

# TCP/IP 通訊協定及應用

Spring 2002

中央大學 吳曉光博士

<http://wmlab.csie.ncu.edu.tw/course/tcp>

*We provide*  
無線網路多媒體實驗室  
*Wireless*  
*Wireless Network & Multimedia Laboratory*  
*Solution*

# Chapter 9: IP Routing

# Introduction

- ◆ Our interest here: How a single IP layer makes its routing decisions

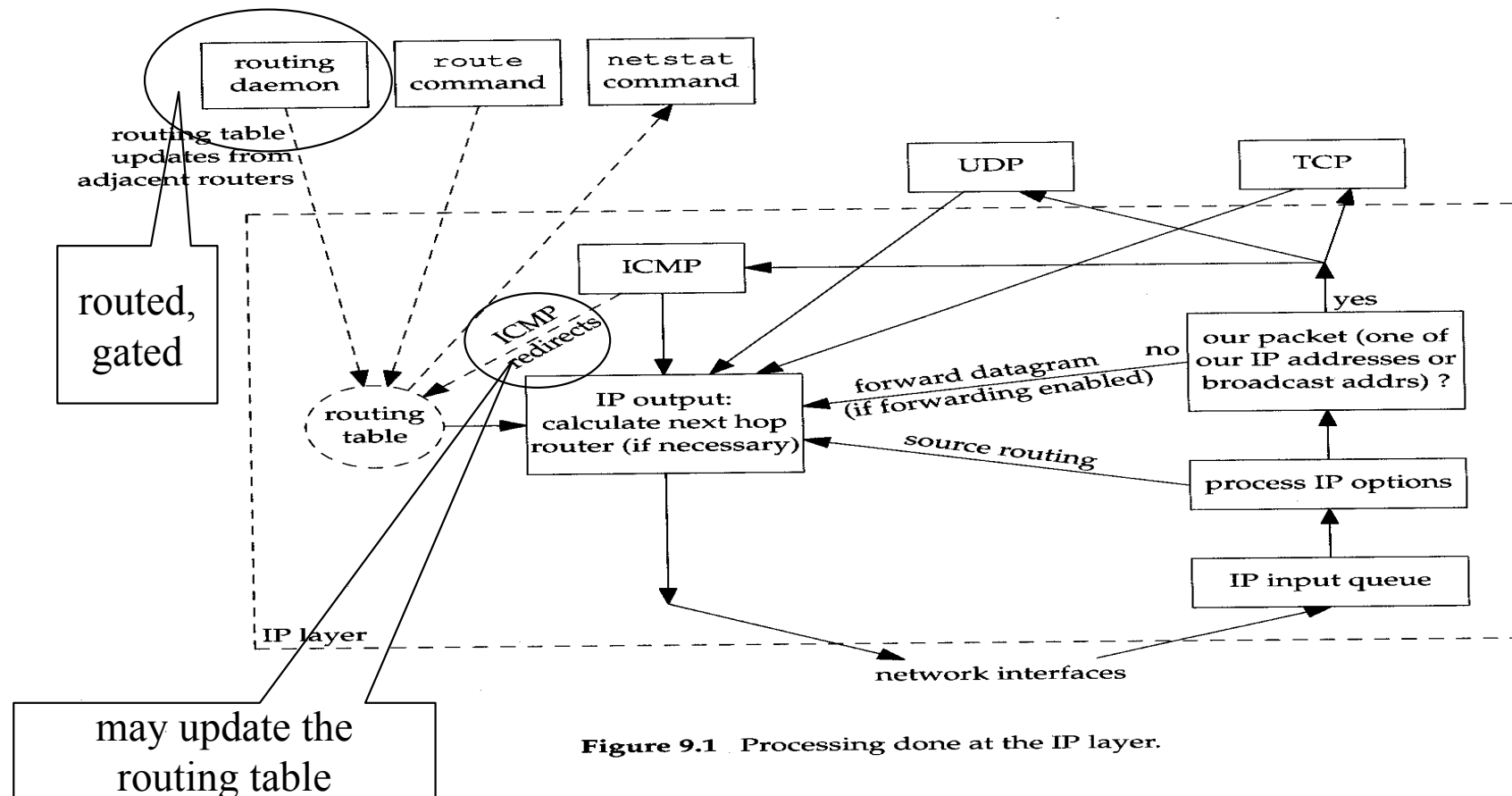
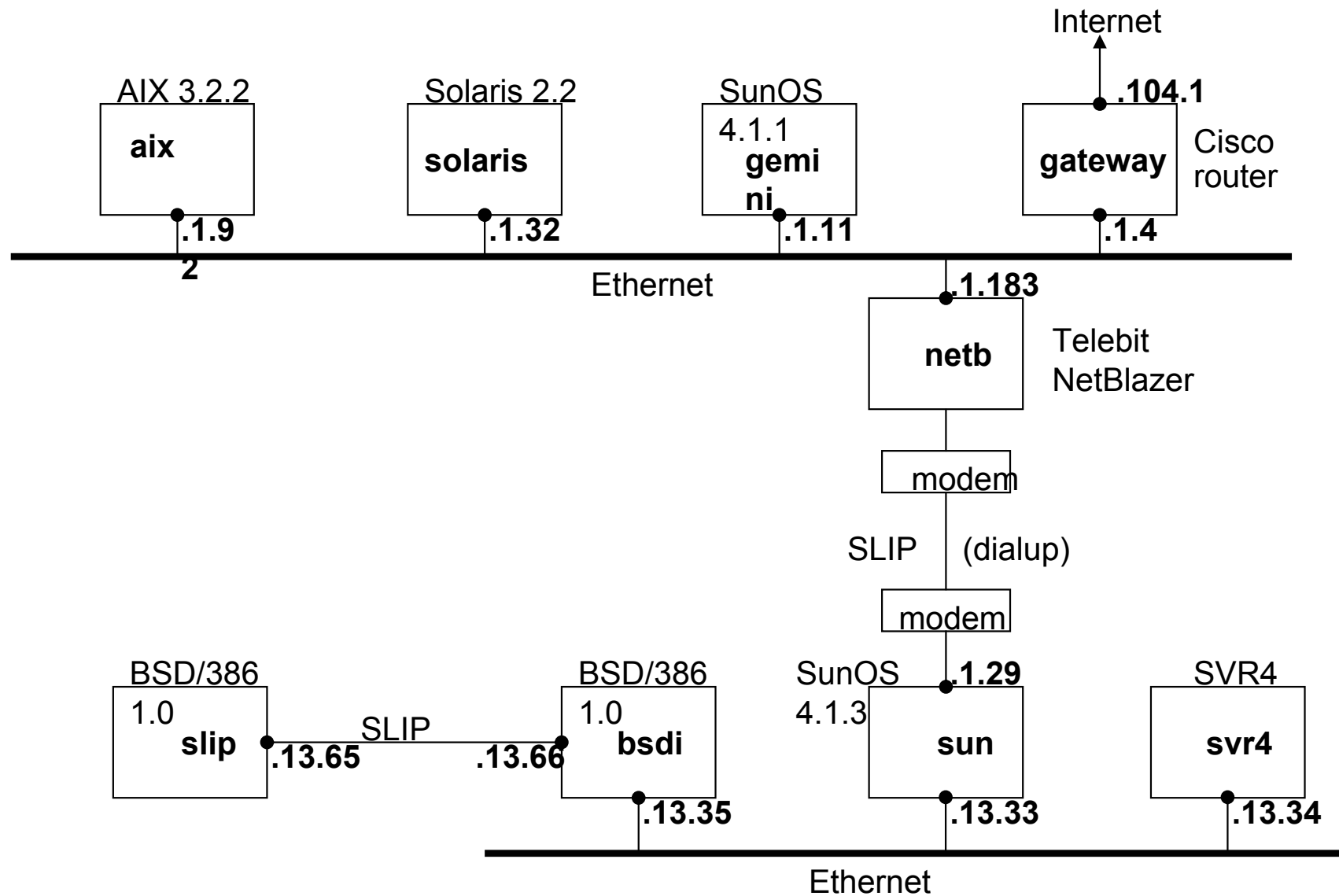
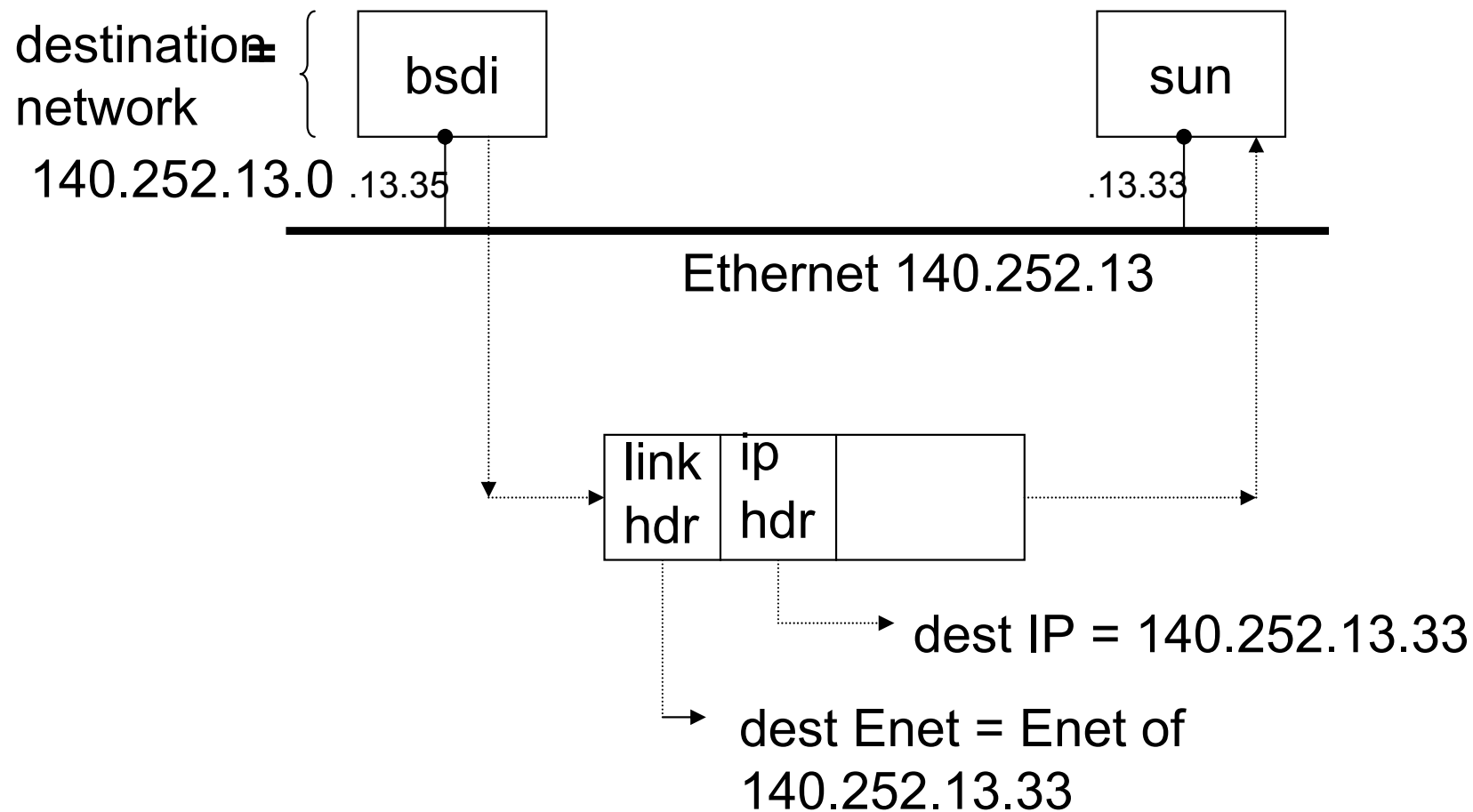


