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

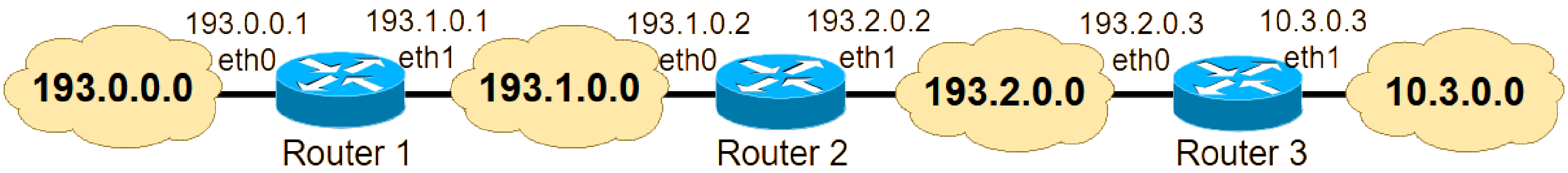
IP Default Routes

- 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 … ← 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

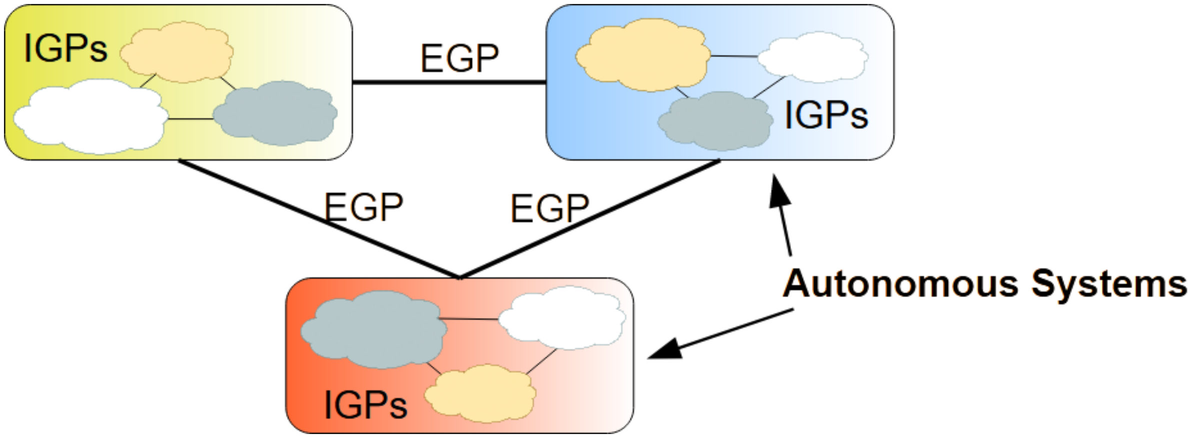
- The squally split of the packets by the different next-hop routers can be packet-based of flow based (next slide)

