ICS 153

Introduction to Computer Networks

Inst: Chris Davison

cbdaviso@uci.edu

ICS 153
Homework
Network Layer: Internet

- Chapter 5 (Internet)
- 36, 39, 45, 46, 53

ICS 153

The Network layer: Internet

- Definitions
  - Computer Network:An interconnected collection of autonomous computers
  - Internet: An interconnected collection of autonomous networks where each machine:
    - Runs TCP/IP
    - Has an IP address
    - Can send IP packets to all other machines on the Internet

The Internet

- There is no fixed topology
- Interconnection of networks is nearly arbitrary
- Large backbones are provided for interconnecting geographically dispersed regions

Protocols used in the Internet

- Network layer protocols
  - IP: Internet network layer protocol
  - ICMP: Internet Control Message Protocol
  - ARP: Address Resolution Protocol
  - RARP: Reverse Address Resolution Protocol
  - OSPF: Internet Routing Protocol
  - RIP: Internet Routing Protocol
  - BGP: Internet routing Protocol

Protocols used in the Internet

- Transport layer protocols
  - TCP: Transmission Control Protocol
  - UDP: User Datagram Protocol
- Application layer protocol
  - DNS: Domain Name Service
  - SNMP: Simple Network Management Protocol

The Internet Protocol (IP)

- Provides delivery of packets from one host on the Internet to another host on the Internet, even if the hosts are on different networks.
- Internet packets are called “datagrams” and may be up to 65,535 bytes in length (although they are typically much shorter)
- Internet IMPs are known as “routers” and they operate in a connectionless mode

The Internet Protocol (IP)

- IP is a “best effort” protocol
  - Delivery is not guaranteed
  - Unreliable, connectionless service (ie. Datagram service)
  - Error recovery is the responsibility of upper layers

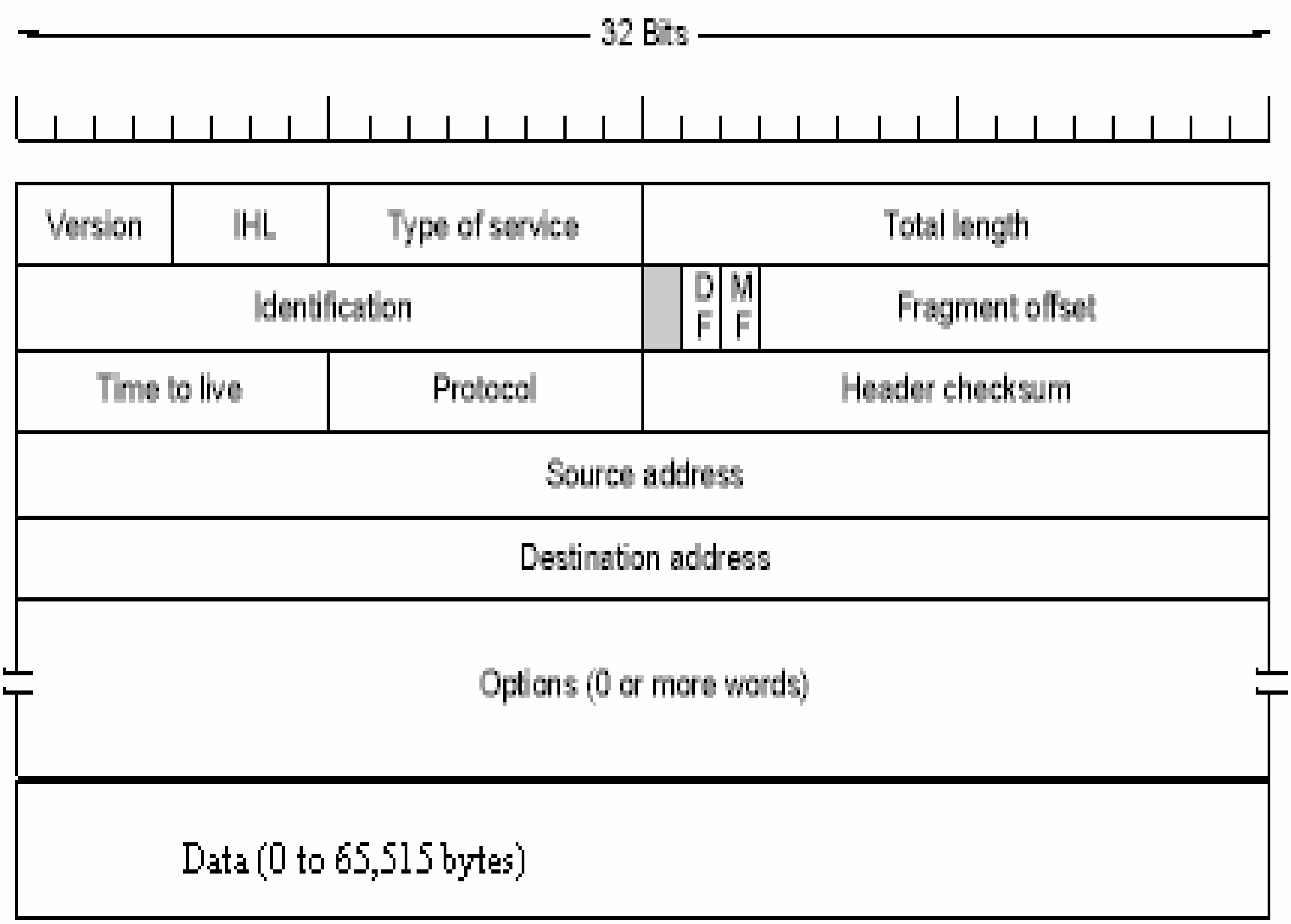
The Internet Protocol (IP)

- Primary IP function:
  - accept data from Transport Layer (TCP or UDP)
  - create a datagram
  - route the datagram through the network
  - deliver it to the recipient host

IP Header

- Minimum of 160 bits
  - 32-bit words times 5
- Options may be added to header
  - Options can be up to 320 bits
    - 32-bit words times 10

IP Packet Format



IP Packet Fields

- Version
  - The IP version number (currently 4)
- IHL
  - IP header length in 32-bit words
- Type of Service
  - Priority Information
  - Routers ignore this field
- Total Length
  - Total length of IP packet including header
  - 2^16 -1 (65,535 octets)

IP Packet Fields

- Identification
  - When an IP packet is segmented into multiple fragments, each fragment is given the same identification
  - This field is used to reassemble fragments
- DF
  - Don’t Fragment
- MF
  - More Fragments
  - When a packet is fragmented, all fragments except the last one have this bit set

IP Packet Fields

- Fragment Offset
  - The fragment’s position within the original packet
- Time to Live
  - Hop count, decremented each time the packet reaches a new router
  - When hop count = 0, packet is discarded
- Protocol
  - Identifies which transport layer protocol is being used for this packet
- Header Checksum
  - Verifies the contents of the IP header
  - Calculated using a one’s complement algorithm

IP Fragmentation

- Fragmentation can be controlled by the DF field
  - may result in suboptimal routing
  - IP standard requires all hosts to accept datagrams of up to 576 octets
- Fragment Offset
  - The fragment’s position within the original packet
  - all fragments (except last) must be a multiple of 8 octets
    - 8 octet chunk of data is called a “fragment block”
  - 13 bits provided for fragment offset
    - max of 8,192 fragment blocks

IP Fragmentation Reconstruction

- A fragmented datagram is reconstructed at the recipient host
  - Pieces of an IP fragment may arrive out of order
- Fragments with matching *Identification, Source IP address, Destination IP address*, and *Protocol* fields belong together and are merged as they arrive

IP Protocol Field

- Assigned by Internet Assigned Numbers Authority (IANA)
  - RFC 1700
- Common IP Protocol Field numbers:
  - 1 ICMP Internet Control Message Protocol -Networking Utilities
  - 2 IGMP Internet Group Management Protocol - Supports Multicasting groups
  - 88 IGRP Cisco's Interior Gateway Routing Protocol - Exchange of router packets

