

Interior Gateway IP Routing Protocols

Redes de Comunicações II

Licenciatura em Engenharia de

Computadores e Informática

Prof. Amaro de Sousa (asou@ua.pt)
DETI-UA, 2024/2025

IP Routing Overview

- Routers must know how to forward packets to any existing IP network.
 - A router immediately knows the IP networks directly connected to each of its interfaces
 - For remote IP networks (i.e., IP networks not directly connected to one of its interfaces), the router must rely on additional information.
- A router can be made aware of remote IP networks through:
 - Static routing: The routing paths from each router to the remote IP networks are administratively configured in the routers.
 - Dynamic routing protocols: Routers communicate between them to learn how to forward packets towards each remote IP network.
 - Policy based routing: Additional routing rules (administratively configured in the routers) to define routing policies other than the basic static/dynamic IP routing.

IP Static Routing

- Static routing does not react to network topology changes:
 - If a link of a static route fails, connectivity is lost.
 - The link failure must be administratively detected and solved (i.e., either solving the failure of configuring a new static route), causing the connectivity lost for a long time.
- Static routing does not scale well when network grows:
 - Administrative burden to maintain routes may become excessive.
- Nevertheless, static routing is useful:
 - when the administrator needs total control over all (or some) routes, for reason like security of QoS (Quality of Service)
 - when a backup to a dynamically learned route is necessary (if the dynamic routing protocol fails)
 - when there is a single routing path towards an IP network
 - if there is no other possible routing path, dynamic routing has no advantage
 - example: a home router that connects to its Internet Service Provider (ISP) only needs to have a default route toward the ISP router

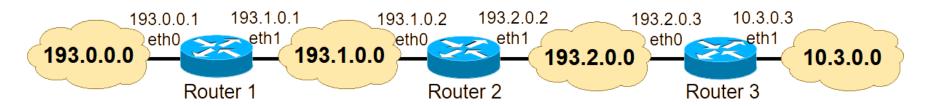
IP Default Routes

- In many cases, a router does not need to recognize each individual remote IP network.
- The router can be configured to send IP packets towards another router for all unknown IP networks (i.e., all IP networks that are not in the IP routing table of the router).
 - This is known as a default route.
- Default routes are either dynamically configured using dynamic routing protocols or statically configured.

IPv4 default route - 0.0.0.0/0

IPv6 default route - ::/0

Static Routing Examples



• Router 2

- Router 2 does not know networks 193.0.0.0/24 and 10.3.0.0/24
- Necessary static routes:
 - 193.0.0.0/24 accessible via 193.1.0.1 (eth1, Router 1)
 - 10.3.0.0/24 accessible via 193.2.0.3 (eth0, Router 3)

Router 1

- Router 1 does not know networks 193.2.0.0/24 and 10.3.0.0/24
- Necessary static routes:
 - 193.2.0.0/24 accessible via 193.1.0.2 (eth0, Router 2)
 - 10.3.0.0/24 accessible via 193.1.0.2 (eth0, Router 2)

OR

• Using default route: 0.0.0.0/0 via through 193.1.0.2 (eth0, Router 2)

IP Dynamic Routing

- Dynamic routing is implemented through protocols running in the routers.
- Dynamic routing protocols, as their name suggests, are used to dynamically exchange routing information between routers.
- Their implementation allows network routing to dynamically adjust to changing network conditions, and to ensure that efficient and redundant routing continues in spite of any changes:
 - When the network topology changes, the new information is dynamically propagated throughout the network, and each router updates its routing table to reflect the changes.
- Routers exchange routing information only with other routers running the same routing protocol.

Longest Match Routing

- Longest Match Routing: when an incoming IP packet in a router has more than one matching entry in its IP routing table, the routing decision uses the entry with the longest prefix
- Example: The current IP routing table of a router is:
 - a) 192.168.1.0/24 via ...
 - b) 192.168.0.0/16 via ...
 - c) 0.0.0.0/0 via ...
- An incoming IP packet to 192.168.1.12 matches all entries:
 - router forwards the packet based on entry: 192.168.1.0/24 via ...
- An incoming IP packet to 192.168.8.14 matches entries b) and c):
 - router forwards the packet based on entry: 192.168.0.0/16 via ...
- An incoming IP packet to 192.1.8.14 matches only entry c):
 - router forwards the packet based on entry: 0.0.0.0/0 via ...

