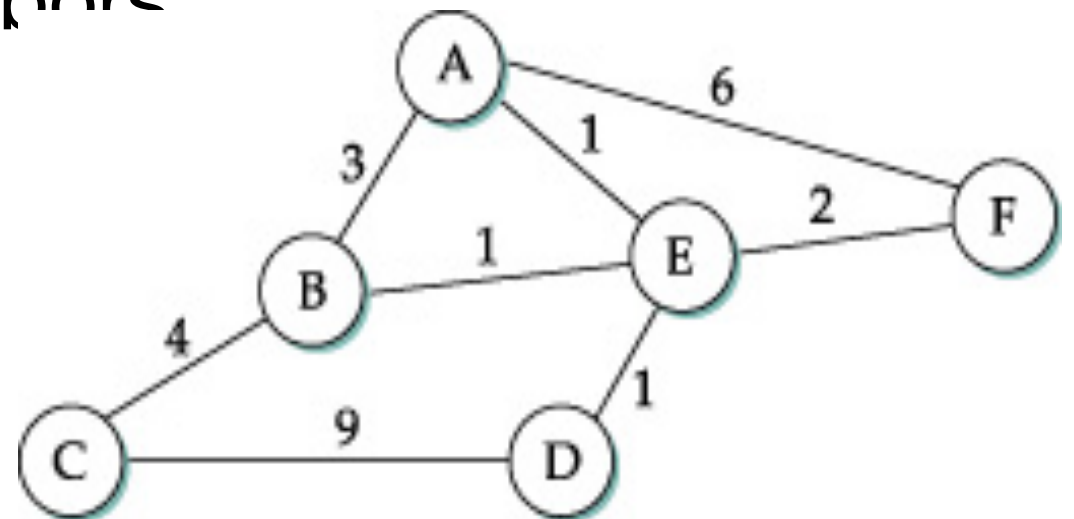
Link State

- Link State Packet (LSP)

  - Distance/Cost of neighbors

  - Flood to all routers

Exercise

1. Make "Routing Table" of network

   Confirmed   Tentative

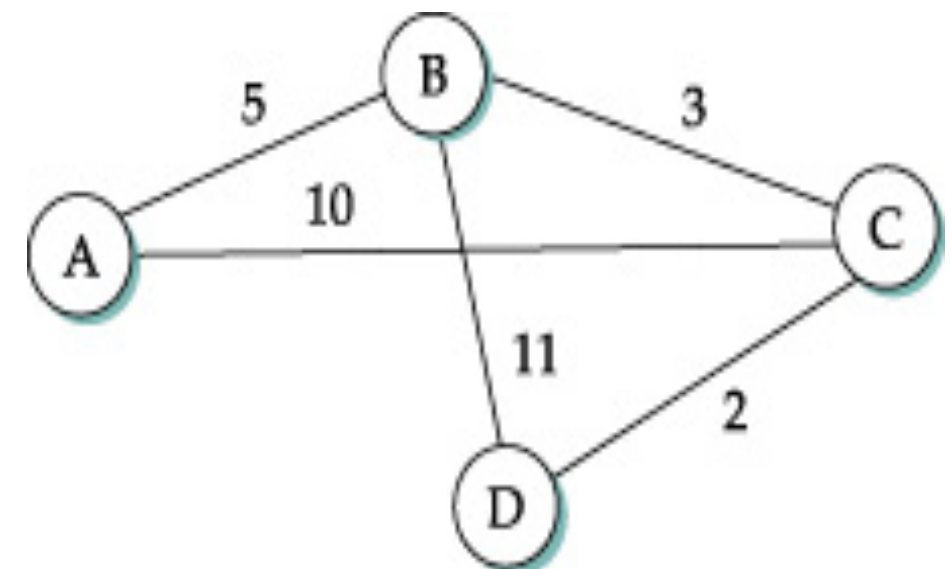2. Flood messages to everyone with neighbor information only, keep your own

