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

Chapter 10: Dynamic Routing Protocols

Introduction

- When do we using dynamic routing?
  - If any of below three conditions is false, dynamic routing is used
    - The network is small
    - A single connection point to other network
    - No redundant routes
- Dynamic Routing
  - It occurs when routers talk to adjacent routers, informing each other of what networks each router is currently conneted to
  - What‘s the routing daemon
    - The process is running protocol, communicating with its neighbor routers
  - If the daemon finds multiple routes to a destination, the daemon choose best best one to insert into the kernel‘s table
  - If the daemon finds that a link has gone down, it can delete routes or alternate routes that bypass the problem

Unix Routing Daemons

- Routing Daemons in Unix System:
  - Unix systems often run the routing daemon named routed
  - An alternative program is gated that supports both IGPs and EGPs

Daemon	Interior Gateway Protocol			Exterior Gateway Protocol	
	HELLO	RIP	OSPF	EGP	BGP
routed		V1			
gated, Version 2	•	V1		•	V1
gated, Version 3	•	V1, V2	V2	•	V2, V3

Figure 10.1 Routing protocols supported by routed and gated.

- RIP (Routing Information Protocol) message format

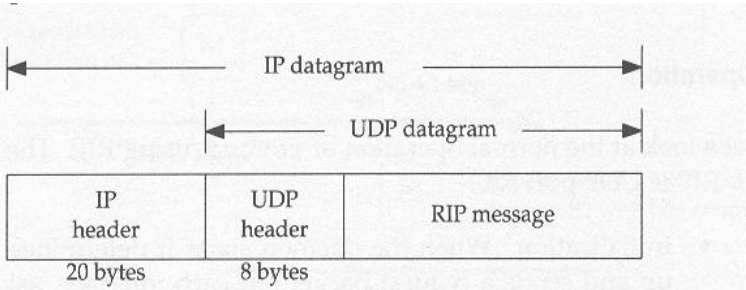

Figure 10.2 RIP message encapsulated within a UDP datagram.