Administrative Distance

- Routers use an **Administrative Distance** value to select the routing paths when they learn the same IP networks from different methods.
 - The method with the lowest Administrative Distance is preferred
- The Administrative Distance has default values but can be configured differently by the administrator.
- Example with default values:
 - Static [1/1] 192.168.1.0/24 via ... \leftarrow Chosen!
 - RIP [**120**/2] 192.168.1.0/24 via ...
 - OSPF [**110**/5] 192.168.1.0/24 via ...
- If Static Routing is required as backup to the dynamic routing protocols, its Administrative Distance must be configured with a value larger than the value(s) of the protocol(s) in use

Equal Cost Multi-Pathing (ECMP)

- **Equal Cost Multi-Pathing**: when the IP routing table of a router has multiple next-hop routers to the same IP destination network, the router applies <u>load balancing</u>:
 - the router splits the packets to that IP destination network equally by the different next-hop routers

Example:

```
R 192.168.30.0/24 [120/1] via 192.168.20.2, 00:00:08, FastEthernet0/1 [120/1] via 192.168.10.1, 00:00:09, FastEthernet0/0
```

Router splits the IP packets towards the remote IP network 192.168.30.0/24 in equal percentage via 192.168.20.1 and via 192.168.10.1

• The squally split of the packets by the different next-hop routers can be packet-based of flow based (next slide)

Load Balancing in ECMP

Packet-based load balancing (the old approach):

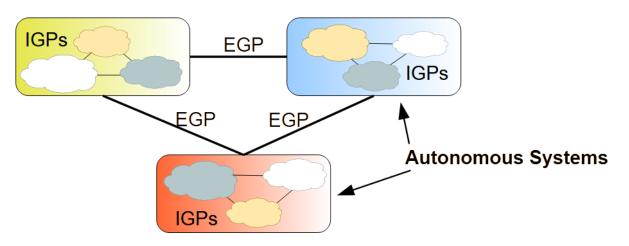
- IP packets are split with a round-robin algorithm
- Introduces undesirable jitter in IP flows when the delay introduced by each routing alternative is significantly different between them
- High jitter and out-of-order reception of IP packets penalises the performance of the TCP congestion-control algorithm

Flow-based load balancing (the current approach):

- IP flows are dynamically assigned among all next-hop routers with the same probability (through hashing)
 - An IP flow is defined by a combination of IP source and destination addresses, protocol (UDP or TCP), source and destination port numbers
- IP packets are forwarded based on the next-hop router assigned to the flow they belong
- It requires hashing, whose complexity is negligible to the processing capability of current routers
- It avoids the disadvantages of packet-based load balancing

Internet Autonomous Systems

- **AS** (**Autonomous System**) set of routers/networks with a common routing policy and under the same administration.
- Routing inside an AS is performed by **IGPs** (**Interior Gateway Protocols**) such as RIP, OSPF, IS-IS and EIGRP.
- Routing between AS is performed by **EGPs** (**Exterior Gateway Protocols**) such as BGP (which has become the standard de facto).
- IGPs and EGPs have different objectives:
 - IGPs: optimize routing performance
 - EGPs: optimize routing performance but highly constrained by economic, political and security policies.



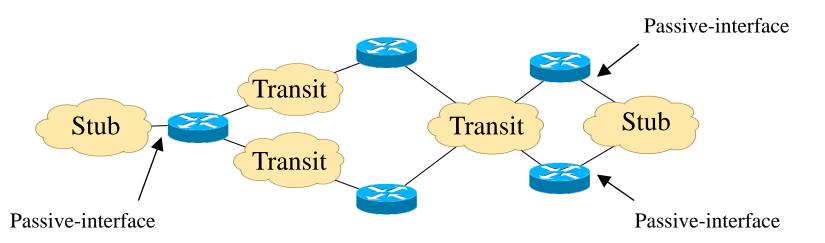
Types of Networks between Routers inside an AS