   From: _____

Neighbor	Cost	NextHop

3. Use messages from others (Dijkstra's Alg.)

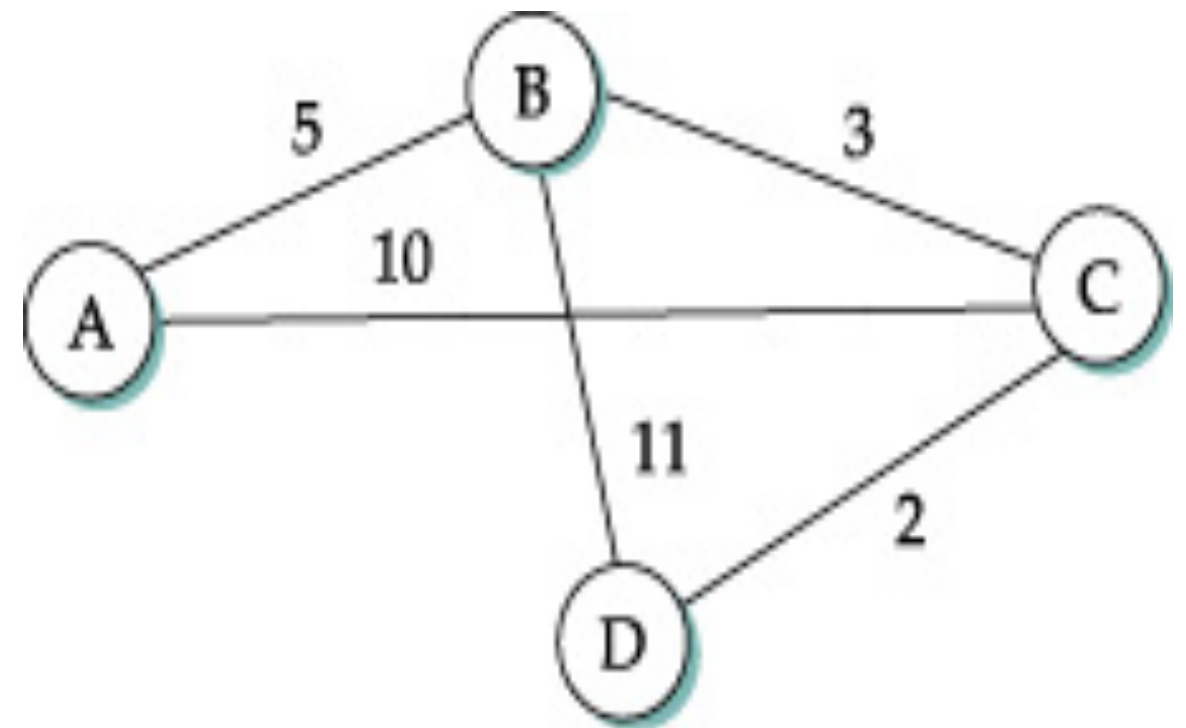4. Repeat from step 2, until tentative is empty

Link State

Step	Confirmed	Tentative	Replaced
1	(A,0,-)		
2	(A,0,-)	(B,5,B) (C,10,C)	
3	(A,0,-) (B,5,B)	(C,8,B) (D,14,B)	(C,10,C)
4	(A,0,-) (B,5,B) (C,8,B)	(D,10,B)	(D,14,B)
5	+(D,10,B)		