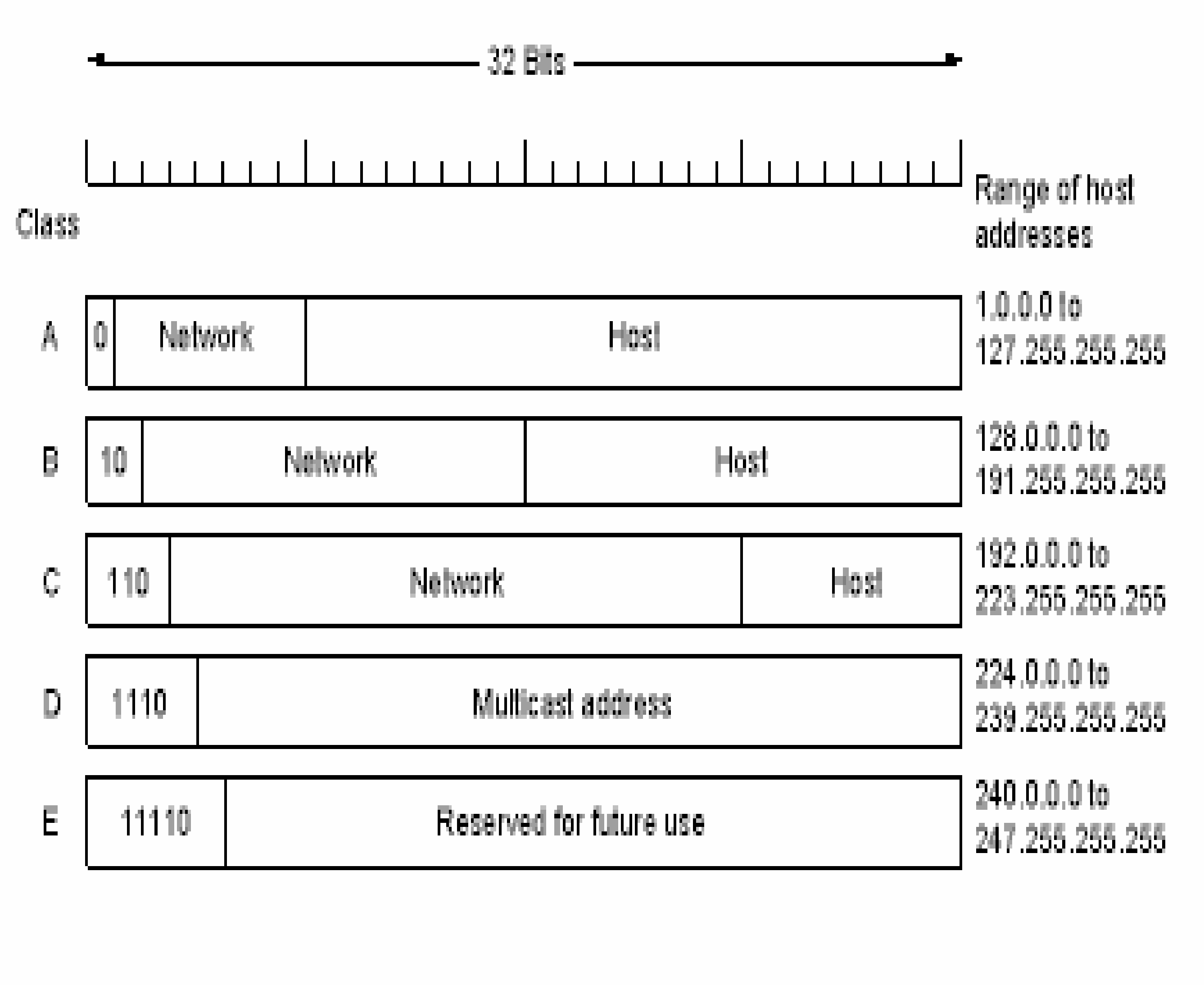
IP Packet Fields

- Source and Destination Addresses
  - Uniquely identify sender and receiver of the packet
  - Given as an IP address (32 bits)
- Options
  - Up to 40 bytes in length
  - Used to extend functionality of IP
  - Examples: source routing, security, record route

IP Addresses

- 32 bits long
- Notation
  - Each byte is written in MSB order, separated by decimals
  - Example
    - 128.195.15.80
    - 0.0.0.0 (lowest) to 255.255.255.255 (highest)
- Address Classes
  - Class A,B,C,D,E
  - loopback (127.any.any.any)
  - broadcast (255.255.255.255)

IP Address Classes



32 Bits

Class				Range of host addresses
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
E	11110	Reserved for future use		240.0.0.0 to 247.255.255.255

IP Address Classes

- Class A
  - For very large organizations
  - 2^24 -2 hosts allowed
- Class B
  - For large organizations
  - 2^16 - 2 hosts allowed
- Class C
  - For small organizations
  - 2^8 - 2 hosts allowed
- Class D
  - Multicast addresses
  - No network/host hierarchy

Special IP Addresses

- Broadcast to a specific subnet
  - All host bits set to 1
  - Example (class B):
    - 128.195.15.255
    - 10000000.11000011.00001111.11111111
- Broadcast to a local subnet
  - All bits set to 1
  - Example:
    - 255.255.255.255
    - 11111111.11111111.11111111.11111111
    - Host uses: 0.0.0.0 as source address
  - Used in DHCP and BOOTP protocols

Special IP Addresses

- Subnet Reference
  - Use a 0 in host byte
  - Example (class B):
    - 128.195.15.0
    - 10000000.11000011.00001111.00000000

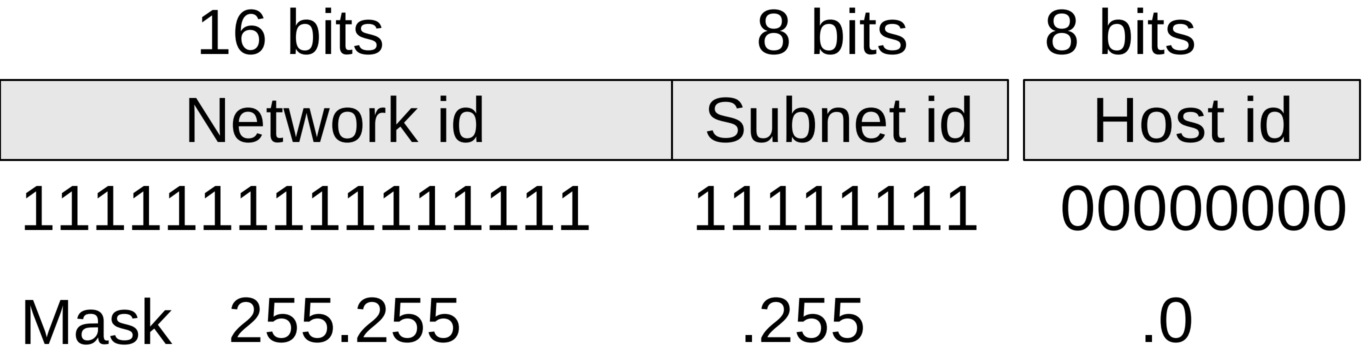
- Loopback Address
  - Processed locally at the host
  - Example:
    - 127.0.0.1

IP Address Hierarchy

- Note that Class A, B, and C addresses only support two levels of hierarchy
- Each address contains a network and a host portion, meaning two levels of hierarchy
- However, the host portion can be further split into “subnets” by the address class owner
- This allows for more than two levels of hierarchy

IP Subnet Mask

- Subnet masks allow hosts to determine if another IP address is on the same subnet or the same network



16 bits	8 bits	8 bits
Network id	Subnet id	Host id
1111111111111111	11111111	00000000
Mask 255.255	.255	.0

IP Subnet Mask

- 8 Bit Subnet Masking is common:

Example (129.195.15.1):

128 .195 .15 .1

10000000.11000011.00001111.00000001

255 .255 .255 .0 (mask)

11111111.11111111.11111111.00000000

Network ID Subnet ID Host ID

IP Subnet Mask

Assume IP addresses A and B share subnet mask M.

Are IP addresses A and B on the same subnet?

1. Compute (A and M). (Boolean AND)
2. Compute (B and M). (Boolean AND)
3. If (A and M) = (B and M) then A and B are on the same subnet.

Example: A and B are class B addresses

A = 165.230.82.52
B = 165.230.24.93
M = 255.255.255.0

Same network?
Same subnet?