Transit Networks:

- used by the routers to route IP packets between other IP networks
- routing protocols use them to exchange routing information

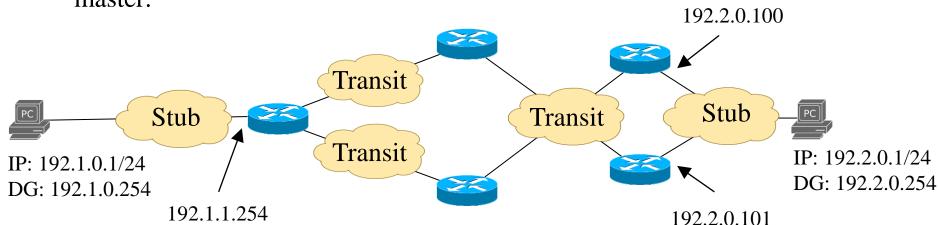
• Stub Networks:

- networks with one attached router
- networks with more than one attached routers, but not used to route IP packets between other IP networks
- in both cases, the router interfaces should be set as passive-interfaces so that routing protocols do not use them to exchange routing information



Virtual Router Redundancy Protocol (VRRP)

- VRRP is a standard protocol defined by the IETF to protected the network against a failure of the Default Gateway.
- A cluster of routers is configured to act as a single router to a LAN with a virtual IP address (to be used as the Default Gateway in the hosts).
- The "best router" (the one with the highest VRRP priority or the highest IP address, if the priority is the same) acts as master and sends periodic VRRP announcement messages to a multicast address:
 - IPv4: 224.0.0.18
 - IPv6: FF02::12
- If the master fails, the "best router" among the remaining ones becomes the master.



Types of IGP Dynamic Routing Protocols

Distance vector

- Each router builds and maintains <u>a vector (i.e., a list) of known IP</u>
 <u>networks</u> and associated distances (costs), based on the information sent
 periodically between neighbour routers.
- Each router determines the shortest paths to all remote IP networks based on the distributed and asynchronous Bellman-Ford algorithm.
- Examples: RIPv2, IGRP, EIGRP.

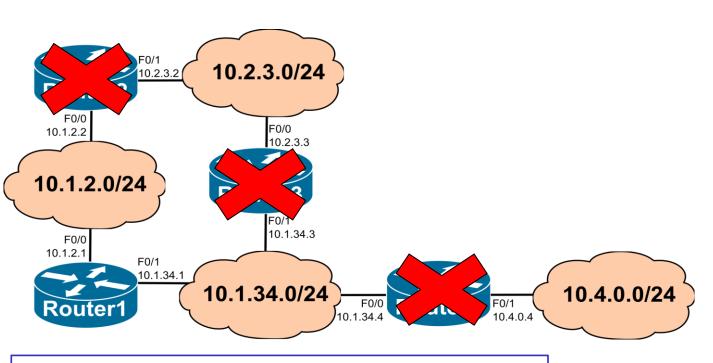
Link state

- Each router learns and maintains <u>a database with the complete network</u> topology view and uses a centralized shortest path algorithm to determine the routing paths from it to all remote IP networks.
- The information required to built the network topology database on each router is obtained by a flooding process.
- Network information is only exchanged on bootstrap and after any topology change.
- Examples: OSPF, IS-IS.

RIP (Routing Information Protocol)

- RIP is a distance vector protocol
 - Each router maintains a list of known IP networks and, for each network,
 an estimation of the cost to reach it this is called a distance vector.
 - Each router periodically sends to its neighbour routers its own distance vector (partially or complete).
 - Each router uses the distance vector received from its neighbours to update its own distance vector.
- The path cost from a router to a remote IP network destination is given by the number of intermediate routers along the path.
 - Maximum cost is 15.
 - A cost of 16 is considered infinite (i.e., an unreachable destination).
- Each router determines the entries of its IP routing table based on its current distance vector:
 - For each remote IP network in the distance vector, the router adds an entry towards each neighbour router providing the lowest cost.