Link State

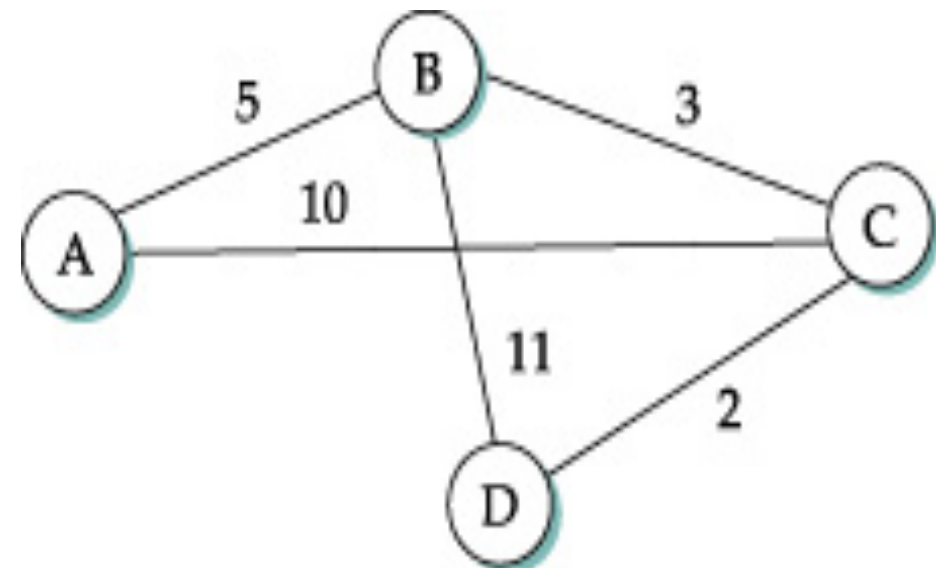| Dest	Cost	Hop
B	5	B
C	8	B
D	10	B

Link State

- Link State Packet (LSP)

  - Distance/Cost of neighbors

  - Flood to all routers