IP Subnet Mask

- Note
  - 0 AND 0 = 0
  - 0 AND 1 = 1 AND 0 = 0
  - 1 AND 1 = 1
- Thus, computing (A and M) results in
  - Network ID = Network ID of A
  - Subnet ID = Subnet ID of A
  - Host ID = 0

IP Subnet Mask

- With 8-bit subnet masking on a class B network: Only 254 hosts per network
  - Recall 0 and 1 are reserved.
- Solution:
  - Variable length Subnet mask

Variable Length Subnet Mask

- When opting for < 8 bit subnets, you are opting for fewer subnets:

Subnet bits	Subnets	Hosts bits	Hosts
7	128	9	510
6	64	10	1022

- Example 7-bit mask (255.255.254.0):
  - Example (128.15.2.1):
    - 128 .15 .2 .1
    - 10000000.00001111.0000001**0.00000001**
    - First host on the subnet
  - Example (128.15.3.254):
    - 128 .15 .3 .254
    - 10000000.00001111.0000001**1.1111110**

Variable Length Subnet Mask

- When opting for > 8 bit subnets, you are opting for fewer hosts

Subnet bits	Subnets	Hosts bits	Hosts
9	512	7	126
10	1024	6	62

- Example 9 bit mask (255.255.255.128) :
  - Example (130.15.1.1)
  - 130 .15 .1 .1
  - 10000010.00001111.00000001.0**0000001**
    - First host on the subnet
  - Example (130.15.1.126)
  - 130 .15 .1 .126
  - 100000010.00001111.00000001.0**1111110**
    - Last host on the subnet

Subnet Mask

- Why do we need subnet mask?
  - When subnetting is introduced, a routing table is modified to include (this-network, subnet, 0) and (this-network, this-subnet, host)

Subnet Mask

- Routing table

network ID	subnet ID	host ID
this network	this subnet	A
this network	this subnet	B
this network	different subnet	0
this network	different subnet	0
different network	0	0

- Subnet mask helps quickly identifying which routing table entry to look up

Subnet Mask

- Real Routing Table of 128.195.7.95:

```
Routing Table:
  Destination          Gateway            Flags  Ref   Use   Interface
-------------------- -------------------- ----- ----- ------ ---------
128.195.2.0          128.195.7.1          UG       0     62
128.195.1.0          128.195.7.1          UG       0   2957
128.195.7.0          128.195.7.95         U        3   3355  le0
128.195.6.0          128.195.7.1          UG       0     11
128.195.38.0         128.195.7.1          UG       0      0
128.195.4.0          128.195.7.1          UG       0   3570
128.195.36.0         128.195.7.1          UG       0      1
128.195.11.0         128.195.7.1          UG       0  14230
128.195.10.0         128.195.7.1          UG       0   5108
128.195.15.0         128.195.7.1          UG       0      1
128.195.18.0         128.195.7.1          UG       0     15
128.195.23.0         128.195.7.1          UG       0      1
128.195.22.0         128.195.7.1          UG       0      9
128.195.21.0         128.195.7.1          UG       0   2399
128.195.20.0         128.195.7.1          UG       0      3
128.195.26.0         128.195.7.1          UG       0  13124
128.195.25.0         128.195.7.1          UG       0   1904
128.195.24.0         128.195.7.1          UG       0     13
 128.195.31.0         128.195.7.1          UG       0    207
128.195.30.0         128.195.7.1          UG       0      0
128.200.0.0          128.195.7.1          UG       0   3599
224.0.0.0            128.195.7.95         U        3      0  le0
default              128.195.7.1          UG       0   1351
127.0.0.1            127.0.0.1            UH       01563556  lo0
```

Extending IPv4 Address Space

- CDIR
  - Classes Inter-Domain Routing
  - RFC 1519
  - Allocate remaining IPv4 addresses in variable-sized blocks.
    - No regard for address class during IP assignment

Extending IPv4 Address Space

- CDIR Addresses the Following:

1)Exhaustion of the class B network address space. One fundamental cause of this problem is the lack of a network class of a size which is appropriate for mid-sized organization; class C, with a maximum of 254 host addresses, is too small, while class B, which allows up to 65534 addresses, is too large for most organizations.

2) Growth of routing tables in Internet routers beyond the ability of current software, hardware, and people to effectively manage.

3) Eventual exhaustion of the 32-bit IP address space.

Extending IPv4 Address Space

- NAT
  - Network Address Translation
  - RFC 3022
  - Basic NAT operation is as follows: A stub domain with a set of private network addresses could be enabled to communicate with external network by dynamically mapping the set of private addresses to a set of globally valid network addresses.

Extending IPv4 Address Space

- NAT

```
\ | /
+---------------+ WAN .                     +/----------------
|Regional Router|---------------------------|Stub Router w/NAT|
+---------------+ .                         +-----------------+
                                                    |
                                                    |

                                                   LAN
```

Traditional NAT Configuration

Extending IPv4 Address Space

- NAT
  - NAT can support 61,440 hosts per IPv4 address
    - 2^16 (TCP sequence numbers)- 4096 (reserved ports)
  - Whenever a NAT device receives an out-going packet it replaces the internal IP source address with the organization’s true IP address.
    - NAT maintains a table of hosts (translation table)
    - NAT modifies the outgoing TCP source port and injects an index number for the translations table
  - Whenever a NAT device receives an incoming packet it replaces the organization’s IP destination address with the entry from the translation table.

Extending IPv4 Address Space: Criticisms of NAT

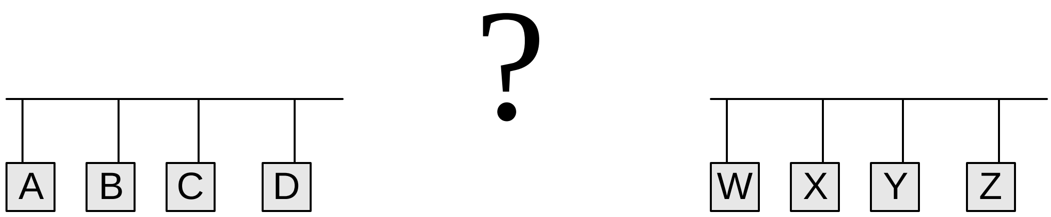
- NAT violates the architectural model of IP
  - One IP address does not uniquely identify a single machine.

- NAT violates the protocol layering of TCP/IP
  - NAT is aware of the TCP layer (source/destination) and modifies these entries.

- RFC 2993: Architectural Implications of NAT

IP Routing

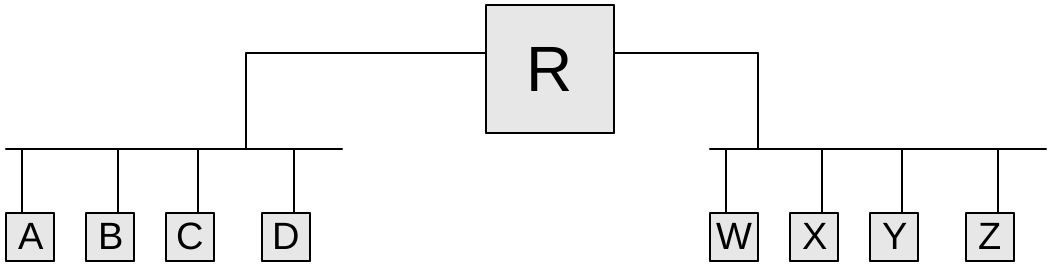
How do you get a packet
from one network to another?



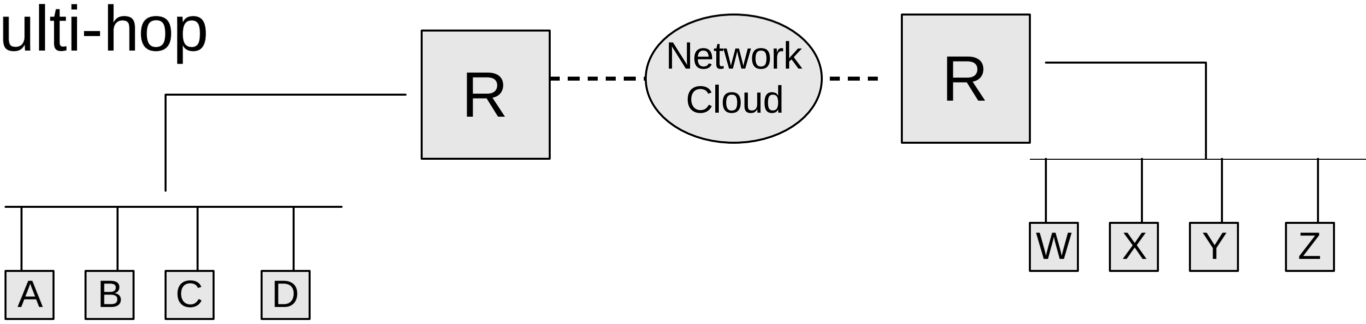
IP Routing

Answer: with a router (or a series of routers)