Figure 9.1 Processing done at the IP layer.

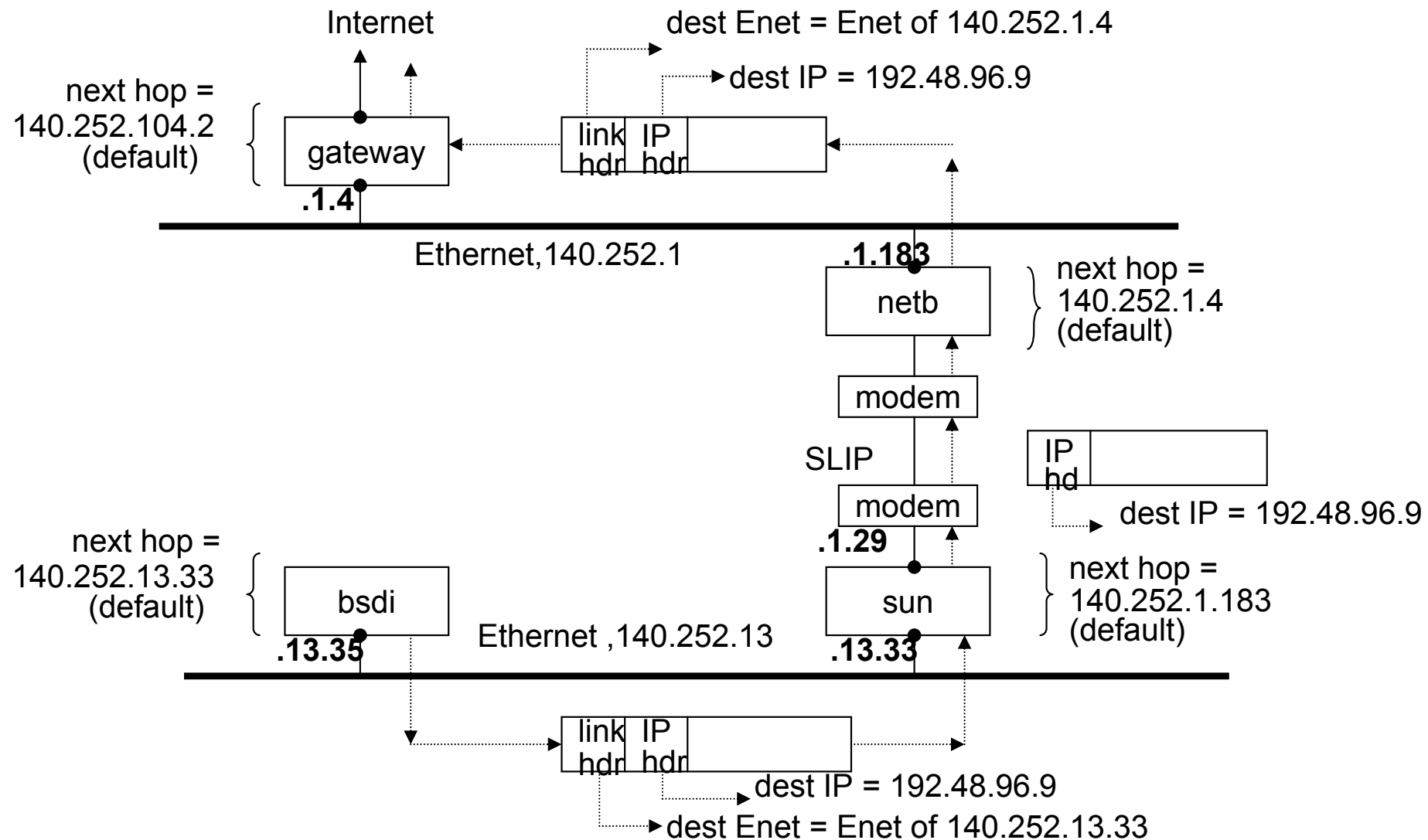
# Study Case



## Delivery of IP datagram