RIP: Routing Information Protocol

- Format of a RIP message

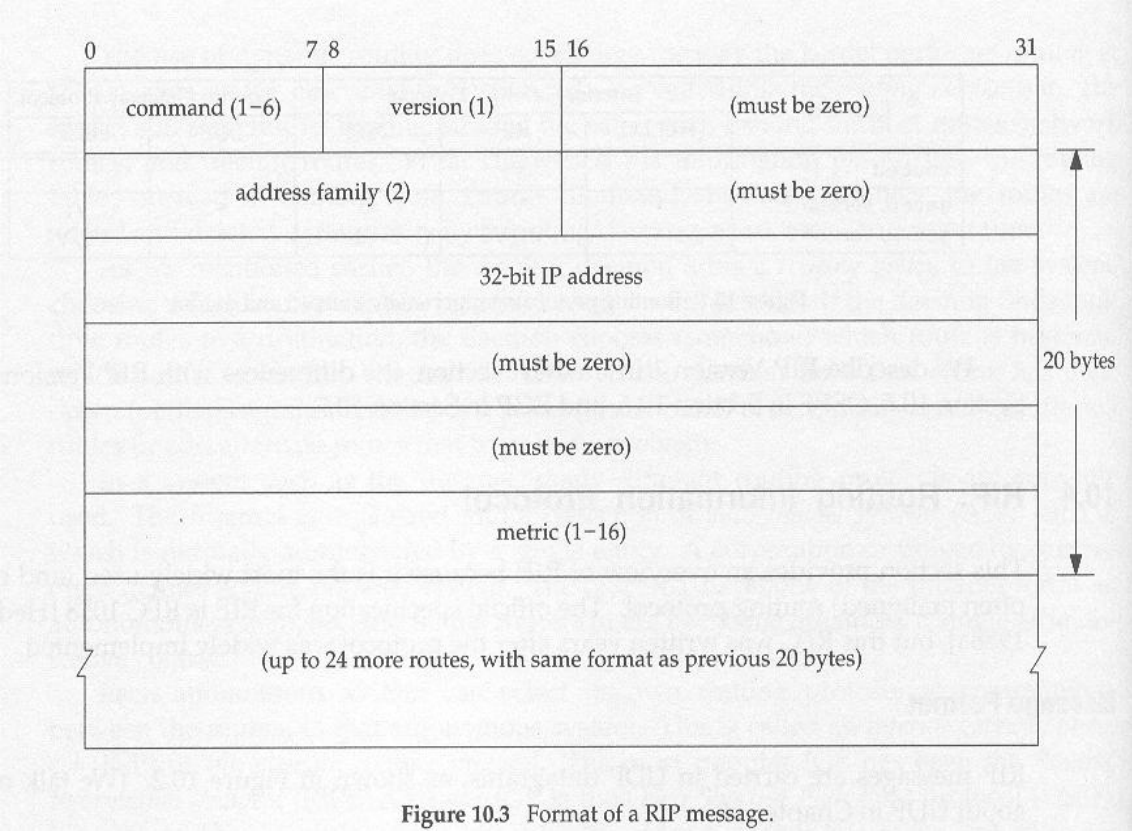

Figure 10.3 Format of a RIP message.

RIP: Routing Information Protocol

- Normal Operation
  - The well-know port number for RIP is UDP port 520
  - Initialization: The daemon send a request packet out each interface, asking for the other router's complete routing table
  - Request received: If we have a route to the specified address, set the metric to our value, else set the metric to 16
  - Response received: The response is validated and may update the routing table
  - Regular routing updates: Every 30 seconds, routing table is sent to every neighbor router
  - Triggered updates: Only those entries that have changed must be transmitted
  - A route has not been updated for 3-min, metric is set to infinity(16)

RIP: Routing Information Protocol

- What's metrics
  - The metrics used by RIP are hop counts.The hop count for all directly connected interfaces is 1

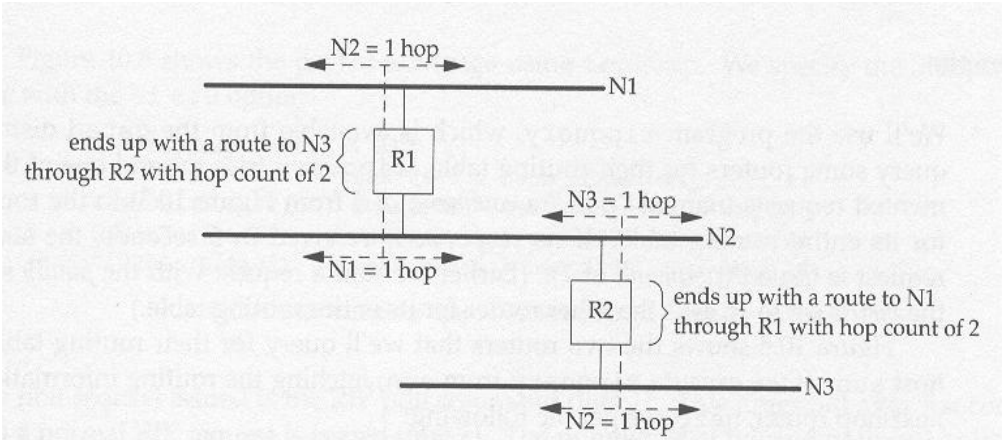

Figure 10.4 Example routers and networks.