Routing Tables with RIP Router1 starts



Router1

C 10.1.2.0/24 is directly connected, FastEthernet0/0 C 10.1.34.0/24 is directly connected, FastEthernet0/1

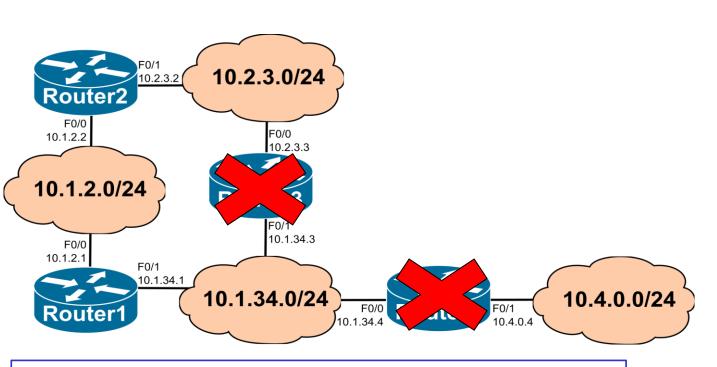
Router2

C 10.1.2.0/24 is directly connected, FastEthernet0/0

R 10.1.34.0/24 [120/1] via 10.1.2.1, 00:00:11, FastEthernet0/0

C 10.2.3.0/24 is directly connected, FastEthernet0/1

Routing Tables with RIP Router2 starts



Router1

C 10.1.2.0/24 is directly connected, FastEthernet0/0

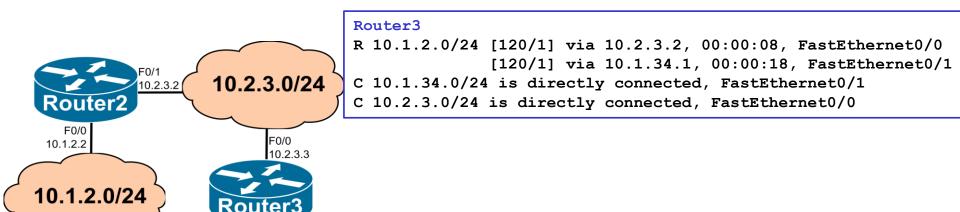
C 10.1.34.0/24 is directly connected, FastEthernet0/1

R 10.2.3.0/24 [120/1] via 10.1.2.2, 00:00:05, FastEthernet0/0

Router2

C 10.1.2.0/24 is directly connected, FastEthernet0/0
R 10.1.34.0/24 [120/1] via 10.2.3.3, 00:00:06, FastEthernet0/1
[120/1] via 10.1.2.1, 00:00:13, FastEthernet0/0
C 10.2.3.0/24 is directly connected, FastEthernet0/1

Routing Tables with RIP Router3 starts



10.4.0.4

10.4.0.0/24

Router1

F0/0 10.1.2.1

Router1

C 10.1.2.0/24 is directly connected, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/1
R 10.2.3.0/24 [120/1] via 10.1.34.3, 00:00:04, FastEthernet0/1
[120/1] via 10.1.2.2, 00:00:12, FastEthernet0/0

F0/0

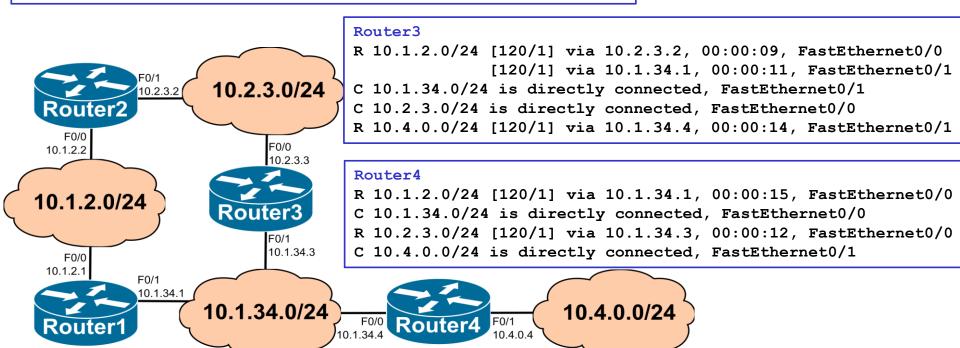
10.1.34.3

10.1.34.0/24

F0/1 10.1.34.1

Router2

Routing Tables with RIP Router4 starts



Router1

C 10.1.2.0/24 is directly connected, FastEthernet0/0 C 10.1.34.0/24 is directly connected, FastEthernet0/1