Metrics

- Bandwidth

- Latency

- "hops"

- Cost (dollars)

- Utilization

- Geo-Political boundaries
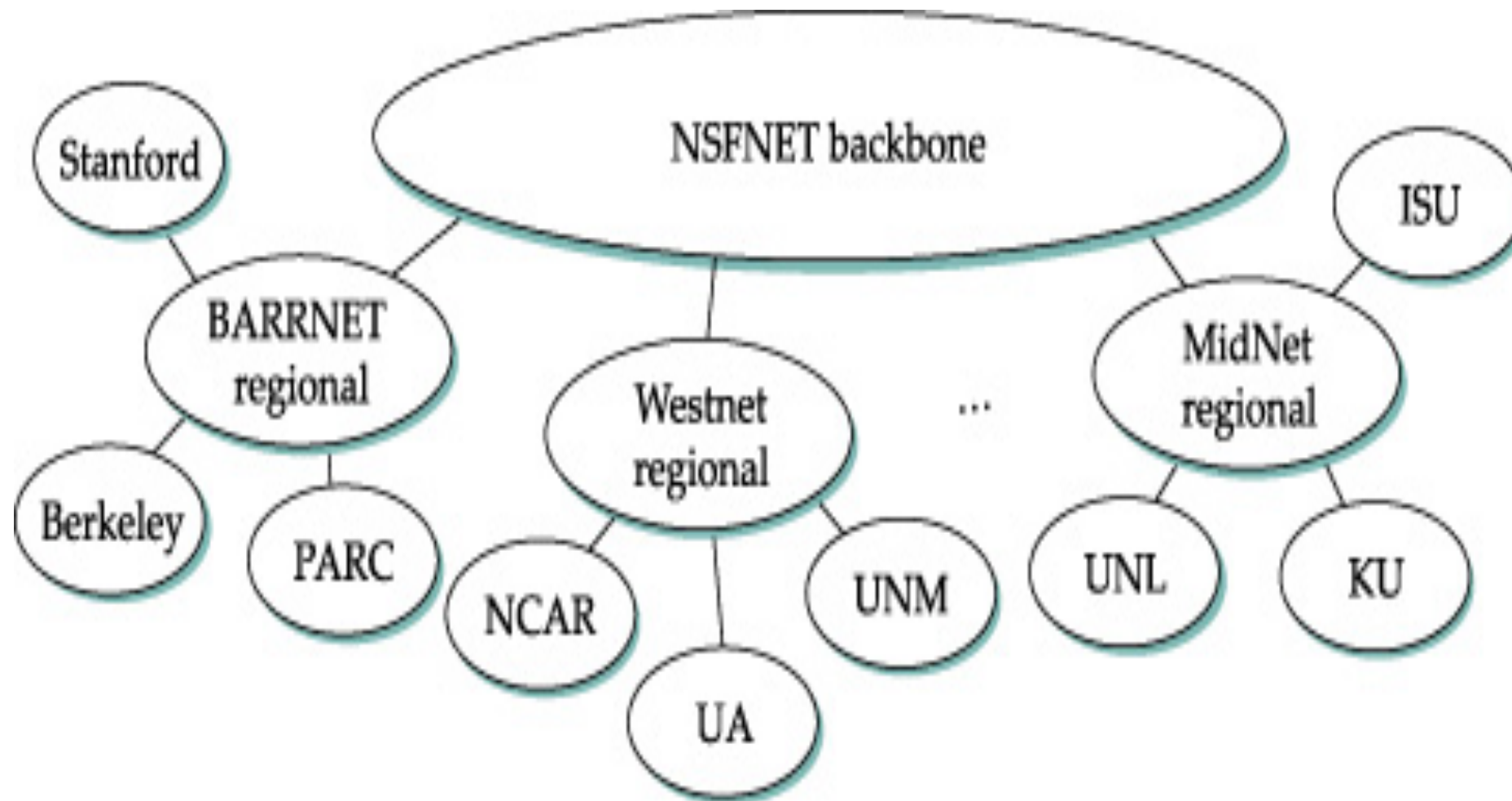
Inter-networking