  - The metric to N1 for R2 is 2, as is the metric to N3 for R1
  - The router selects the path with the smallest hop count and ignores the other paths
  - maximum number of hops between hosts is 15

RIP: Routing Information Protocol

- Drawbacks of RIP
  - RIP has no knowledge of subnet addressing
  - It takes a long time to stabilize after the failure of a router or a link
  - A maximum of 15 for the metric limits sometime is not enough
- Example 1
  - ripquery tries to send one of the undocumented requests (name “poll”,a command of 5)
  - If no response is received in 5 seconds, the standard RIP request is issued (command of 1)

```
sun % ripquery -n netb
504 bytes from netb (140.252.1.183):      first message contains 504 bytes
                                          lots of other lines deleted
        140.252.1.0, metric 1             the top Ethernet in Figure 10.5
        140.252.13.0, metric 1            the bottom Ethernet in Figure 10.5
244 bytes from netb (140.252.1.183):      second message with remaining 244 bytes
                                          lots of other lines deleted
```

RIP: Routing Information Protocol

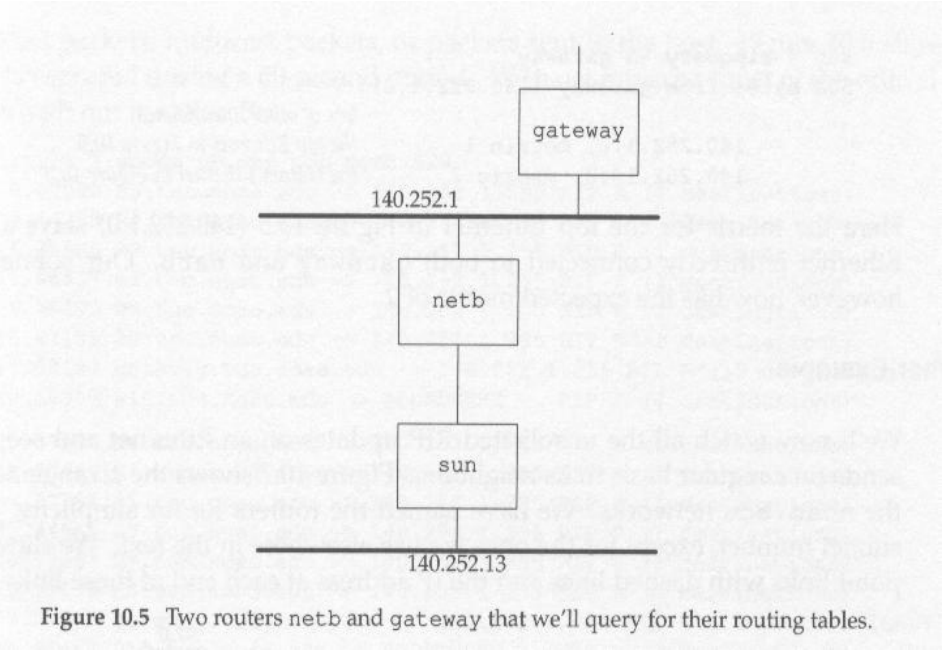

Figure 10.5 Two routers `netb` and `gateway` that we'll query for their routing tables.

```
   sun % tcpdump -s600 -i sl0
1  0.0                    sun.2879 > netb.route: rip-poll 24
2  5.014702 (5.0147)      sun.2879 > netb.route: rip-req 24
3  5.560427 (0.5457)      netb.route > sun.2879: rip-resp 25:
4  5.710251 (0.1498)      netb.route > sun.2879: rip-resp 12:
```

Figure 10.6 `tcpdump` output while running `ripquery` program.

RIP: Routing Information Protocol

```
sun % ripquery -n gateway
504 bytes from gateway (140.252.1.4):
                                        lots of other lines deleted
        140.252.1.0, metric 1           the top Ethernet in Figure 10.5
        140.252.13.0, metric 2          the bottom Ethernet in Figure 10.5
```

- Here the metric in 140.252.1.0 stays at 1, since that Ethernet is directly connected to both gateway and netb
- Our subnet 140.252.13.0 has the expected metric of 2

- Example 2
  - We'll run the Solaris 2.x program snoop, which is similar to tcpdump, on the host solaris
  - Figure 10.8 shows the packets captured during a 60-second period