from bsd1 to sun



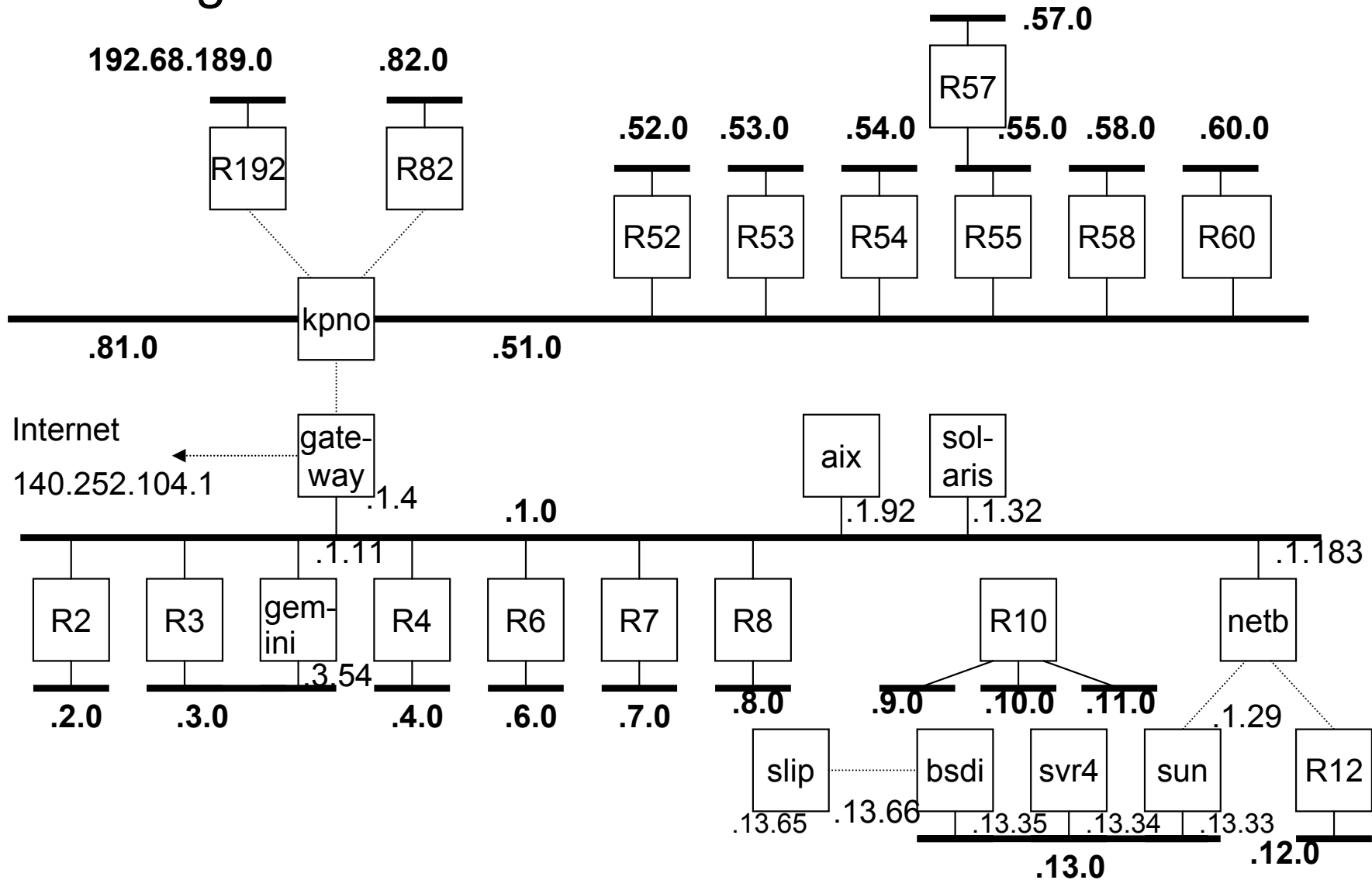
# Initial path of datagram from bsdi to ftp.uu.net(192.48.96.9)