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

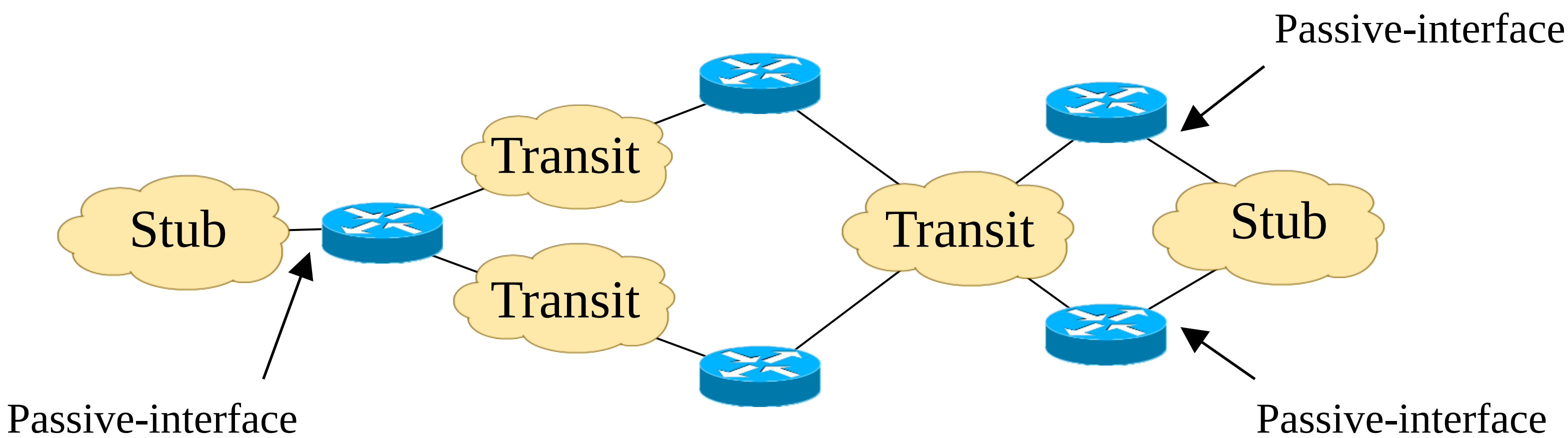
Internet Autonomous Systems

- **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.



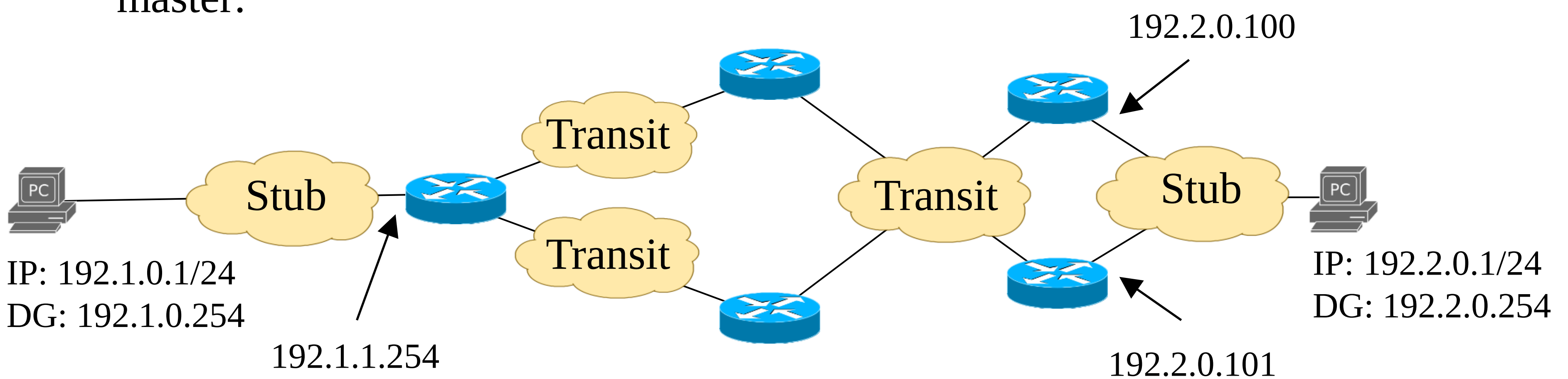
Types of Networks between Routers inside an AS

- 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.