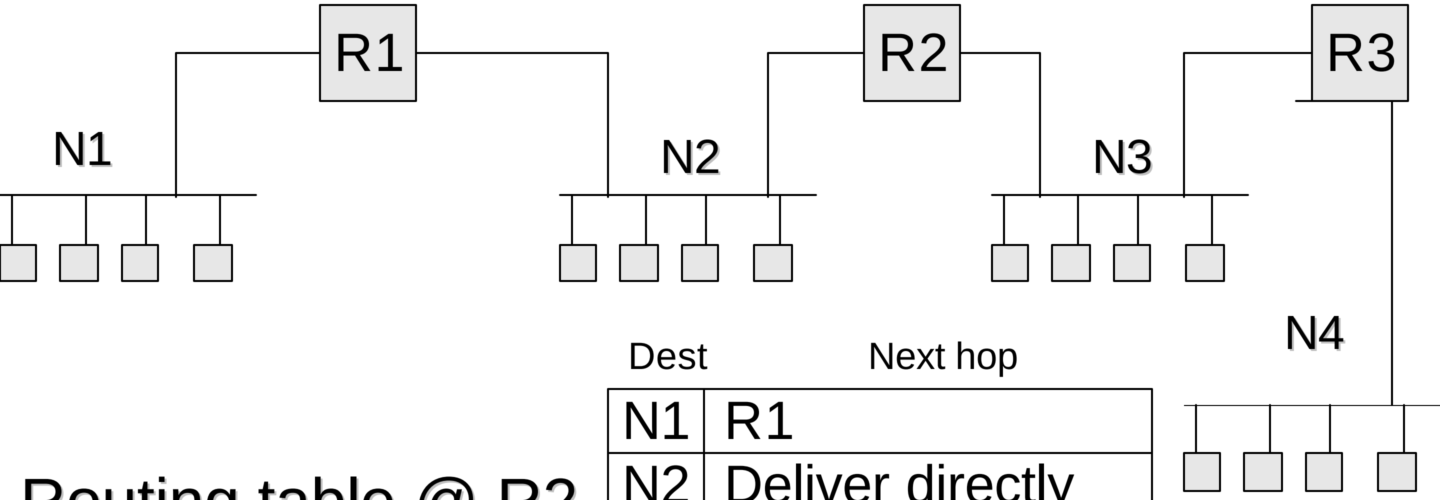
Case 1:
Single hop



Case 2:
Multi-hop



IP Routing



Routing table @ R2

Dest	Next hop
N1	R1
N2	Deliver directly
N3	Deliver directly
N4	R3

Actual routing table contains IP addresses, Flags indicating type of entries, net mask etc. (see slide 34).

Searching the Routing table

- First, search for a matching host address
  - Flag H is set
- Second, search for a matching network address
  - Need to know the number of bits to use for network ID
- Third, search for a default entry
    - Execute netstat -rn on your machine and find the contents of the routing table
  - Default entry allows for a single entry for a list of entries that have the same next-hop value

Searching the Routing table

- The goal of routing lookup is to find the most specific address that matches the given destination
  - host address is preferred over a network address which is preferred over a default address
- An entry in the routing table matches a search key (destination) if the search key logically ANDed with the network mask of the entry equals the entry itself
  - search key could have multiple matches; take the more

Internet Control Protocols

- ICMP
- ARP
- RARP
- RIP
- OSPF
- BGP

ICMP

- Internet Control Message Protocol
  - RFC 792
- ICMP information is carried in the data portion of an IP datagram
- Handles special Internet control functions
  - When an IP datagram has been discarded (link down, router crash, etc.) ICMP reports problems to the source address.

ICMP

- Responsibilities:
  - – Reporting unreachable destinations
  - – Reporting IP packet header problems
  - – Reporting routing problems
  - – Reporting echoes (pings)
- ICMP Message Format:

8 bits	8 bits	16 bits
Type	Code	Checksum

The rest of the message depends on the ICMP type

ICMP

- Protocol for error detection and reporting
    - tightly coupled with IP, unreliable
- ICMP messages delivered in IP packets
- ICMP functions:
    - Announce network errors
    - Announce network congestion
    - Assist trouble shooting
    - Announce timeouts

ICMP Types

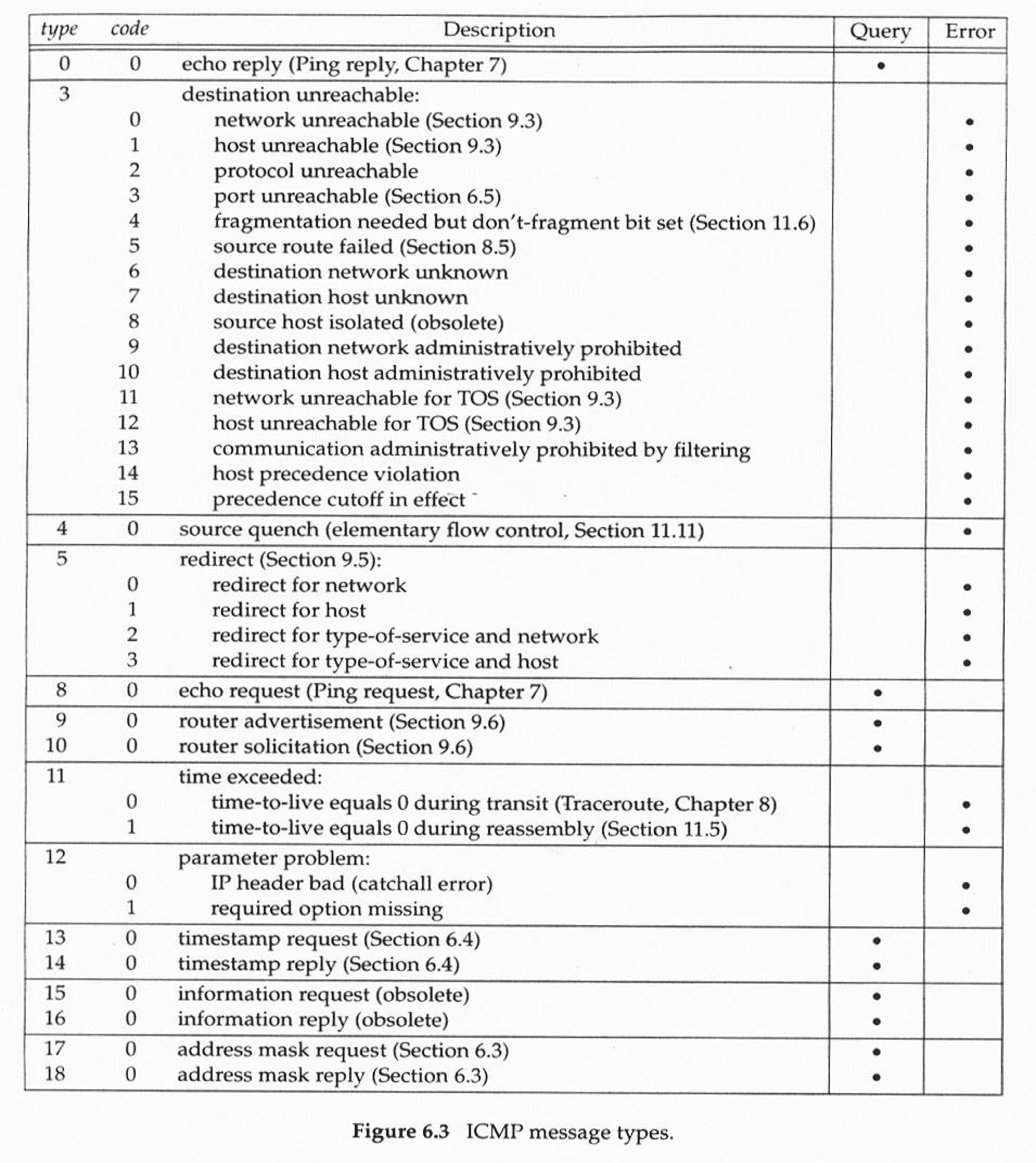
Type	Meaning
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect
8	Echo Request
11	Time Exceeded
12	Parameter Problem

ICMP Code Field

- Many ICMP Types have an associated code.
- Example: Type 3 (Destination Unreachable) has the following codes:

  0 Net Unreachable
  1 Host Unreachable
  2 Protocol Unreachable
  3 Port Unreachable
  4 Fragmentation Needed and Don't Fragment was Set
  5 Source Route Failed
  6 Destination Network Unknown
  7 Destination Host Unknown
  8 Source Host Isolated
  9 Communication with Destination Network is Administratively Prohibited
  10 Communication with Destination Host is Administratively Prohibited
  11 Destination Network Unreachable for Type of Service
  12 Destination Host Unreachable for Type of Service