✓Internet Protocol (IP)

✓Routing

★The Internet

• Multicast*

• Multi-protocol Label Switching*

The Internet

- Construction

  - Autonomous Systems

- Subnetting

- Classless Routing (CIDR)

- Border Gateway Routing (BGP)

- IPv6

What does it look like

Subnetting

- Network part of address uniquely identifies a physical network

  - Class B network (16 & 16 bits) - unused

  - Class C network (24& 8 bits) - too small

Subnetting

- Allocate one "network" to several physical networks

- internal to network

- outside routes same

Network number	Host number

Class B address

111111111111111111111111	00000000

Subnet mask (255.255.255.0)

Network number	Subnet ID	Host ID

Subnetted address

Subnetting

Subnet mask: 255.255.255.128
Subnet number: 128.96.34.0

128.96.34.15
H1

128.96.34.1
R1

128.96.34.130

Subnet mask: 255.255.255.128
Subnet number: 128.96.34.128

128.96.34.129

128.96.34.139
H2

H3

128.96.34.129
R2

128.96.33.1

128.96.33.14

Subnet mask: 255.255.255.0
Subnet number: 128.96.33.0
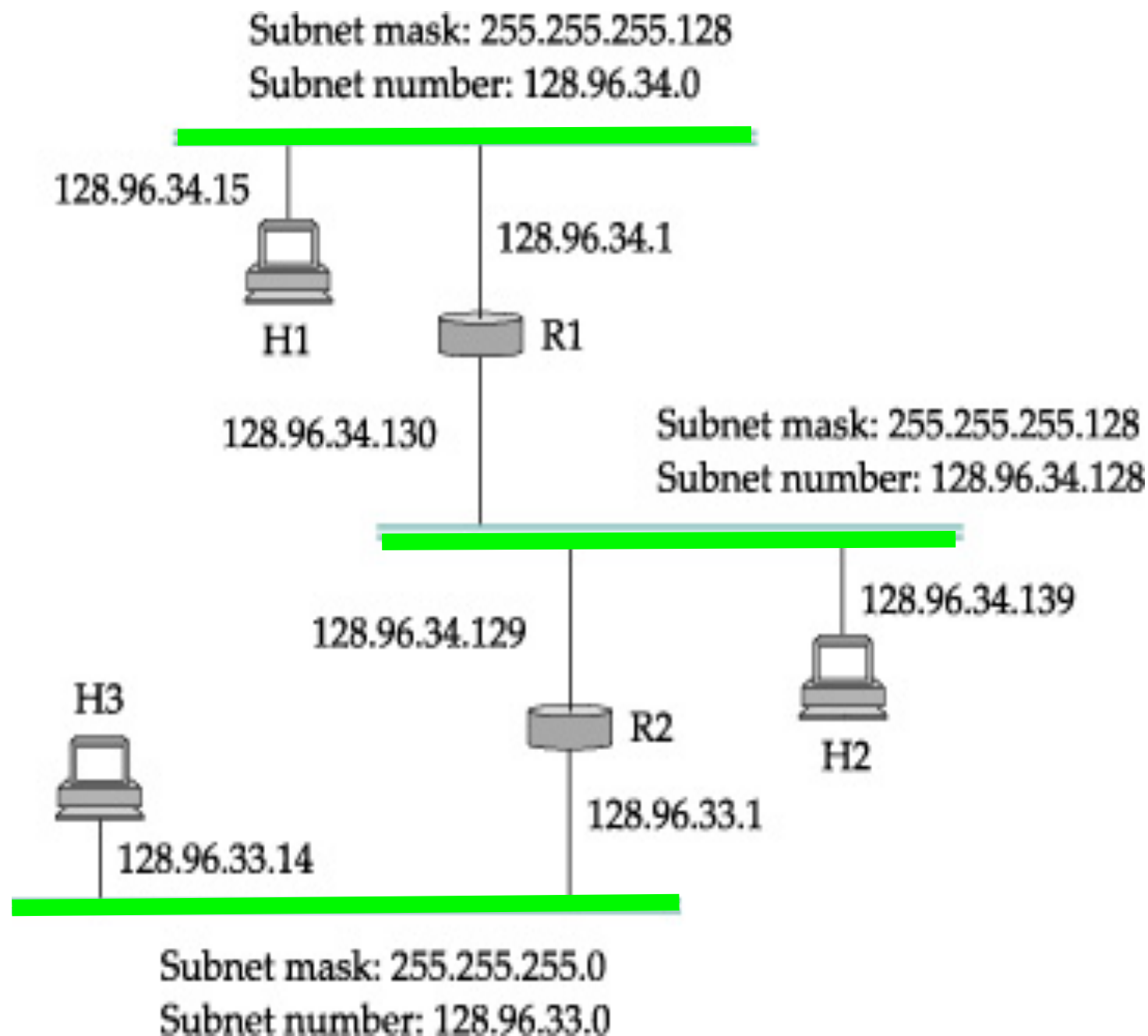
Subnetting