# Subnetting a class B address

	16 bits	8 bits	8 bits
<b>Class B</b>	<b>netid = 140.252</b>	<b>subnetid</b>	<b>hostid</b>

# Arrangement of most of the **noao.edu** 140.252 subnets



# Routing Principles

- ◆ A matching host address is always used before a matching network address: routing mechanism:
  - 1. Search for a matching host address.
  - 2. Search for a matching network address.
  - 3. Search for a default entry.

# Simple Routing Table

- ◆ A example routing table:
- ◆ netstat options:
  - -r option: list the routing table
  - -n option: print IP address in numeric format

```
svr4 % netstat -rn
```

```
Routing tables
```

Destination	Gateway	Flags	Refcnt	Use	Interface
140.252.13.65	140.252.13.35	UGH	0	0	emd0
✓ 127.0.0.1	127.0.0.1	UH	1	0	lo0
default	140.252.13.33	UG	0	0	emd0
140.252.13.32	140.252.13.34	U	4	25043	emd0

# Simple Routing Table

◆ (continued...)

◆ Flags:

- U            the route is up
- **G**        the route is gateway (router)
- H            the route is to a host
- D            the route was created by a redirect
- M            the route was modified by a redirect

```
svr4 % netstat -rn
```

```
Routing tables
```

Destination	Gateway	Flags	Refcnt	Use	Interface
140.252.13.65	140.252.13.35	UGH	0	0	emd0
✓ 127.0.0.1	127.0.0.1	UH	1	0	lo0
default	140.252.13.33	UG	0	0	emd0
140.252.13.32	140.252.13.34	U	4	25043	emd0

# Simple Routing Table

- ◆ The G Flag:
  - not set for a direct route, else for indirect route
- ◆ A packet goes out a direct route:
  - both the IP address and the link-layer address specifying the destination
- ◆ A packet goes out a indirect route:
  - the IP address specifies the final destination but the link-layer address specifies the gateway(the next-hop router)
- ◆ The H Flag:
  - set for a host address, searched first by the routing principle

# Simple Routing Table

- ◆ The complexity of a routing table:
  - simplest: not connected to the network: only a single entry for the loopback interface
  - connected to a single LAN: only able to access hosts on that LAN: two entries, one for loopback, one for LAN
  - other networks are reachable through a single router: normally handled with a default entry pointing to that router
  - host-specific or network-specific routes are added

# Simple Routing Table

- ◆ illustration by a example (the host svr4):
  - the destination is the host sun, 140.252.13.33 ... not match any host entry ... matches 140.252.13.32 (the network IDs and subnet IDs match) ... so emd0 interface is used
  - the destination is the host slip, 140.252.13.65 ... matches the host entry ... and then indirect route
  - the destination is the host aw.com (192.207.117.2) ... not match any host entry ... not match any network entry ... searched for a default route ... then indirect route to the gateway 140.252.13.33 using the interface emd0

# Initializing a Routing Table

- ◆ an interface initialized, then automatically create a direct route in the routing table
- ◆ use route command to assign a route explicitly
  - Route add default sun 1
- ◆ run a routing daemon or use the newer router discovery protocol

# A More Complex Routing Table

```
sun % netstat -rn
Routing tables
Destination      Gateway          Flags    Refcnt  Use    Interface
140.252.13.65    140.252.13.35   UGH      0        171    le0
127.0.0.1        127.0.0.1       UH        1        766    lo0
140.252.1.183    140.252.1.29    UH        0         0      s10
default          140.252.1.183   UG        1       2955    s10
140.252.13.32    140.252.13.33   U         8      99551    le0
```

```
sun % ifconfig s10
s10: flags=1051<UP,POINTOPOINT,RUNNING>
      inet 140.252.1.29 --> 140.252.1.183 netmask ffffffff0
```

# No Route to Destination

- ◆ If there is NO default route:
  - for the IP datagram was generated on the host:: an error the application that sent the datagram: “host unreachable” or “network unreachable”
  - for the IP datagram is being forwarded: an ICMP host unreachable is sent back to original sender