ICMP



type	*code*	Description	Query	Error
0	0	echo reply (Ping reply, Chapter 7)	•	
3		destination unreachable:		
	0	network unreachable (Section 9.3)		•
	1	host unreachable (Section 9.3)		•
	2	protocol unreachable		•
	3	port unreachable (Section 6.5)		•
	4	fragmentation needed but don't-fragment bit set (Section 11.6)		•
	5	source route failed (Section 8.5)		•
	6	destination network unknown		•
	7	destination host unknown		•
	8	source host isolated (obsolete)		•
	9	destination network administratively prohibited		•
	10	destination host administratively prohibited		•
	11	network unreachable for TOS (Section 9.3)		•
	12	host unreachable for TOS (Section 9.3)		•
	13	communication administratively prohibited by filtering		•
	14	host precedence violation		•
	15	precedence cutoff in effect		•
4	0	source quench (elementary flow control, Section 11.11)		•
5		redirect (Section 9.5):		
	0	redirect for network		•
	1	redirect for host		•
	2	redirect for type-of-service and network		•
	3	redirect for type-of-service and host		•
8	0	echo request (Ping request, Chapter 7)	•	
9	0	router advertisement (Section 9.6)	•	
10	0	router solicitation (Section 9.6)	•	
11		time exceeded:		
	0	time-to-live equals 0 during transit (Traceroute, Chapter 8)		•
	1	time-to-live equals 0 during reassembly (Section 11.5)		•
12		parameter problem:		
	0	IP header bad (catchall error)		•
	1	required option missing		•
13	0	timestamp request (Section 6.4)	•	
14	0	timestamp reply (Section 6.4)	•	
15	0	information request (obsolete)	•	
16	0	information reply (obsolete)	•	
17	0	address mask request (Section 6.3)	•	
18	0	address mask reply (Section 6.3)	•	

Figure 6.3 ICMP message types.

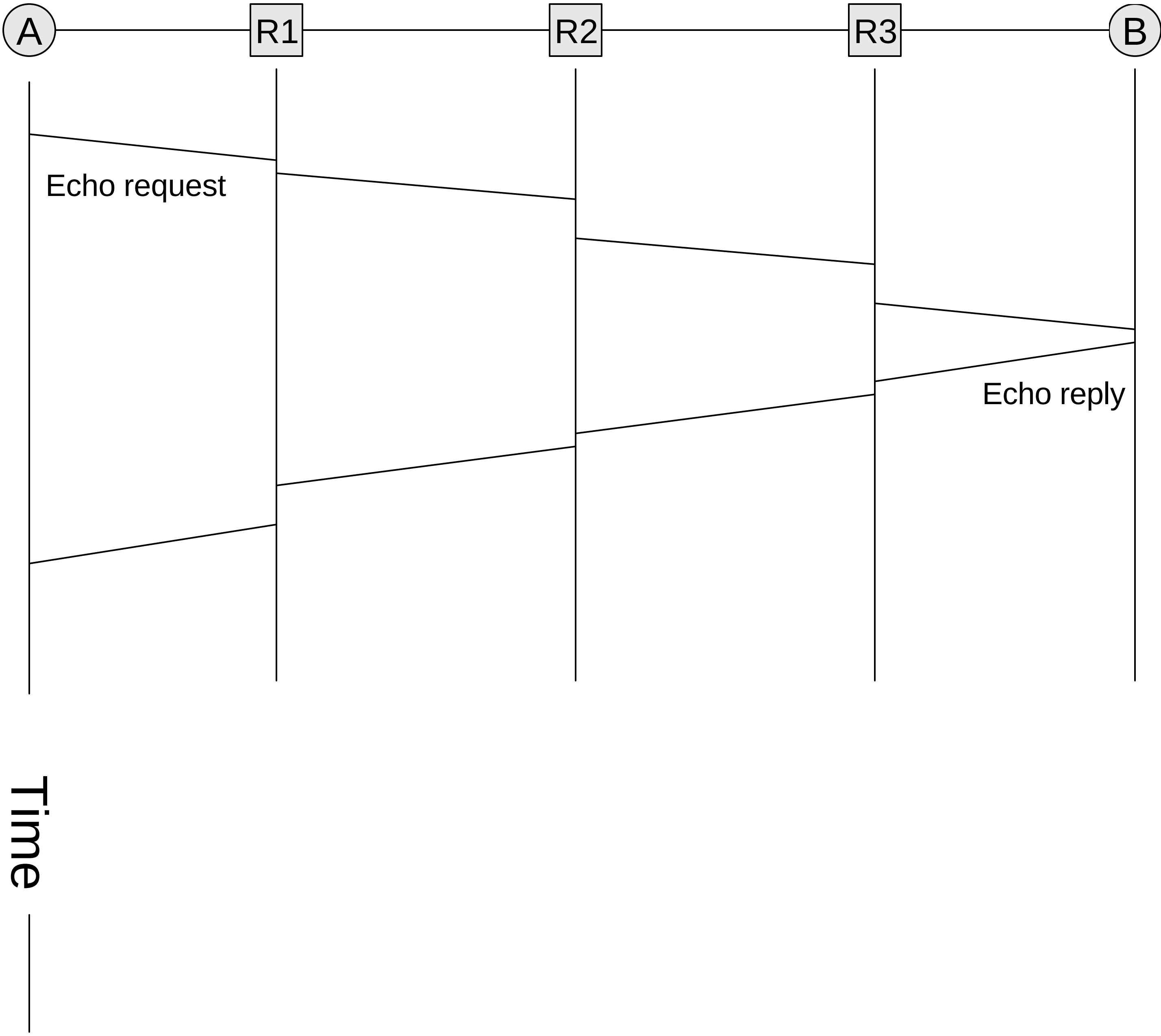
ICMP

- Echo request/reply
  - Can be used to check if a host is alive
- Address mask request/reply
  - Learn the subnet mask
- Destination unreachable
  - Invalid address and/or port
- Source quench
  - choke packet
- TTL expired
  - Routing loops, or too far away

ICMP (ping)

- Uses ICMP echo request/reply
- Source sends ICMP echo request message to the destination address
  - Echo request packet contains sequence number and timestamp
- Destination replies with an ICMP echo reply message containing the data in the original echo request message
- Source can calculate round trip time (RTT) of packets
- If no echo reply comes back then the destination is unreachable