RIP: Routing Information Protocol

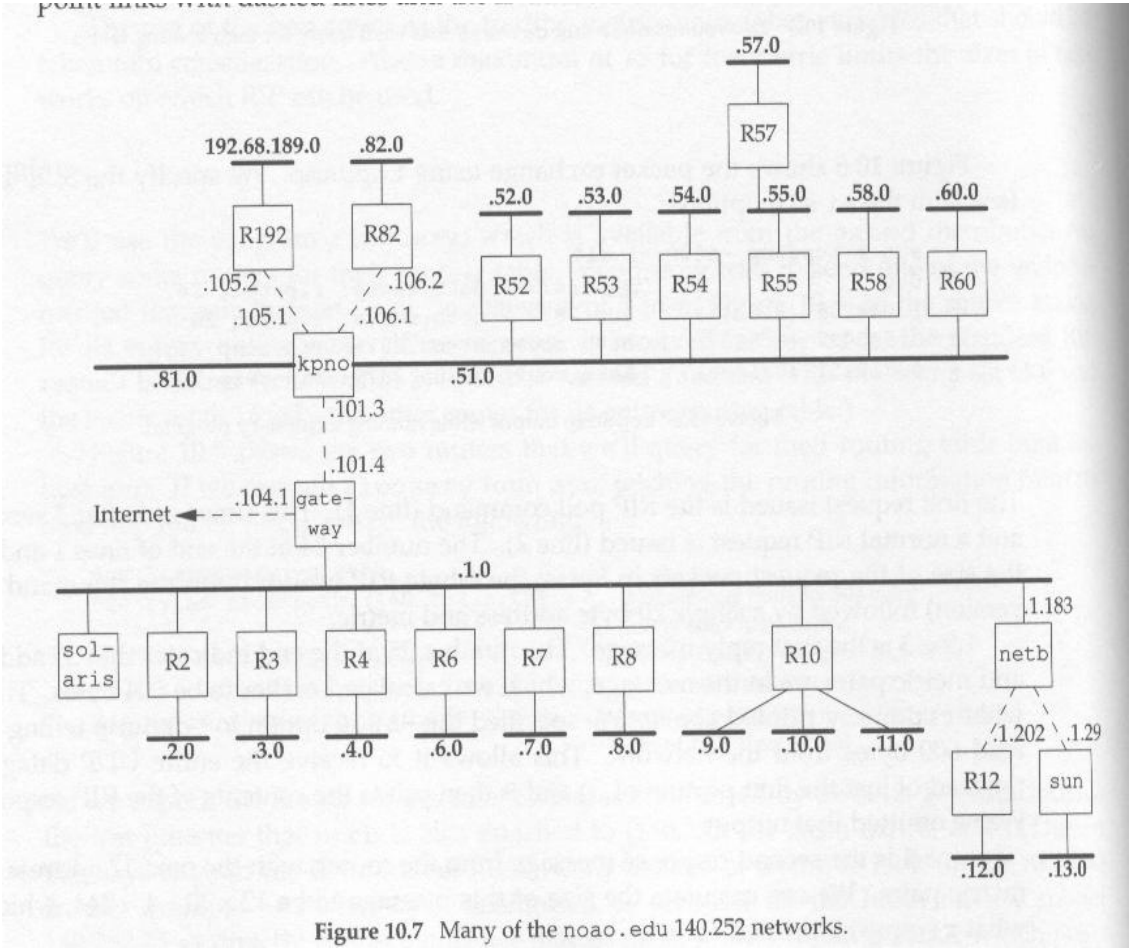

Figure 10.7 Many of the noao.edu 140.252 networks.