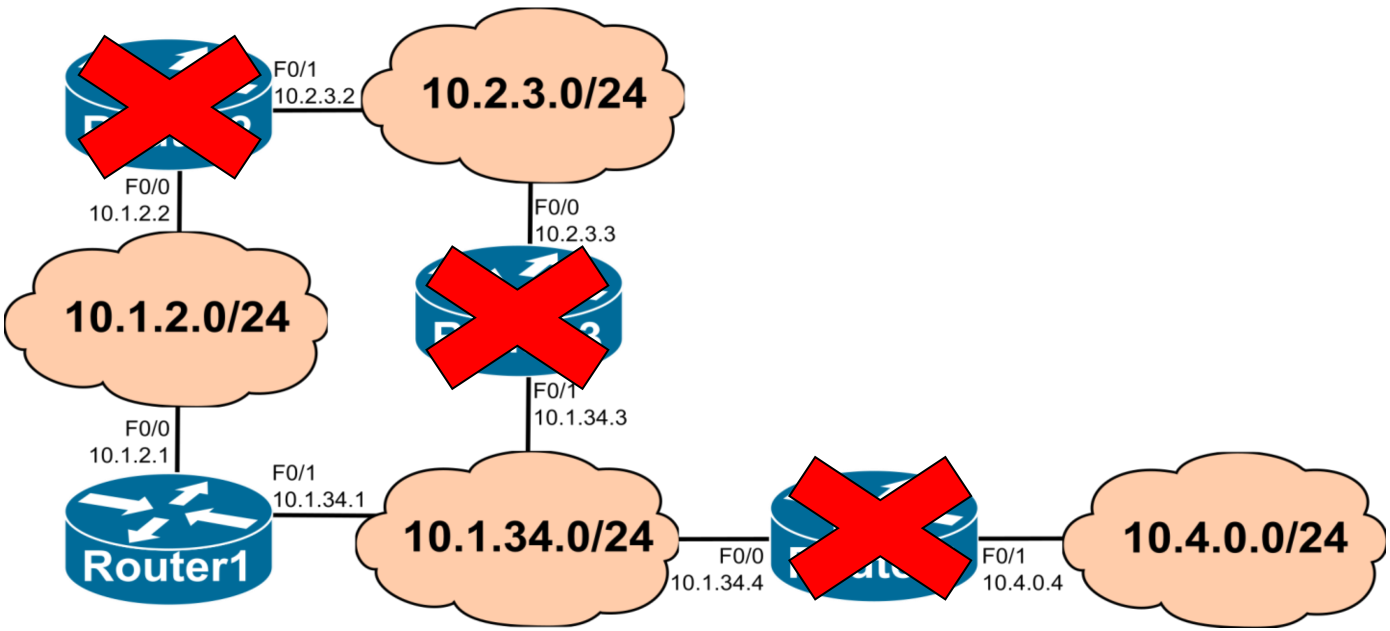
Types of IGP Dynamic Routing Protocols

- Distance vector
  - Each router builds and maintains a vector (i.e., a list) of known IP networks 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 a database with the complete network 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* routing 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.

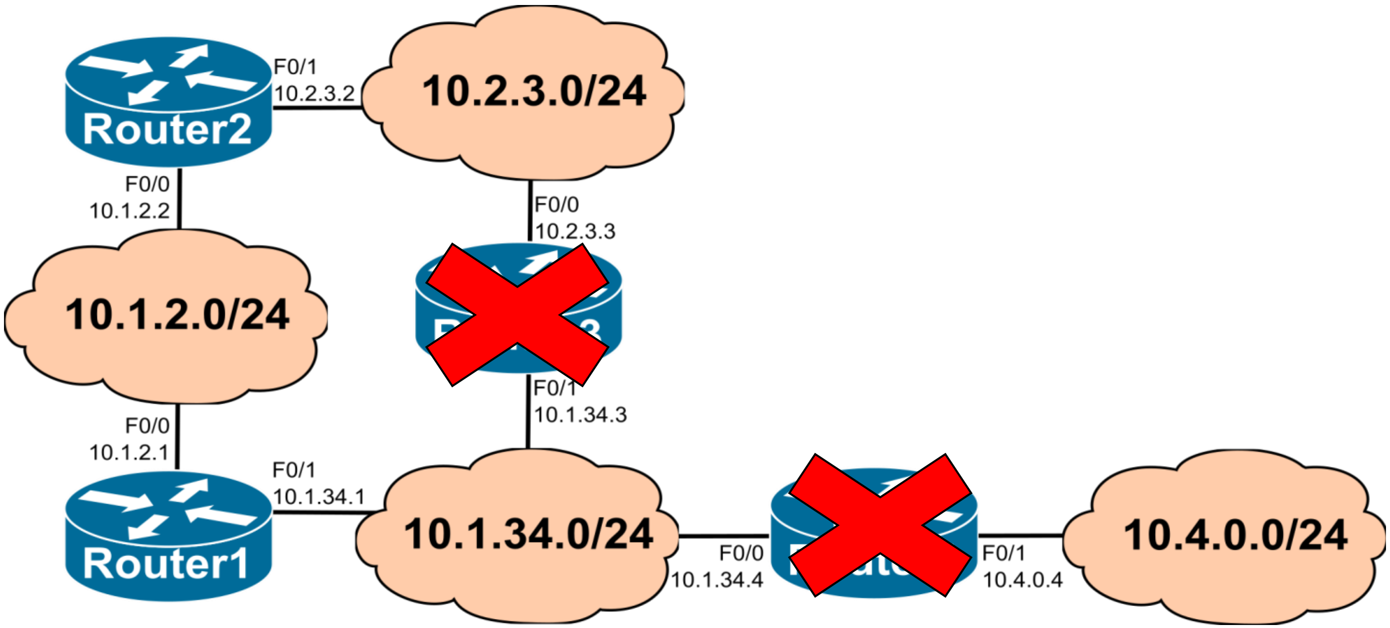
Routing Tables with RIP
Router1 starts




```
Router1
C 10.1.2.0/24 is directly connected, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/1
```