# ICMP Host and Network Unreachable Errors



- ◆ The ICMP “host unreachable” error message is sent by a router when it receives an IP datagram that it cannot deliver or forward:

```
svr4 % ping gemini
ICMP Host Unreachable from gateway sun (140.252.13.33)
ICMP Host Unreachable from gateway sun (140.252.13.33)
^?                                     type interrupt key to stop
```

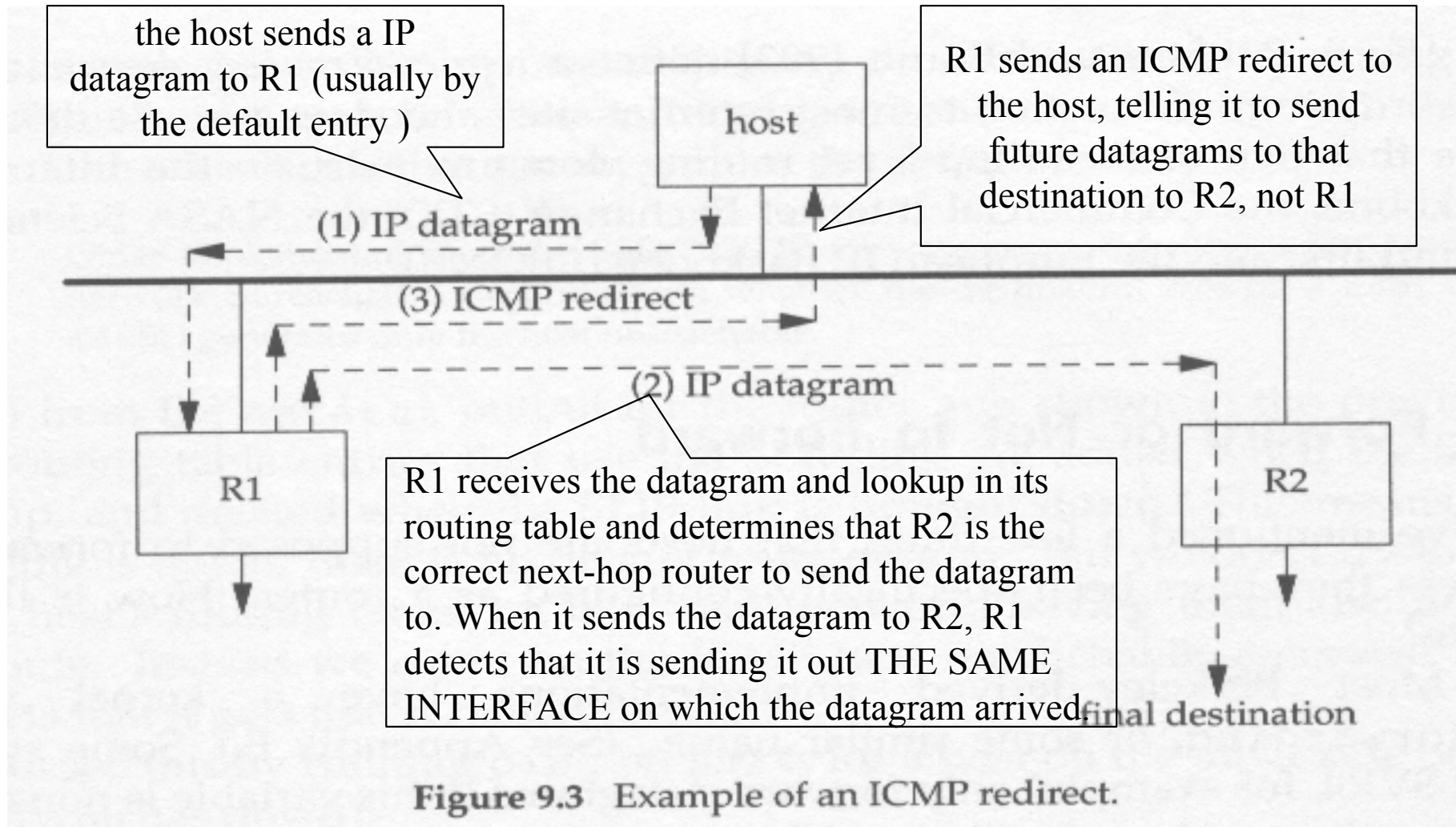
Test how far to know the host is not connected to the Internet: in this case about 6 routers

```
sun % ping 192.82.148.1                this IP address is not connected to the Internet
PING 192.82.148.1: 56 data bytes
ICMP Host Unreachable from gateway enss142.UT.westnet.net (192.31.39.21)
for icmp from sun (140.252.1.29) to 192.82.148.1
```