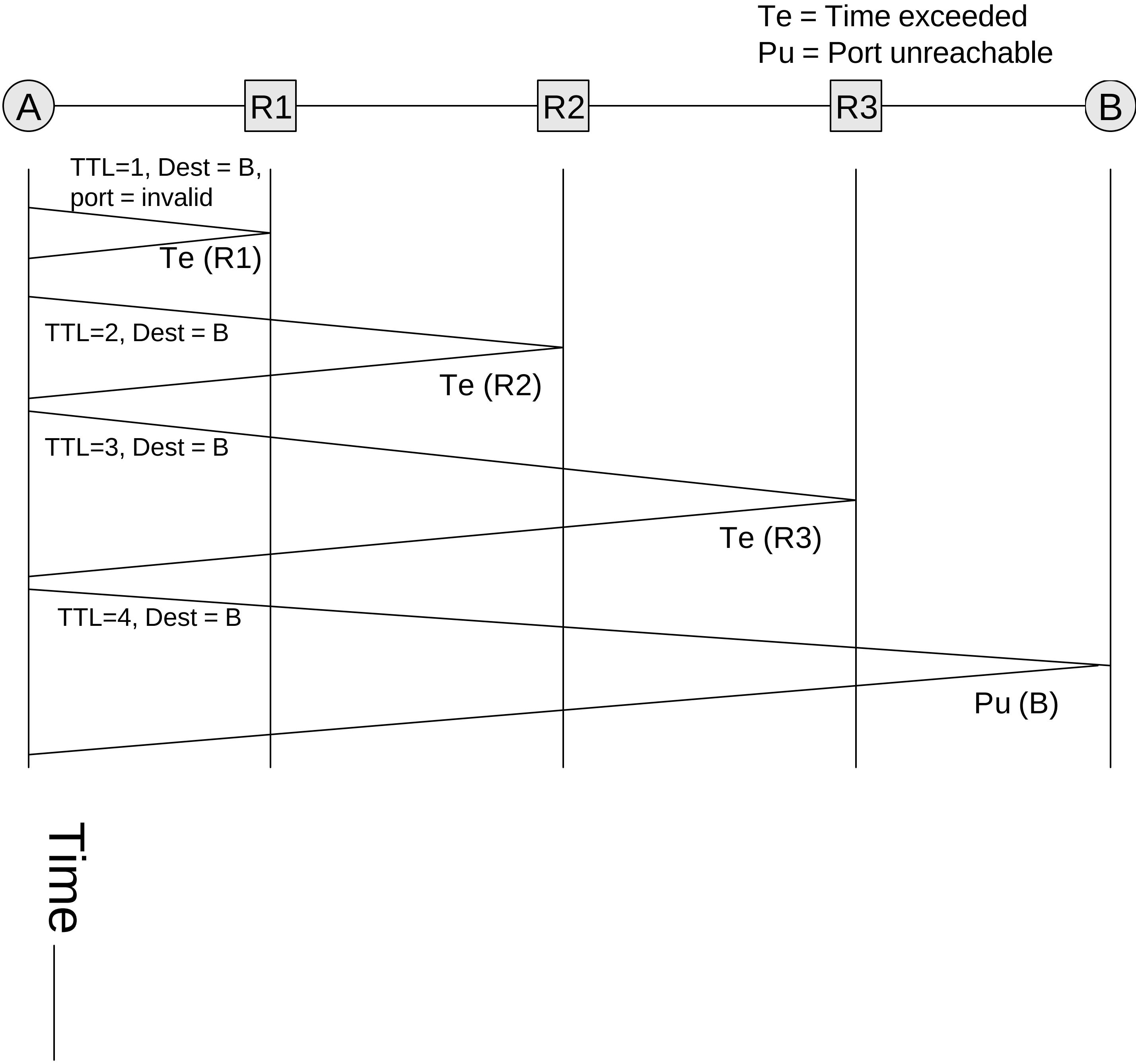
ping



traceroute

- Traceroute records the route that packets take
- A clever use of the TTL field
- When a router receives a packet, it decrements TTL
- If TTL=0, it sends an ICMP time exceeded message back to the sender
- To determine the route, progressively increase TTL
  - Every time an ICMP time exceeded message is received, record the sender’s (router’s) address
  - Repeat until the destination host is reached or an error message occurs

traceroute


Te = Time exceeded
Pu = Port unreachable
A
R1
R2
R3
B
TTL=1, Dest = B, port = invalid
Te (R1)
TTL=2, Dest = B
Te (R2)
TTL=3, Dest = B
Te (R3)
TTL=4, Dest = B
Pu (B)
Time

ARP

- Address Resolution Protocol
  - RFC 826
- Returns a MAC sublayer address when given an Internet address
- Commonly used in broadcast LANs so two hosts can communicate using IP addresses instead of MAC sublayer addresses

ARP

- Most machines on the Internet run ARP.
- ARP example:
  - Host A = 128.195.15.80
  - Host B = 128.195.15.81
    - Host A has a packet for Host B
    - Host A broadcasts an ARP packet asking for the Ethernet Address of Host B
    - Host B responds to Host A with its Ethernet Address

ARP
MAC Address

- Windows
  - – ipconfig /all
    - Returns the IP address of the host and the associated MAC address
- Example:

```
Ethernet adapter SMCPWRII1:
    Description . . . . . . . . . : SMC EtherPower II 10/100
    Physical Address. . . . . . . : 00-E0-29-14-49-D9
    DHCP Enabled. . . . . . . . . : No
    IP Address. . . . . . . . . . : 128.195.7.189
    Subnet Mask . . . . . . . . . : 255.255.255.0
    Default Gateway . . . . . . . : 128.195.7.1
    Primary WINS Server . . . . . : 128.195.7.70
    Secondary WINS Server . . . . : 128.195.4.70
```

ARP
MAC Address

- Unix
  - arp -a
    - Returns the arp table

Example

```
packrat.ics.uci.edu (128.195.1.16) at 8:0:20:85:6e:67
rodan.ics.uci.edu (128.195.1.64) at 8:0:20:9f:48:58
ics.uci.edu (128.195.1.1) at 8:0:47:0:1b:22
octavian.ics.uci.edu (128.195.1.18) at 8:0:20:1b:d4:6f
banzai.ics.uci.edu (128.195.1.3) at 8:0:20:87:3:c0
pacific.ics.uci.edu (128.195.1.20) at 8:0:20:77:c:2
annex5.ics.uci.edu (128.195.1.52) at 8:0:4c:0:21:6a
hindenburg.ics.uci.edu (128.195.1.25) at 8:0:20:83:4c:65
godzilla.ics.uci.edu (128.195.1.58) at 8:0:20:86:45:da
drivel.ics.uci.edu (128.195.1.13) at 8:0:20:75:7c:46
cs1-rsm.gw.nts.uci.edu (128.195.1.61) at 0:90:92:c8:7c:0
binky.ics.uci.edu (128.195.1.14) at 8:0:20:11:67:b9
```

RARP

- Reverse Address Resolution Protocol
  - RFC 903
- RARP performs the inverse action of ARP
- RARP returns an IP address for a given MAC sublayer address
- Operationally, RARP is the same as ARP

RARP

- Used for diskless clients
- RARP example:
  - Host A -No IP address
  - Host B - RARP server
    - Host A broadcasts its physical address to the subnet and requests its IP address
    - Host B responds with Host A’s IP address

RARP

- Disadvantages of RARP
  - Due to its broadcast nature, it is not routed.
- RARP alternative bootstrap protocol: BOOTP and DHCP

DHCP

- Dynamic Host Configuration Protocol
  - www.isc.org
    - Distributes a free version of DHCP
  - Routable
    - Utilizes a relay agent to forward dhcp requests
  - Does not require a MAC address to IP address table
    - Not required but may have one

DHCP

- Disadvantages of DHCP
  - Security
    - Will give any requesting host a valid IP address, unless MAC address registration or other controls are implemented
  - Host accountability

    Normally, only the requestor’s MAC address, date, and time are logged when dhcp grants a “lease”

      -Malicious users may take advantage of DHCP’s generosity.

Internet
Routing Protocols

- Interior Gateway Protocol
    - Routing algorithm *within* an AS
        - RIP
        - OSPF
- Exterior Gateway Protocol
    - Routing algorithm *between* AS
        - BGP

RIP

- Routing Information Protocol
  - RFC 1058
- One of the routing algorithms used by the Internet
- Based on distance vector routing (dervived from XNS (Xerox Network Systems) routing protocol)
- Most unix systems ship with RIP (routed)

RIP

- Did not scale well
  - suffers from the count-to-infinity problem
  - maximum path metric is 15
  - does not respond to load or delay

- RIP is slowly being phased out

RIP Distance Vector Routing

- RIP computes routes using a simple distance vector algorithm
  - Every hop in the network is assigned a cost (usually 1)
  - The total cost of a path is the sum of the hop costs
  - RIP chooses the next hop so the datagrams will follow a least-cost path.

RIP

- Positive aspects of RIP:
  - Simple
  - Available

- RIP shortcomings:
  - Maximum path metric is 15
    - Large networks may need > 16 hops
  - Cannot perform load balancing
  - Count to Infinity problem