- Solves scalability problem

  - improve address assignment efficiency

  - aggregates information, from a distance

Classless Routing (CIDR)

- Scalability problems in the backbone

  - Single location with multiple C nets

  - Lots of entries to maintain outside

  - Assignment of Class B wastes

CIDR

- Solution: aggregate routes

  - 192.4.16/20 = 192.4.16 to 192.4.31
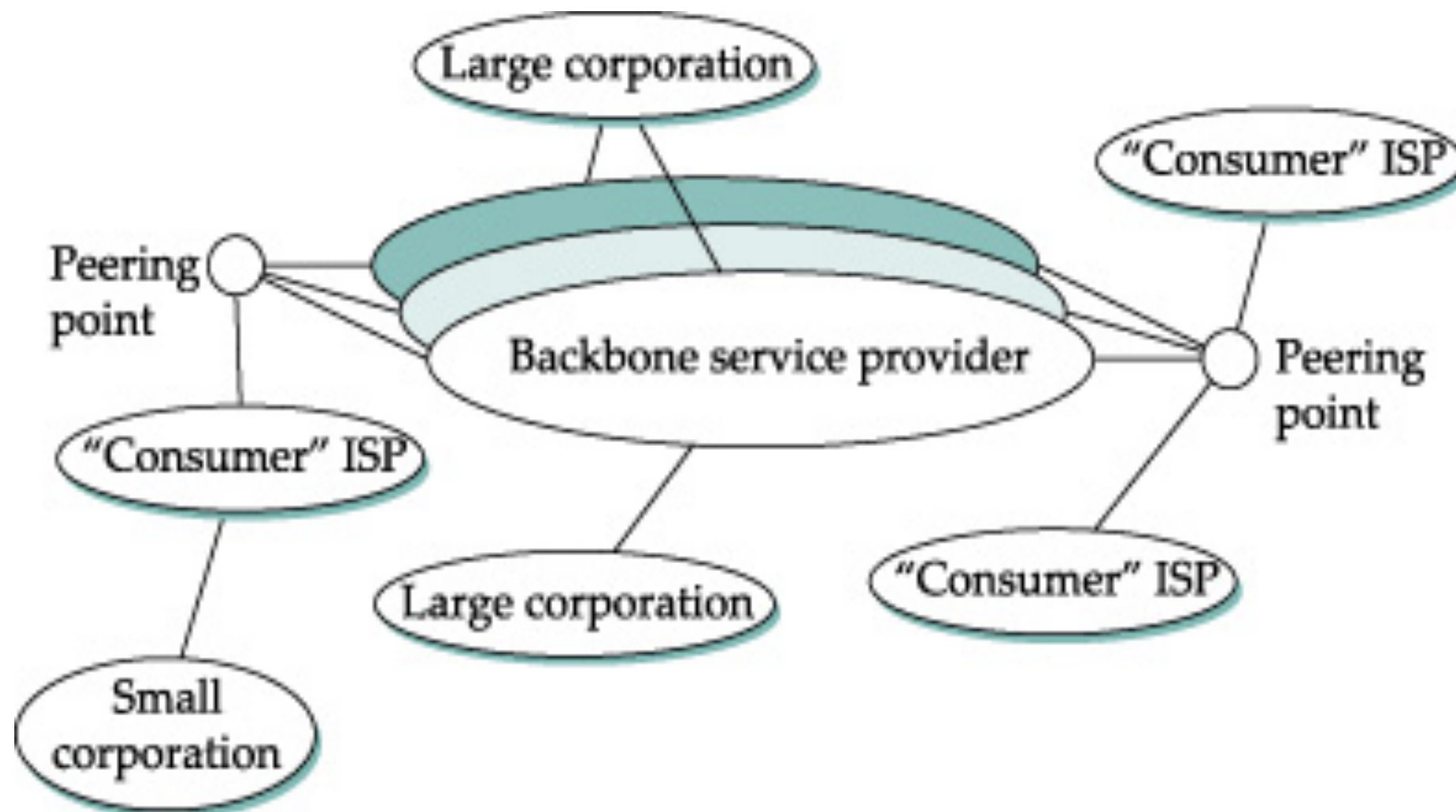
  - 192.4.16/24 = 192.4.16