R 10.2.3.0/24 [120/1] via 10.1.34.3, 00:00:20, FastEthernet0/1 [120/1] via 10.1.2.2, 00:00:04, FastEthernet0/0

R 10.4.0.0/24 [120/1] via 10.1.34.4, 00:00:12, FastEthernet0/1

Router2 (AFTER 3 minutes TIMEOUT)

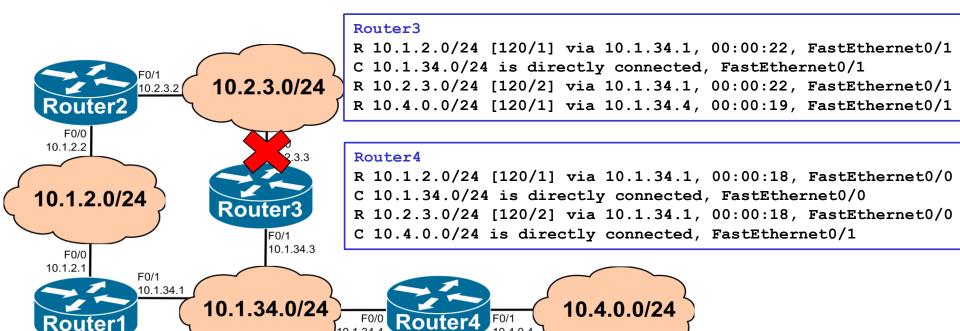
C 10.1.2.0/24 is directly connected, FastEthernet0/0

R 10.1.34.0/24 [120/1] via 10.1.2.1, 00:00:25, FastEthernet0/0

C 10.2.3.0/24 is directly connected, FastEthernet0/1

R 10.4.0.0/24 [120/2] via 10.1.2.1, 00:00:25, FastEthernet0/0

Routing Tables with RIP (after an interface failure)



Router1

C 10.1.2.0/24 is directly connected, FastEthernet0/0

C 10.1.34.0/24 is directly connected, FastEthernet0/1

R 10.2.3.0/24 [120/1] via 10.1.2.2, 00:00:01, FastEthernet0/0

R 10.4.0.0/24 [120/1] via 10.1.34.4, 00:00:24, FastEthernet0/1

RIP Messages

RIP Response message

- It contains the distance vector of the sender router. It is sent:
 - 1. Periodically (~30 seconds by default).
 - 2. When some information changes (*triggered updates*).
 - 3. In response to a RIP Request from a neighbour router.
- In RIP version 2, the RIP Response messages are sent to:
 - the multicast address 224.0.0.8 (in cases 1 and 2)
 - the unicast address of the router that sent the RIP Request (in case 3)
- RIP Request message (Optional)
 - Sent by a router when it starts (bootstrap) or, when the validity of some of the distance vector information has expired (default timeout = 180 seconds)
 - It may request specific information (a specific network), or the complete distance vector of the neighbour router.
- Both RIP messages are sent through UDP (port number = 520)

RIPv2 Response messages without split-horizon

10.4.0.4

- The RIP Response messages are sent by each router with its complete distance vector through all interfaces
- The metric announced to each IP network is the number of hops that the neighbour router will have if the neighbour uses it as the next hop to the network

10.2.3.0/24 10.2.3.2 Router2 F0/0 10.1.2.2 10.2.3.3 10.1.2.0/24 Router3 F0/1 10.1.34.3 F0/0 10.1.2.1 F0/1 10.1.34.1 10.1.34.0/24 10.1.34.4