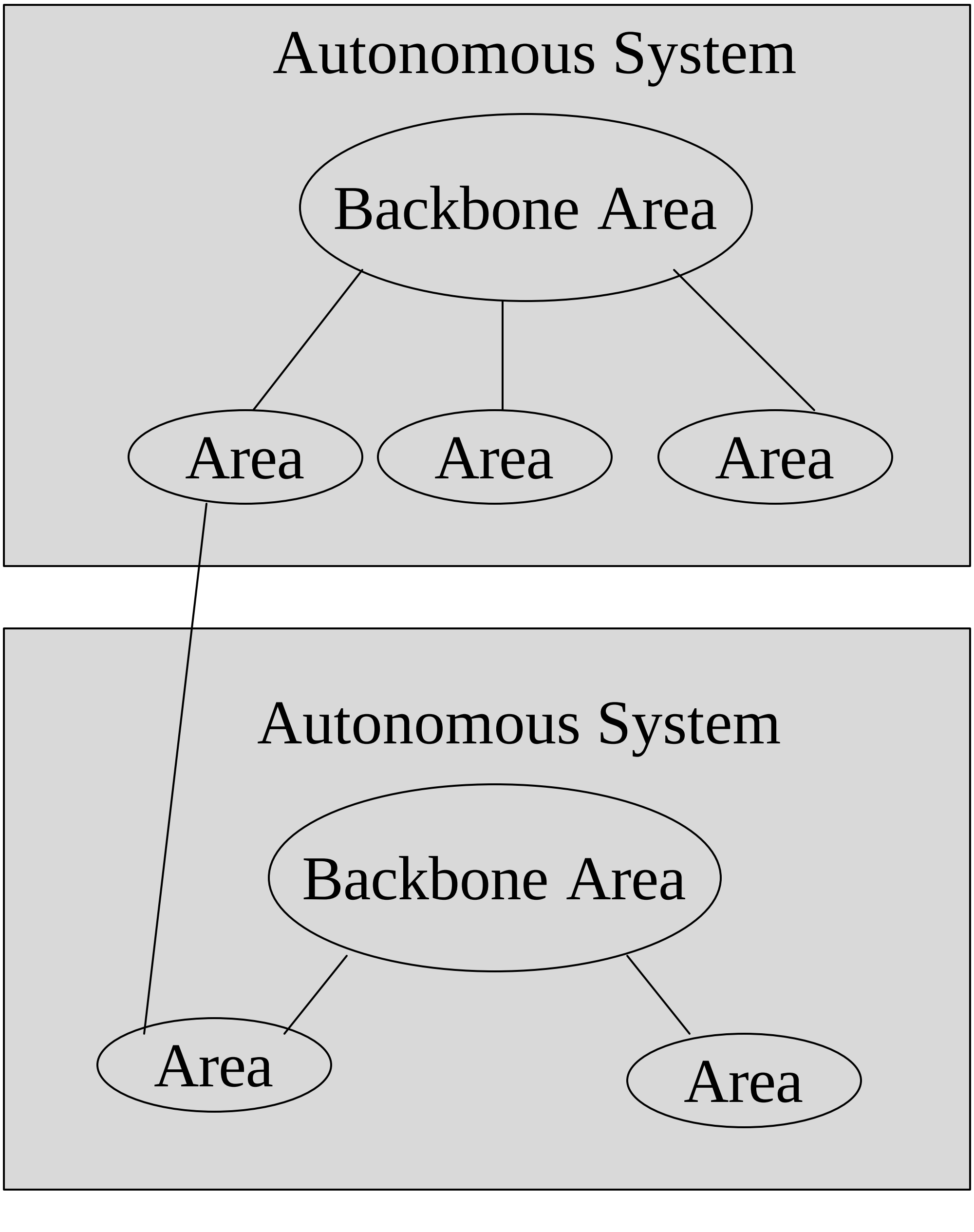
OSPF

- Open Shortest Path First
  - RFC 1247
- Routing algorithm used in the Internet within Autonomous Systems (AS)
  - Interior Gateway Protocol
- OSPF uses the Link State Routing algorithm with modifications to support:
  - Multiple distance metrics (geographical distance, delay, and throughput)
  - Support for real-time traffic
  - Hierarchical routing

OSPF

- OSPF divides the network into several hierarchies
  - Autonomous Systems (AS)
    - Groups of subnets
  - Areas
    - Groups of routers within an AS
  - Backbone Areas
    - Groups of routers that connect other areas together
    - Also known as Area 0

OSPF


Autonomous System
Backbone Area
Area
Area
Area
Autonomous System
Backbone Area
Area
Area

OSPF

- Routers are distinguished by the functions they perform
  - Internal routers
    - Only route packets within one area
  - Area Border routers
    - Connect areas together
  - Backbone routers
    - Reside only in the backbone area
  - AS boundary routers
    - Routers that connect to a router outside the AS

OSPF Example

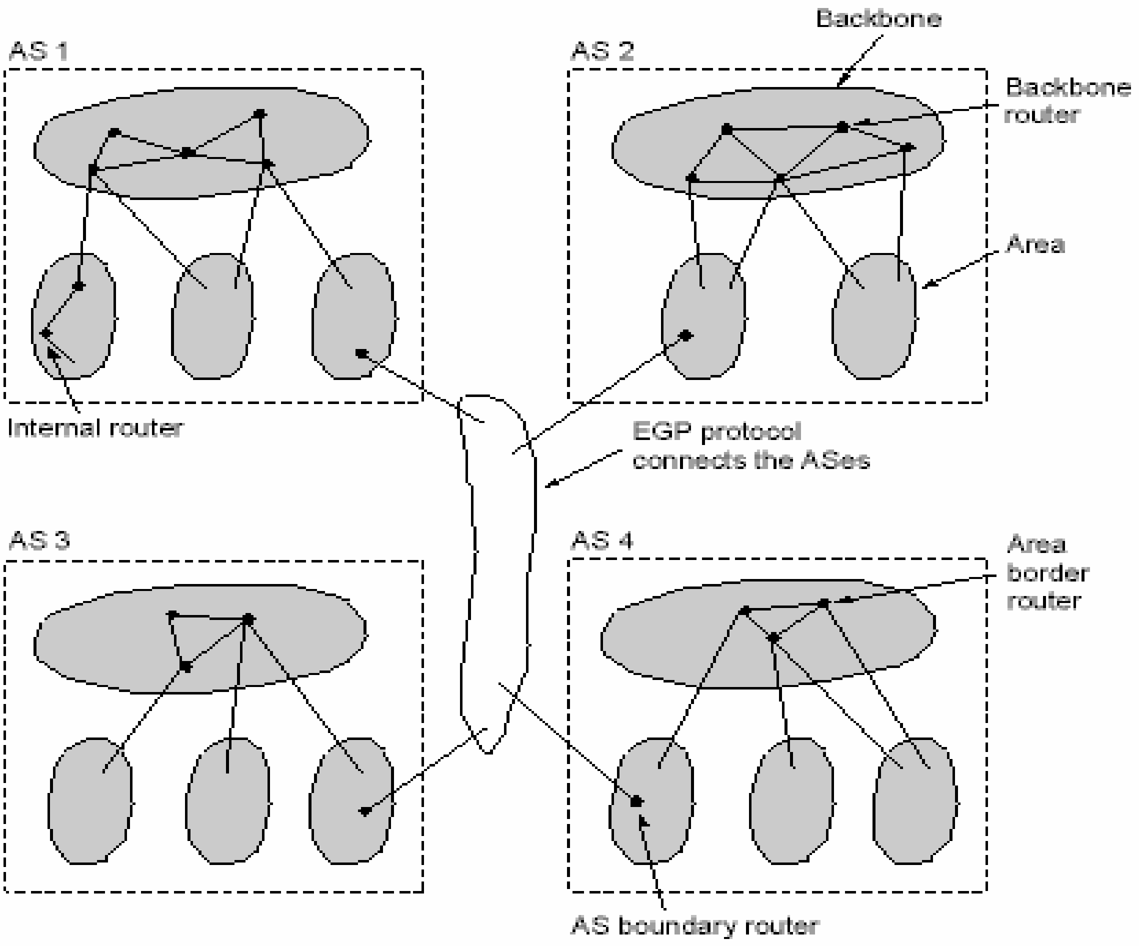


Fig. 5-53. The relation between ASes, backbones, and areas in OSPF.