Routing Tables with RIP
Router2 starts

```
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
```

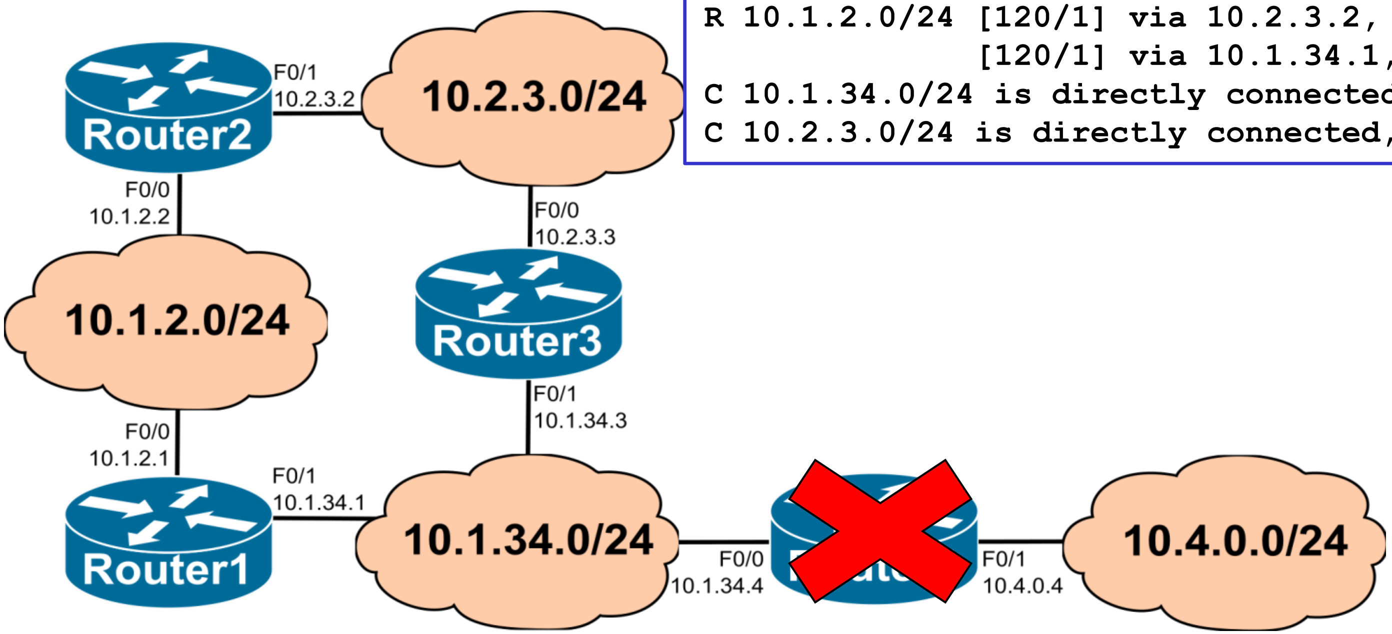



```
Router1
C 10.1.2.0/24 is directly connected, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/1
R 10.2.3.0/24 [120/1] via 10.1.2.2, 00:00:05, FastEthernet0/0
```

Routing Tables with RIP

Router3 starts

```
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
```




```
Router3
R 10.1.2.0/24 [120/1] via 10.2.3.2, 00:00:08, FastEthernet0/0
              [120/1] via 10.1.34.1, 00:00:18, FastEthernet0/1
C 10.1.34.0/24 is directly connected, FastEthernet0/1
C 10.2.3.0/24 is directly connected, FastEthernet0/0
```