Sent by Router 3 through F0/0 Internet Protocol Version 4, Src: 10.2.3.3, Dst: 224.0.0.9 ▶ User Datagram Protocol, Src Port: 520, Dst Port: 520 ▼ Routing Information Protocol Command: Response (2) Version: RIPv2 (2) ▼ IP Address: 10.1.2.0, Metric: 2 Address Family: IP (2) Route Tag: 0 IP Address: 10.1.2.0 Netmask: 255.255.255.0 Next Hop: 10.2.3.2 Metric: 2 ▼ IP Address: 10.1.34.0, Metric: 1 Address Family: IP (2) Route Tag: 0 IP Address: 10.1.34.0 Netmask: 255.255.255.0 Next Hop: 0.0.0.0 Metric: 1 ▼ IP Address: 10.2.3.0, Metric: 1 Address Family: IP (2) Route Tag: 0 IP Address: 10.2.3.0 Netmask: 255.255.255.0 Next Hop: 0.0.0.0 Metric: 1 ▼ IP Address: 10.4.0.0, Metric: 2 Address Family: IP (2) Route Tag: 0 IP Address: 10.4.0.0 Netmask: 255.255.255.0 Next Hop: 0.0.0.0 Metric: 2 10.4.0.0/24

RIPv2 Response messages with split-horizon

- The RIP Response messages sent by a router through an interface include only the IP networks whose shortest path is not through the interface.
 - The split-horizon shortens the average number of iterations required for the convergence of the IP routing tables when the network topology changes
- In the example, Router 3 does not announce 10.1.2.0/24 through F0/0 because one of the two shortest paths from Router 3 to 10.1.2.0/24 is through its F0/0 interface.

F0/0 10.1.2.0/24 F0/0 10.1.2.0/24 F0/0 10.1.34.3 F0/0 10.1.34.3 F0/0 10.1.34.4 F0/0 10.1.34.4

Sent by Router 3 through F0/0

Internet Protocol Version 4, Src: 10.2.3.3, Dst: 224.0.0.9

User Datagram Protocol, Src Port: 520, Dst Port: 520

Routing Information Protocol
 Command: Response (2)
 Version: RIPv2 (2)

▼ IP Address: 10.1.34.0, Metric: 1
 Address Family: IP (2)
 Route Tag: 0
 IP Address: 10.1.34.0
 Netmask: 255.255.255.0
 Next Hop: 0.0.0.0
 Metric: 1

▼ IP Address: 10.4.0.0, Metric: 2
 Address Family: IP (2)

10.4.0.0/24

Route Tag: 0

Metric: 2

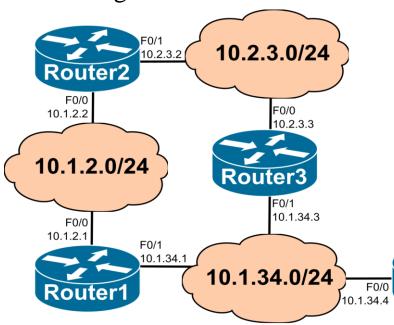
10.4.0.4

IP Address: 10.4.0.0 Netmask: 255.255.255.0 Next Hop: 0.0.0.0

Triggered Updates

- RIP Response messages are sent periodically to continuously check the connectivity between neighbour routers
- *Triggered updates* are RIP Response messages sent by a router to its neighbour routers immediately after a change of its current distance vector (instead of waiting for the time instant of the next periodic RIP Response)
 - If the cost of a known IP network changes
 - If a new IP network is added
 - If a previously known IP network is removed
- In this way, neighbour routers update their distance vector faster and overall convergence becomes also faster.

- The interface F0/1 of Router 4 fails, causing Router 4 to lose connectivity with network 10.4.0.0/24
- Router 4 immediately sends through F0/0 a triggered update announcing the network 10.4.0.0/24 with cost 16 (which means not reachable)
- In this case, the network 10.4.0.0/24 will be quickly removed from the IP routing table of all other routers



Triggered Updates (Illustration)