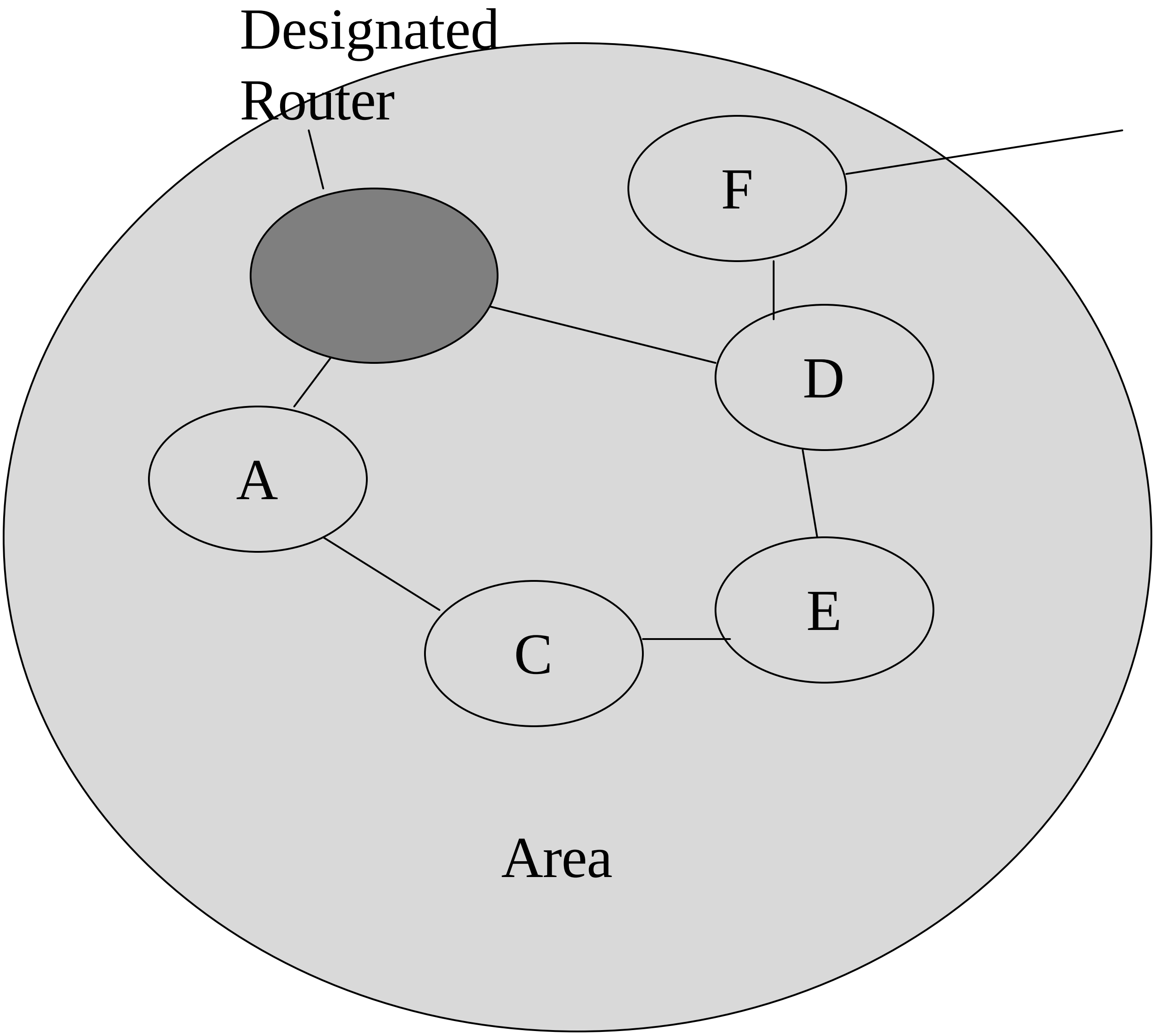
OSPF: Modified Link State Routing

- Recall:
  - In link state routing, routers flood their routing information to all other routers in the network
- In OSPF, routers only send their information to “adjacent routers”, not to all routers
- Adjacent does NOT mean nearest-neighbor in OSPF
- One router in each area is marked as the “designated router”
- Designated routers are considered adjacent to all other routers in the area

OSPF: Modified Link State Routing

- OSPF combines link state routing with centralized adaptive routing
  - Recall: centralized adaptive routing has a routing control center somewhere in the network.

OSPF: Adjacency



C is adjacent to B, but not to A or E

B is adjacent to ALL routers in the area

BGP

- Border Gateway Protocol
  - RFC 1771
- Routing algorithm used in the Internet between AS’s
  - Exterior Gateway Protocol

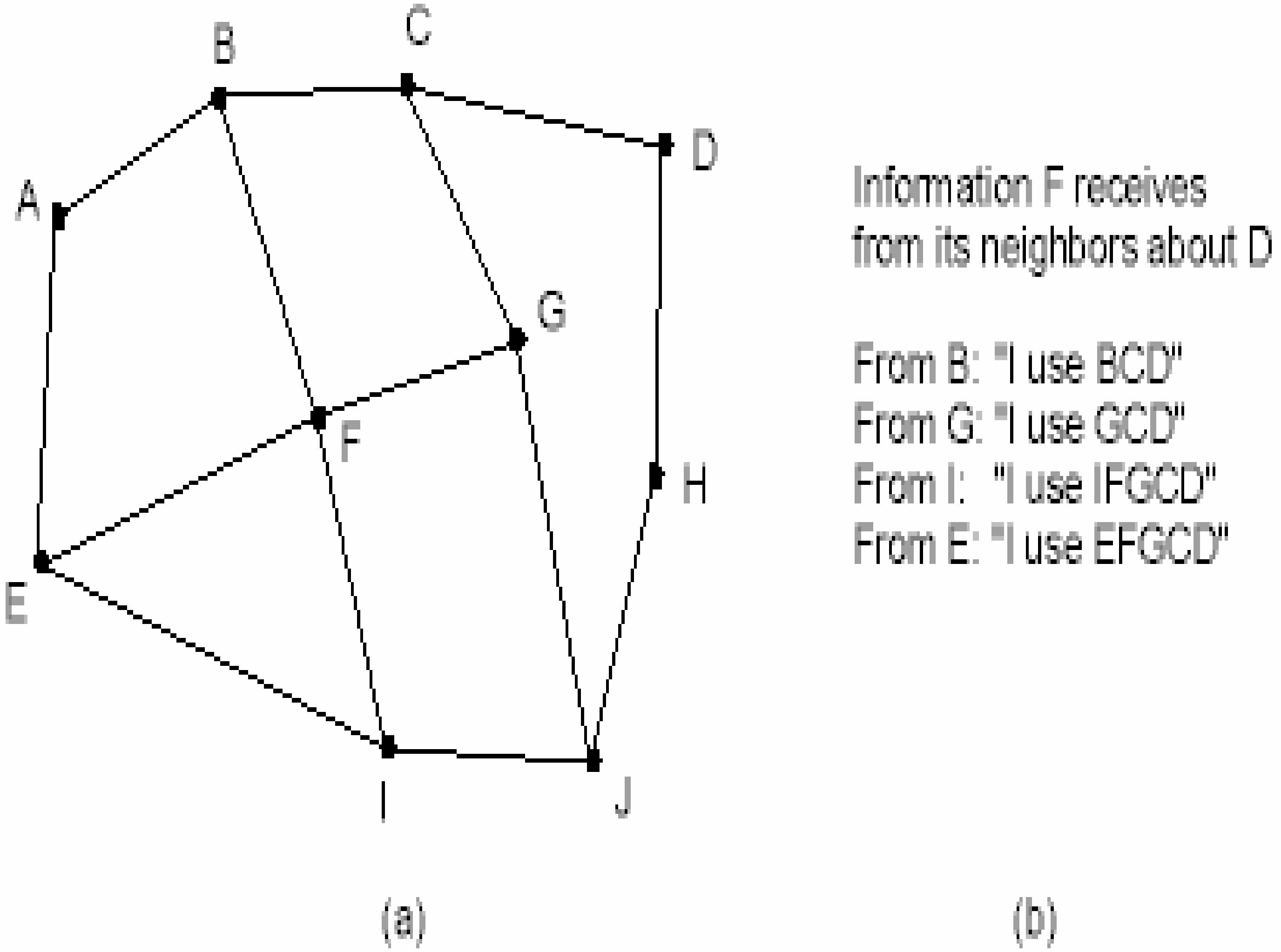
BGP

- Distance Vector protocol
  - transmits entire path to destination
- Contains manually configured routing polices with consideration given to:
  - politics
  - security
  - economics
    - Example:
      - Never allow traffic from IRAQ

BGP

- BGP routers exchange entire route information between themselves.
  - Done with Update Messages
  - Port 179
- Any route that violates policy constraints will not be used.

BGP Example



Information F receives
from its neighbors about D

From B: "I use BCD"
From G: "I use GCD"
From I: "I use IFGCD"
From E: "I use EFGCD"

Fig. 5-55. (a) A set of BGP routers. (b) Information sent to *F*.

Recent Developments: IPv6

- IPv4 (standard IP protocol) is limited
- Most importantly, IP is running out of addresses. 32 bits is not enough.
- Real-time traffic and mobile users are becoming common
- IP version 6 is also called IP-ng or IP Next Generation

IPv6: The Changes

- Large address space:
  - 128 bit address space (16 bytes)
    - Virtually unlimited addresses
- Fixed length headers
  - 40 byte headers
  - Improves the speed of packet processing in routers

IPv6: The Changes

- Support for “flows”
  - Flows help support real-time service in the Internet
  - A flow is a number in the IPv6 header that can be used by routers to see which packets belong to the same stream
  - Guarantees can then be assigned to certain flows
  - Example:
    - Packets from flow 10 should receive rapid delivery
    - Packets from flow 12 should receive reliable delivery

IPv6: The Changes

- Supports Anycasting
  - Like multicasting but delivery of packet is limited to just one host
    - Usually the nearest one
  - Telnet to an Anycast address and be routed to the closest server

IPv6: The Changes

- Header Checksum Removed
  - The data link layer and other upper layers already perform checksum calculations.

IPv6: The Changes

- Support for “Extension Headers:”
  - Fields immediately following the 40 byte fixed header giving additional IP information
    - Hop-by-Hop options
    - Routing
    - Fragmentation
    - Authentication
    - Encrypted Security Payload
    - Destination Options

IPv6

- When?
  - No set date for implementation
  - Will be phased in gradually
- Many aspects to transition:
  - Hardware:
    - Routers
    - Switches
  - Software
    - DNS
    - DHCP

IPv6
More Info

- RFC 1884
  - describes IPv6 addresses
- RFC 1883
  - outlines IPv6 protocol
- RFC 1885
  - ICMPv6
- RFC 1886
  - DNS extensions
- RFC 1887
  - address allocation