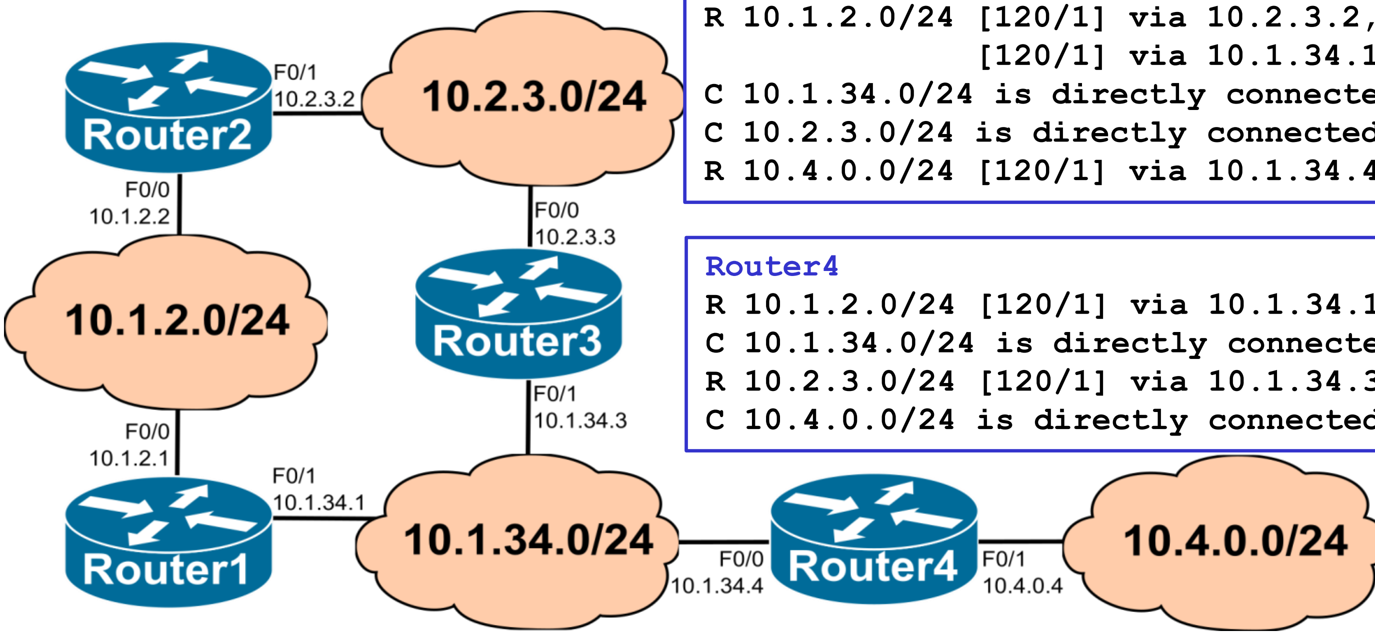
```
Router1
C 10.1.2.0/24 is directly connected, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/1
R 10.2.3.0/24 [120/1] via 10.1.34.3, 00:00:04, FastEthernet0/1
              [120/1] via 10.1.2.2, 00:00:12, FastEthernet0/0
```

Routing Tables with RIP
Router4 starts

```
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, 00:00:14, FastEthernet0/1
               [120/1] via 10.1.2.1, 00:00:07, FastEthernet0/0
C 10.2.3.0/24 is directly connected, FastEthernet0/1
R 10.4.0.0/24 [120/2] via 10.2.3.3, 00:00:14, FastEthernet0/1
              [120/2] via 10.1.2.1, 00:00:07, FastEthernet0/0
```

```
Router3
R 10.1.2.0/24 [120/1] via 10.2.3.2, 00:00:09, FastEthernet0/0
              [120/1] via 10.1.34.1, 00:00:11, FastEthernet0/1
C 10.1.34.0/24 is directly connected, FastEthernet0/1
C 10.2.3.0/24 is directly connected, FastEthernet0/0
R 10.4.0.0/24 [120/1] via 10.1.34.4, 00:00:14, FastEthernet0/1
```