RIP: Routing Information Protocol

```
solaris % snoop -P -tr udp port 520
  0.00000 R6.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
  4.49708 R4.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
  6.30506 R2.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 11.68317 R7.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 16.19790 R8.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 16.87131 R3.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 17.02187 gateway.tuc.noao.edu -> 140.252.1.255 RIP R (15 destinations)
 20.68009 R10.tuc.noao.edu -> BROADCAST    RIP R (4 destinations)

 29.87848 R6.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 34.50209 R4.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 36.32385 R2.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 41.34565 R7.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 46.19257 R8.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 46.52199 R3.tuc.noao.edu -> 140.252.1.255 RIP R (1 destinations)
 47.01870 gateway.tuc.noao.edu -> 140.252.1.255 RIP R (15 destinations)
 50.66453 R10.tuc.noao.edu -> BROADCAST    RIP R (4 destinations)
```

Figure 10.8 RIP broadcasts captured at solaris over a 60-second period.

- Why R10 is advertising four networks ?

```
RIP:  Address          Metric
RIP:  140.251.0.0      16 (not reachable)
RIP:  140.252.9.0      1
RIP:  140.252.10.0     1
RIP:  140.252.11.0     1
```

RIP: Routing Information Protocol

- The router getway advertises 15 routes.We can run snoop with the -v flag and see the entire contents of the RIP message

```
solaris % snoop -P -v -tr udp port 520 host gateway
                                                    many lines deleted
RIP:  Opcode = 2 (route response)
RIP:  Version = 1
RIP:  Address          Metric
RIP:  140.252.101.0    1
RIP:  140.252.104.0    1
RIP:  140.252.51.0     2
RIP:  140.252.81.0     2
RIP:  140.252.105.0    2
RIP:  140.252.106.0    2
RIP:  140.252.52.0     3
RIP:  140.252.53.0     3
RIP:  140.252.54.0     3
RIP:  140.252.55.0     3
RIP:  140.252.58.0     3
RIP:  140.252.60.0     3
RIP:  140.252.82.0     3
RIP:  192.68.189.0     3
RIP:  140.252.57.0     4
```

Figure 10.9 RIP response from gateway.