CIDR

- Solution: Assign **blocks** of Class C addresses

  - creates aggregate routes to AS

Interdomain Routing

- Each Autonomous System (AS)

  - Has a "border"

  - Handles its own internal routing

  - Has its own policies for in & out

Interdomain Routing

Interdomain Routing

- Find *some* path to destination

- Avoid *loops*

- Compliant with *policies* of ASs

Interdomain Routing

- Border Gateway Protocol (BGP)

  - AS numbers are *unique*

  - One node is the *speaker* for the network

  - Advertises *complete paths* to networks

  - Prevent *loops*

  - Path *withdrawn* messages

Interdomain Routing

- Scalable (by hierarchy)

  - AS has own internal policies and routing

  - Only AS borders run BGP on backbone

  - Aggregate networks

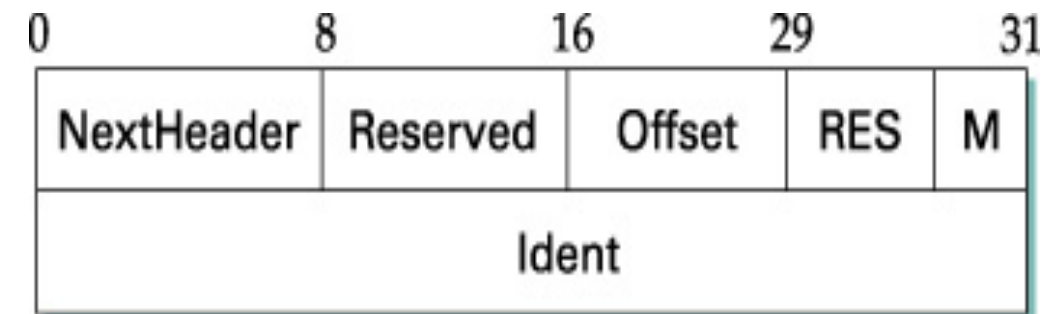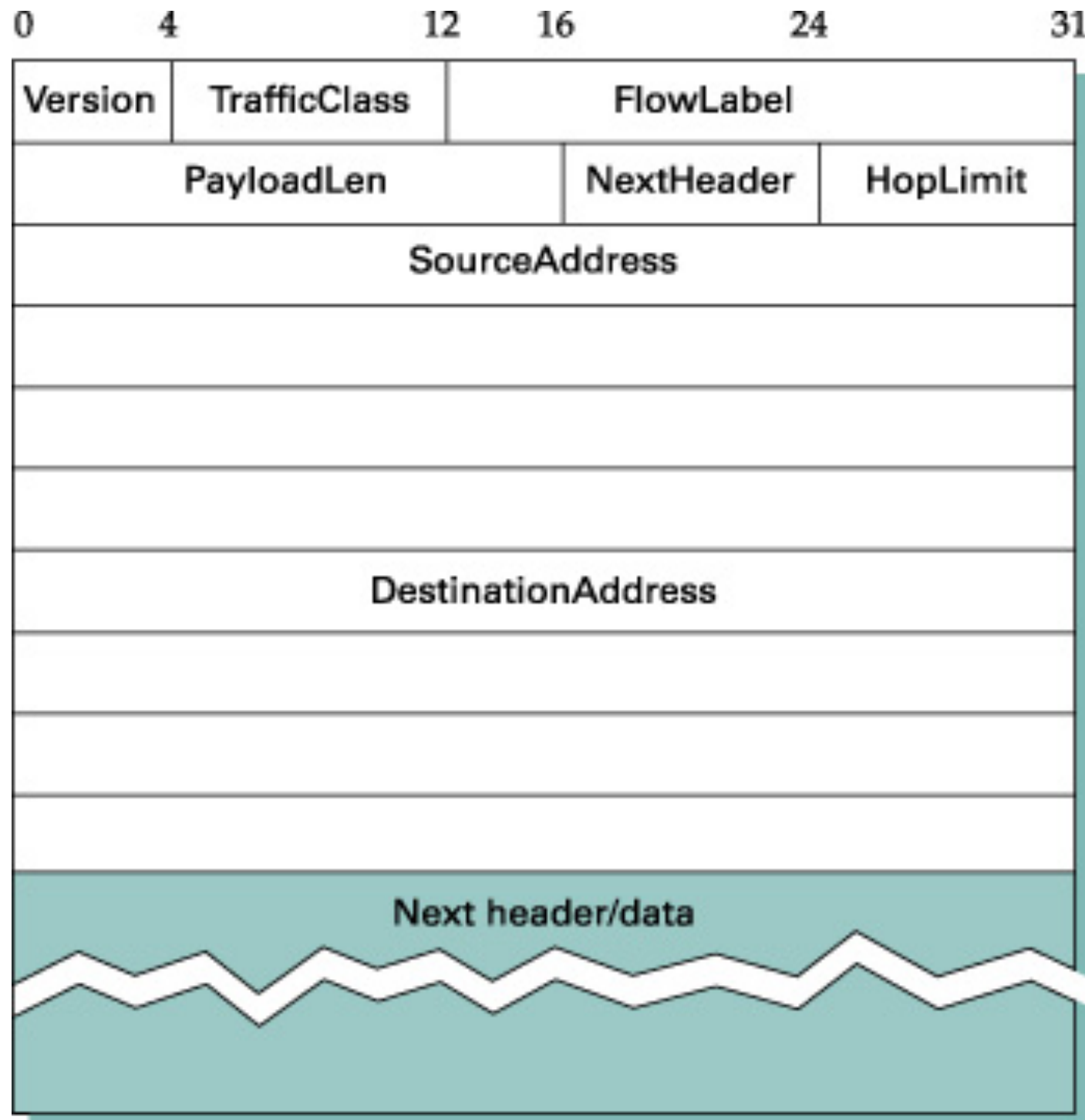 - Only need to find path to border