Sent by Router 4 through F0/0

No.	Time	Source	Destination	Protocol
Г	6 19.989963	10.1.34.4	224.0.0.9	RIPv2
	13 48.085996	10.1.34.4	224.0.0.9	RIPv2
	22 75.094251	10.1.34.4	224.0.0.9	RIPv2
	29 104.786980	10.1.34.4	224.0.0.9	RIPv2
	38 131.405015	10.1.34.4	224.0.0.9	RIPv2
	44 157.130810	10.1.34.4	224.0.0.9	RIPv2
	54 182.953939	10.1.34.4	224.0.0.9	RIPv2
	60 212.648749	10.1.34.4	224.0.0.9	RIPv2
	69 241.319650	10.1.34.4	224.0.0.9	RIPv2
	71 248.829658	10.1.34.4	224.0.0.9	RIPv2
	78 268.235135	10.1.34.4	224.0.0.9	RIPv2
L	86 296.796580	10.1.34.4	224.0.0.9	RIPv2

- Frame 71: 66 bytes on wire (528 bits), 66 bytes captured ▶ Ethernet II, Src: ca:04:56:ad:00:08, Dst: 01:00:5e:00:00:09
- Internet Protocol Version 4, Src: 10.1.34.4, Dst: 224.0.0.9
- ▶ User Datagram Protocol, Src Port: 520, Dst Port: 520
- ▼ Routing Information Protocol

Command: Response (2) Version: RIPv2 (2)

▼ IP Address: 10.4.0.0, Metric: 16

10.4.0.0/24

Address Family: IP (2)

Route Tag: 0

IP Address: 10.4.0.0 Netmask: 255.255.255.0

Next Hop: 0.0.0.0

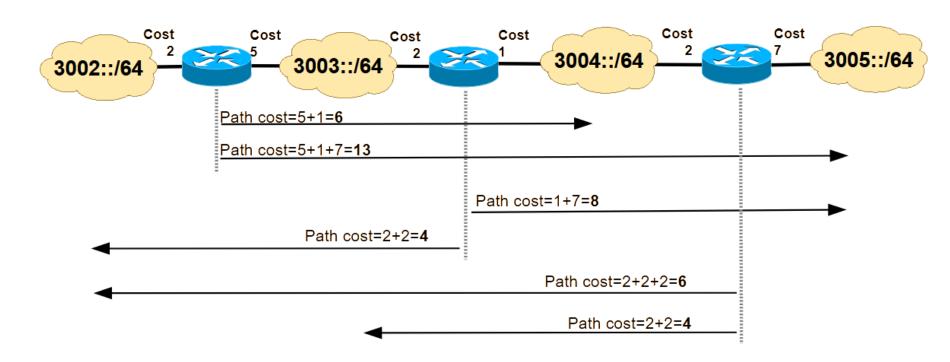
Metric: 16

RIP Next Generation (RIPng) for IPv6 Routing

- RIPng (for IPv6) is similar to RIP version 2 (RIPv2):
 - Distance-vector concept, maximum cost is 15 (16 is infinite), splithorizon, triggered updates, messages transported over UDP
- Differences from RIPv2 to RIPng:
 - RIPng messages are sent through link-local IPv6 addresses (FE80::/64).
 - Uses multicast group address FF02::9 as the destination address for RIP Response messages.
 - A cost is assigned to each output interface.
 - The cost of a routing path is the sum of the costs of the output interfaces towards the destination (illustration in the next slide):
 - allows the operator to control the resulting routing paths.
 - In the split-horizon configuration, routers also include the networks connected to the sending interface:
 - since messages are sent through link-local IPv6 addresses, each router does not know if the neighbour routers on a given interface are also connected to the same global IPv6 networks.

RIPng Path Costs

- Each router interface has an associated RIPng cost (default value = 1).
- The total cost of a routing path from a router to a remote IPv6 network is the sum of the costs of the output interfaces of each router along the path.
- With the infinity metric value at 16, the costs must be carefully configured to prevent any routing path to cost more than 15.