RIP Version 2

- Compare RIP and RIP-2

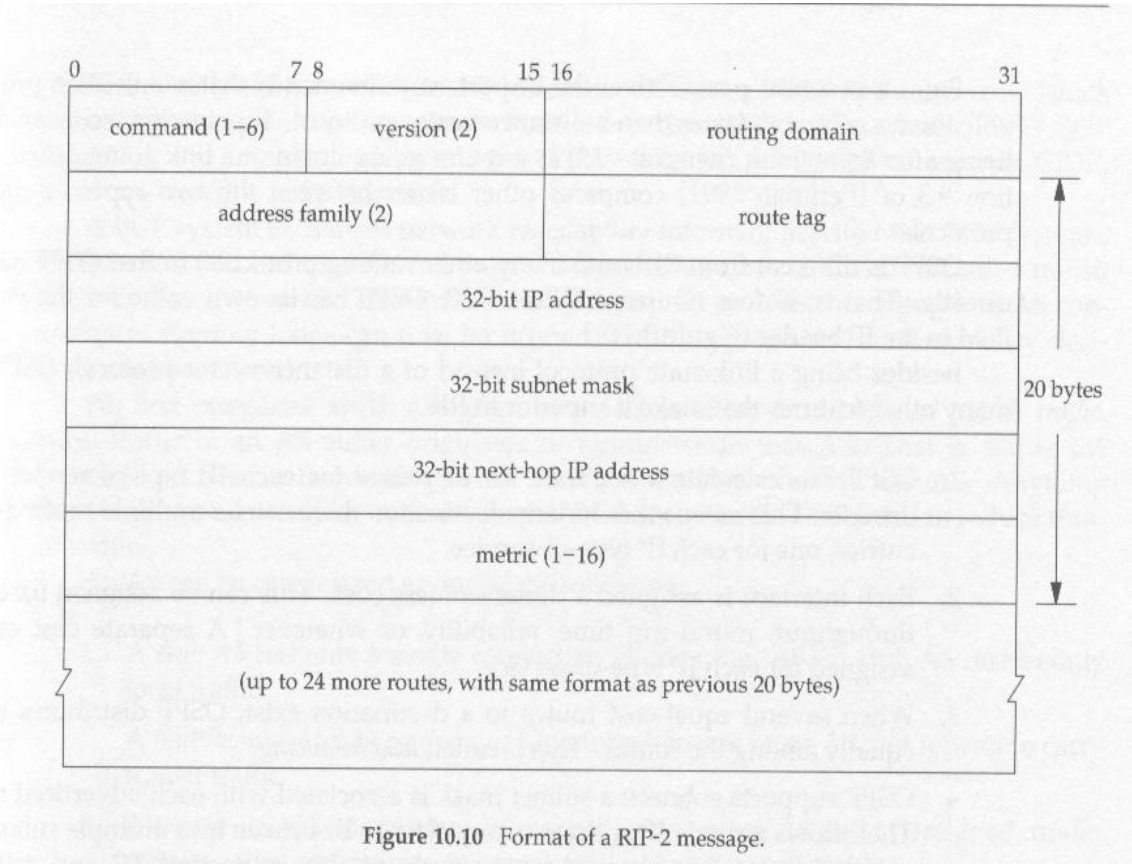

Figure 10.10 Format of a RIP-2 message.

- RIP-2 supports multicasting.This can reduce the load on hosts that not listening for RIP-2 messages

OSPF: Open Shortest Path First

- The difference of OSPF and RIP
  - OSPF is a link-state protocol,RIP is a distance-vector protocol
    - The distance-vector means each router updates from its neighbors
    - In a link-state protocol,each router actively tests the status of its link to each of its neighbors, sends this information to its other neighbors
  - A link-state protocol will always converge faster than a distance-vector protocol
- Features of OSPF that superior to RIP
  - OSPF can calculate a separate set of routes for each IP TOS
  - Each interface is assigned a dimensionless cost
  - OSPF distributes traffic equally among the routes (load balancing)
  - OSPF supports subnets
  - PPP links between routers do not need an IP address at each end

BGP: Border Gateway Protocol

- A cleartext password can be specified, similar to the RIP-2 scheme
- OSPF uses multicasting instead of broadcasting to reduce the load

- Local traffic and transit traffic
  - Local traffic in an AS either originates or terminates in that AS
  - Anything else is called transit traffic
  - A major goal of BGP is to reduce transit traffic
- Category of an AS
  - A stub AS has only a single connection to one other AS
  - A multihomed AS has connections to more than one other AS
  - A transit AS has connections to more than one other AS and is designed, to carry both local and transit traffic

BGP: Border Gateway Protocol

- BGP allows for policy-based routing
- BGP uses TCP as its transport protocol
- BGP is a distance vector protocol
- BGP detects the failure of either the link or the host by sending keepalive message to its neighbor on a regular basis

- What's CIDR?
  - CIDR is a way to prevent explosion in the size of the Internet routing tables.It is also called supernetting
  - The basic concept in CIDR is to allocate multiple IP addresses summarization into a smaller number of routing table entries
- Three features are needed to allow this summarization
  - Multiple IP addresses must share the same high-order bits

CIDR: Classless Interdomain Routing

- The routing tables and routing algorithms must be extended to base their routing decisions on a 32-bit IP address and a 32-bit mask
- The routing protocols being used must be extended to carry the 32-bit mask in addition to the 32-bit address

- CIDR will slow down the growth of the Internet routing tables

- Summary
  - Two routing protocols: IGPs and EGPs
  - The most popular IGP is RIP
  - OSPF being widespread use
  - The popular EGP is BGP
  - CIDR is to reduce the size of the size of the Internet routing tables
  - Other protocols: IDRP,IS-IS