IPv6
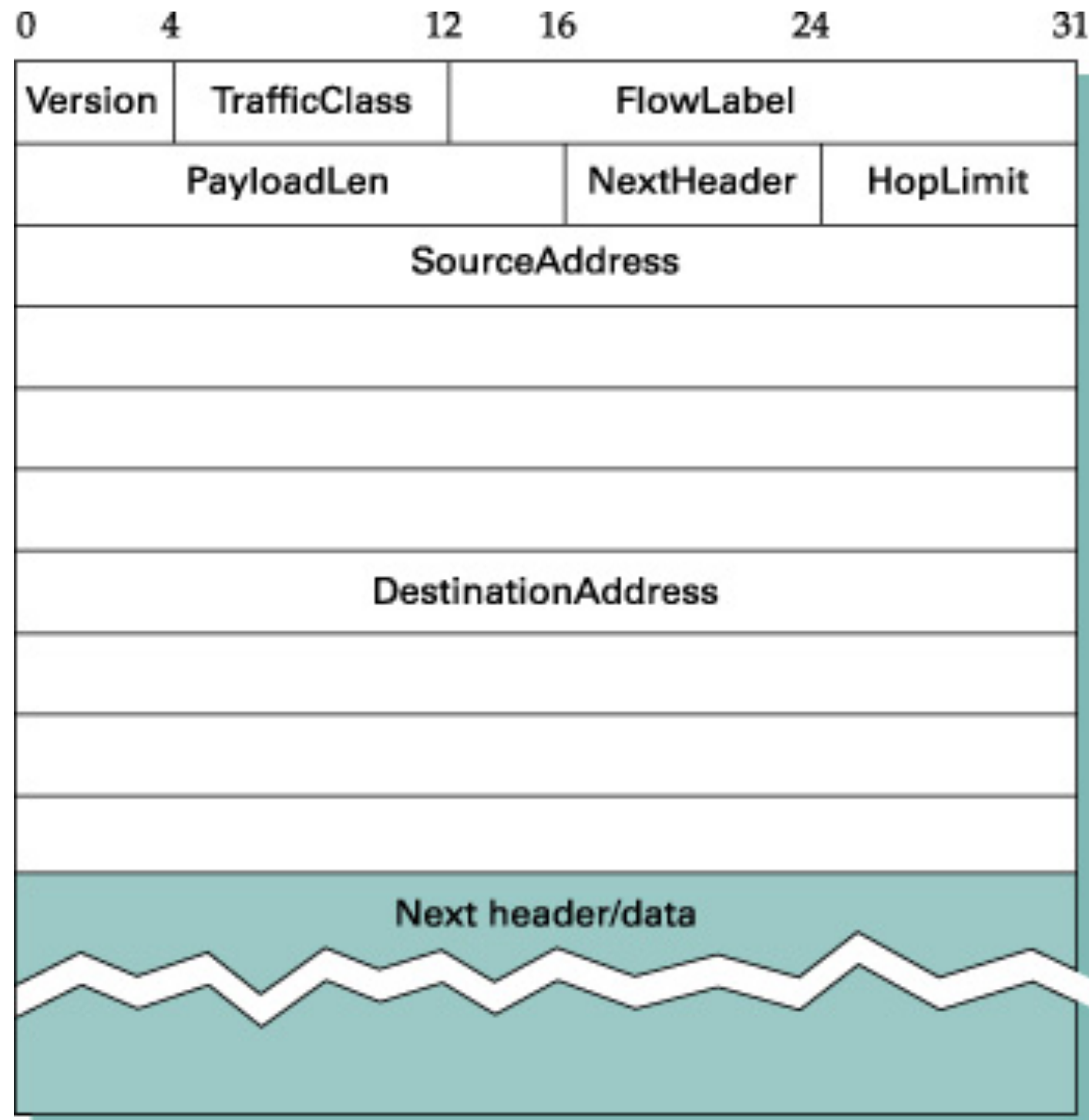
IPv6 - Why?

- Address exhaustion (really Class B)

- Real-time services

- Security

- Configuration

- Routing - mobile hosts

Addresses

- 128 bits = 3.4×10^{38} nodes

- Notation = x:x:x:x:x:x:x:x

  - :: = zeros; 45de:1230:FG::11:12

  - IPv4 (zero extended) ::FFFF:128.96.33.81

Packet Format

Configuration

- No more DHCP servers...

- Prefixes are assigned to networks (at AS)

  - Obtain interface ID

  - Obtain address prefix for subnet

    - local link only (1111 1110 10)

    - routers periodically advertise prefix

Inter-networking

✓ Internet Protocol (IP)

✓ Routing

✓ The Internet

• Multicast*

• Multi-protocol Label Switching*

Inter-networking

fin