RIPng Response Messages with Split-Horizon

Sent by Router 2 through F0/0

```
Internet Protocol Version 6, Src: fe80::c802:54ff:fef5:6, Dst: ff02::9

User Datagram Protocol, Src Port: 521, Dst Port: 521

▼ RIPng
    Command: Response (2)
    Version: 1
    Reserved: 0000

▼ Route Table Entry: IPv6 Prefix: 2001:1:2::/64 Metric: 1
    IPv6 Prefix: 2001:1:2::
    Route Tag: 0x0000
    Prefix Length: 64
    Metric: 1
```

The IPv6 global network 2001:2:3::/64 is included in the RIPng Response messages exchanged through this network

▼ Route Table Entry: IPv6 Prefix: 2001:2:3::/64 Metric: 1 IPv6 Prefix: 2001:2:3:: Route Tag: 0x0000 Prefix Length: 64

Sent by Router 3 through F0/0

```
▶ Internet Protocol Version 6, Src: fe80::c803:56ff:fe0a:8, Dst: ff02::9
                                                                  ▶ User Datagram Protocol, Src Port: 521, Dst Port: 521
   Metric: 1
                                                                  ▼ RIPnq
                                                                      Command: Response (2)
                                                                      Version: 1
                                                                      Reserved: 0000
                                                                    ▼ Route Table Entry: IPv6 Prefix: 2001:2:3::/64 Metric: 1
                                                                        IPv6 Prefix: 2001:2:3::
                                                                        Route Tag: 0x0000
                          2001:2:3::/64
                2001:2:3::2
                                                                        Prefix Length: 64
                                                                        Metric: 1
                                                                    ▼ Route Table Entry: IPv6 Prefix: 2001:1:34::/64 Metric: 1
                                                                        IPv6 Prefix: 2001:1:34::
     F0/0
                                                                        Route Tag: 0x0000
                                 F0/0
 2001:1:2::2
                                 2001:2:3::3
                                                                        Prefix Length: 64
                                                                        Metric: 1
                                                                    ▼ Route Table Entry: IPv6 Prefix: 2001:4::/64 Metric: 2
2001:1:2::/64
                                                                        IPv6 Prefix: 2001:4::
                          Router3
                                                                        Route Tag: 0x0000
                                                                        Prefix Length: 64
                                                                        Metric: 2
                                 2001:1:34::3
     F0/0
 2001:1:2::1
               F0/1
               2001:1:34::
                      2001:1:34::/64
                                                                         2001:4::/64
                                          2001:1:34::4
```

C 2001:1:2::/64 [0/0] via FastEthernet0/0, directly connected R 2001:1:34::/64 [120/2] via FE80::C801:54FF:FE41:8, FastEthernet0/0 via FE80::C803:56FF:FE0A:8, FastEthernet0/1 C 2001:2:3::/64 [0/0] via FastEthernet0/1, directly connected R 2001:4::/64 [120/3] via FE80::C801:54FF:FE41:8, FastEthernet0/0 via FE80::C803:56FF:FE0A:8, FastEthernet0/ 2001:2:3::/64 2001:1:2::2 2001:2:3::3 2001:1:2::/64 2001:1:34::3 2001:1:2::1 F0/1 2001:4::/64 2001:1:34::/64

IPv6 Routing Tables with RIPng

```
Router3
R 2001:1:2::/64 [120/2]
via FE80::C802:54FF:FEF5:6, FastEthernet0/0
via FE80::C801:54FF:FE41:6, FastEthernet0/1
C 2001:1:34::/64 [0/0]
via FastEthernet0/1, directly connected
C 2001:2:3::/64 [0/0]
via FastEthernet0/0, directly connected
R 2001:4::/64 [120/2]
via FE80::C804:56FF:FEAD:8, FastEthernet0/1
```

```
Router1
```

Router2

```
C 2001:1:2::/64 [0/0]
   via FastEthernet0/0, directly connected
C 2001:1:34::/64 [0/0]
   via FastEthernet0/1, directly connected
R 2001:2:3::/64 [120/2]
   via FE80::C802:54FF:FEF5:8, FastEthernet0/0
   via FE80::C803:56FF:FE0A:6, FastEthernet0/1
R 2001:4::/64 [120/2]
   via FE80::C804:56FF:FEAD:8, FastEthernet0/1
```

```
R 2001:1:2::/64 [120/2]
    via FE80::C801:54FF:FE41:6, FastEthernet0/0
C 2001:1:34::/64 [0/0]
    via FastEthernet0/0, directly connected
R 2001:2:3::/64 [120/2]
    via FE80::C803:56FF:FE0A:6, FastEthernet0/0
C 2001:4::/64 [0/0]
```

via FastEthernet0/1, directly connected

Router4