# To Forward or Not to Forward

- ◆ Hosts are not supposed to forward IP datagrams unless they have been specifically configured as a router:
  - Most Berkeley-derived implementations have a kernel variable named `ipforwarding`, or some similar name
  - Older 4.2 BSD hosts forwarded datagrams by default, which caused lots of problems for systems configured improperly. That's why this kernel option must always default to "never forward" unless the system administrator specifically enables forwarding

# ICMP Redirect Errors



# ICMP Redirect Errors

## ◆ An Example:

```
solaris % netstat -rn
```

Routing Table:

Destination	Gateway	Flags	Ref	Use	Interface
-----	-----	-----	-----	-----	-----
127.0.0.1	127.0.0.1	UH	0	848	lo0
140.252.1.0	140.252.1.32	U	3	15042	le0
224.0.0.0	140.252.1.32	U	3	0	le0
default	140.252.1.4	UG	0	5747	

```
solaris % ping -sv bsdi
```

PING bsdi: 56 data bytes

ICMP Host redirect from gateway gateway (140.252.1.4)

to netb (140.252.1.183) for bsdi (140.252.13.35)

64 bytes from bsdi (140.252.13.35): icmp\_seq=0. time=383. ms

64 bytes from bsdi (140.252.13.35): icmp\_seq=1. time=364. ms

64 bytes from bsdi (140.252.13.35): icmp\_seq=2. time=353. ms

^?

*type interrupt key to stop*

----bsdi PING Statistics----

4 packets transmitted, 3 packets received, 25% packet loss

round-trip (ms) min/avg/max = 353/366/383

# ICMP Redirect Errors

- ◆ (continued...)

```
solaris % netstat -rn
```

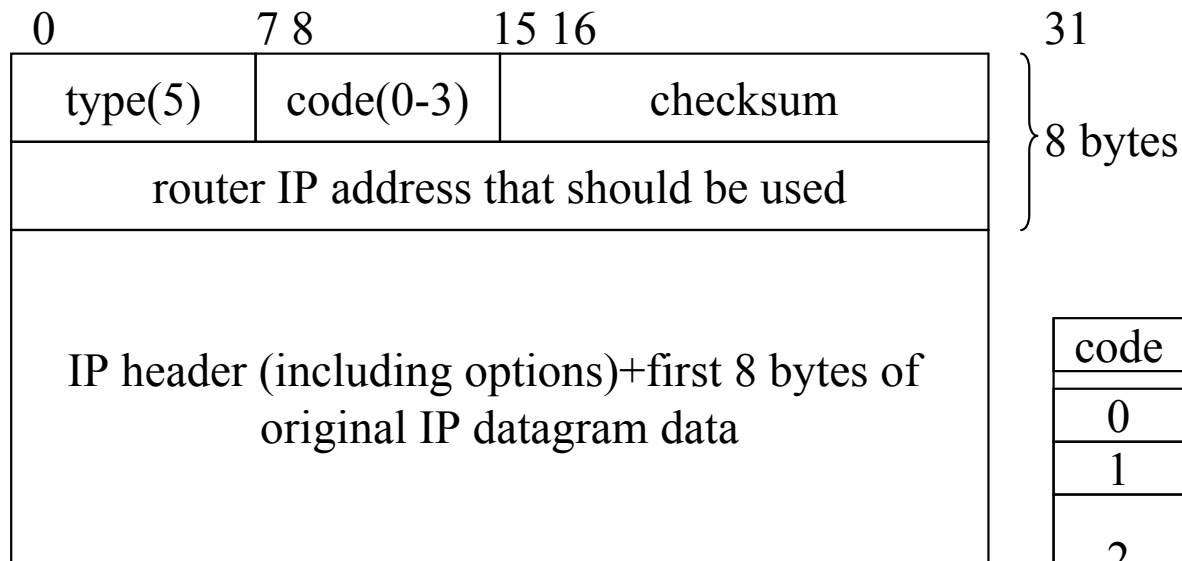
Routing Table:

Destination	Gateway	Flags	Ref	Use	Interface
127.0.0.1	127.0.0.1	UH	0	848	lo0
140.252.13.35	140.252.1.183	UGKD	0	2	
140.252.1.0	140.252.1.32	U	3	15045	le0
224.0.0.0	140.252.1.32	U	3	0	le0
default	140.252.1.4	UG	0	5749	

the route was installed by an ICMP redirect

# ICMP Redirect Errors

## ◆ More Details:



code	Description
0	redirect for network
1	redirect for host
2	redirect for type-of-service and network
3	redirect for type-of-service and host

the rules about ICMP redirects:

1 generated only by routers, not by hosts

2 intended to be used by hosts, not routers

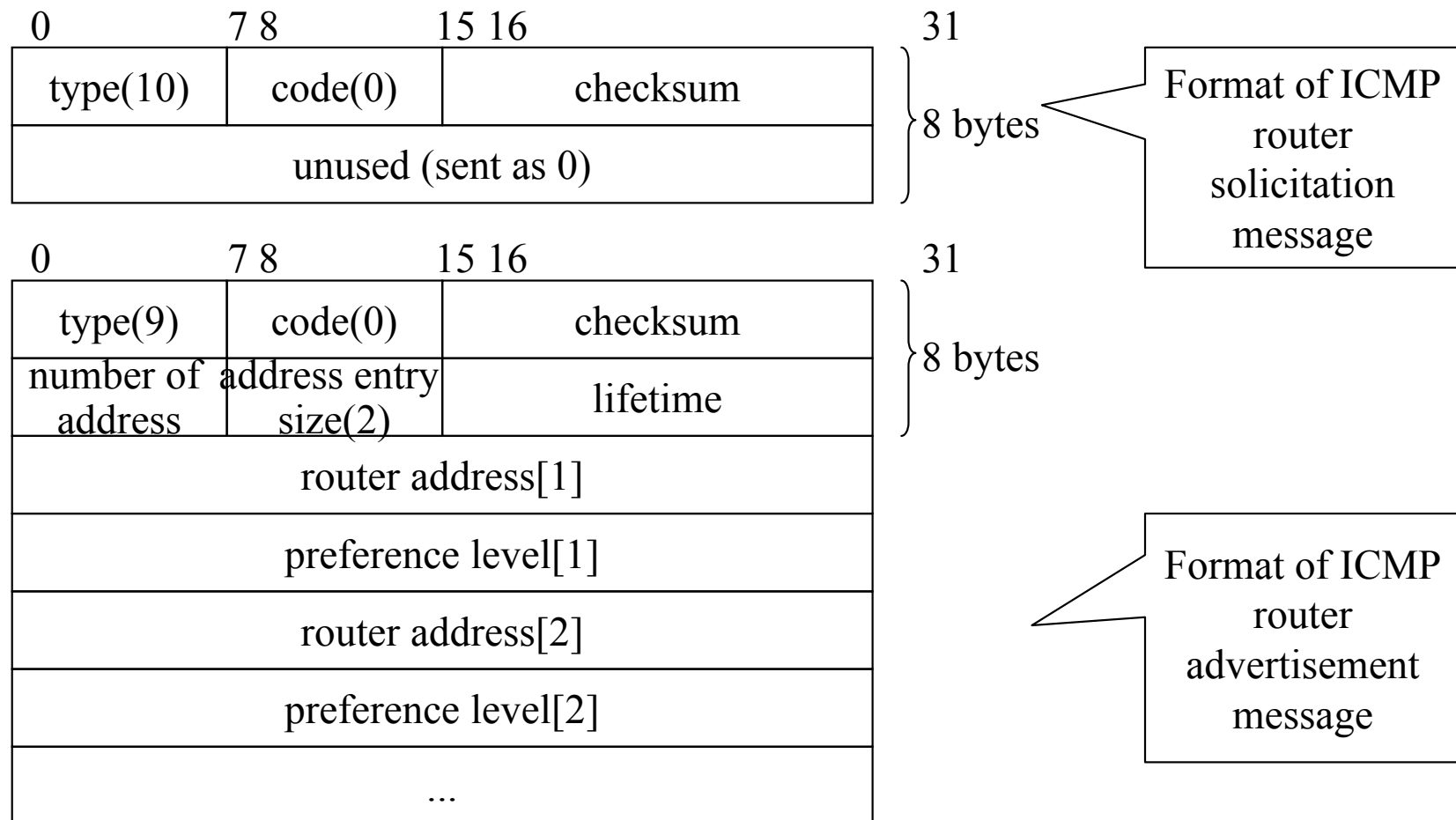
3 the routing protocol should obviate the need for redirects

# ICMP Router Discovery Messages

- ◆ Initialize a routing table by using newer ICMP router advertisement and solicitation messages:
  - bootstrapping ... broadcast/multicast solicitation message ... one or more routers responds with a router advertisement message
  - Additionally, the routers periodically broadcast/multicast their router advertisements, allowing any hosts that are listening to update their routing table accordingly
  - RFC 1256 specifies the format of these two ICMP messages:

# ICMP Router Discovery Messages

◆ (continued...)



# ICMP Router Discovery Messages

## ◆ Router Operation:

- When a router starts up it transmits periodic advertisements on all interfaces capable of broadcasting or multicasting
- these advertisements are not exactly periodic, but are randomized

## ◆ Host Operations:

- Upon bootstrap a host normally transmit three router solicitations, 3 seconds apart.
- A host also listens for advertisements from adjacent routers.

## ◆ Implementation:

- The router discovery message are normally generated by and processed by a user process (a daemon)

# Summary

- ◆ How a single IP layer makes its routing decisions
- ◆ Routing principles
- ◆ ICMP host and network unreachable errors
- ◆ ICMP redirect errors
- ◆ ICMP Router Discovery Messages