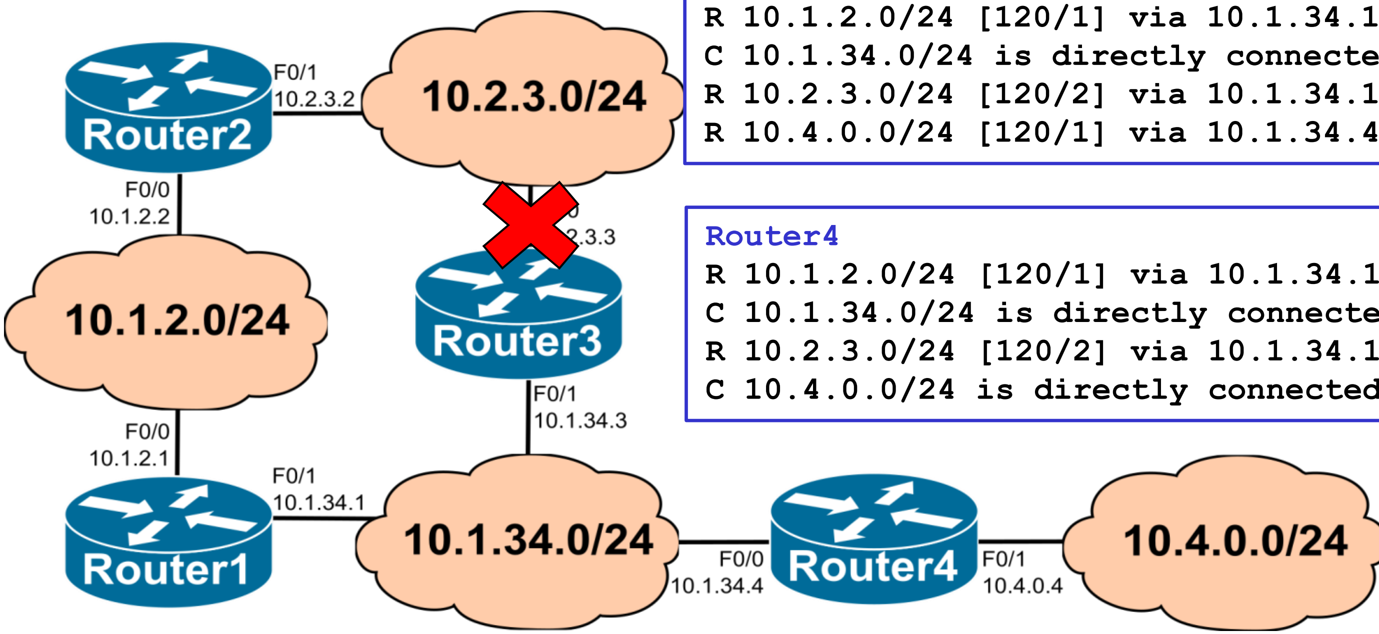
```
Router4
R 10.1.2.0/24 [120/1] via 10.1.34.1, 00:00:15, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/0
R 10.2.3.0/24 [120/1] via 10.1.34.3, 00:00:12, FastEthernet0/0
C 10.4.0.0/24 is directly connected, FastEthernet0/1
```




```
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, 00:00:20, 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
```

Routing Tables with RIP (after an interface failure)

```
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
```




```
Router3
R 10.1.2.0/24 [120/1] via 10.1.34.1, 00:00:22, FastEthernet0/1
C 10.1.34.0/24 is directly connected, FastEthernet0/1
R 10.2.3.0/24 [120/2] via 10.1.34.1, 00:00:22, FastEthernet0/1
R 10.4.0.0/24 [120/1] via 10.1.34.4, 00:00:19, FastEthernet0/1
```

```
Router4
R 10.1.2.0/24 [120/1] via 10.1.34.1, 00:00:18, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/0
R 10.2.3.0/24 [120/2] via 10.1.34.1, 00:00:18, FastEthernet0/0
C 10.4.0.0/24 is directly connected, FastEthernet0/1
```

```
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 10.1.2.2, 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)

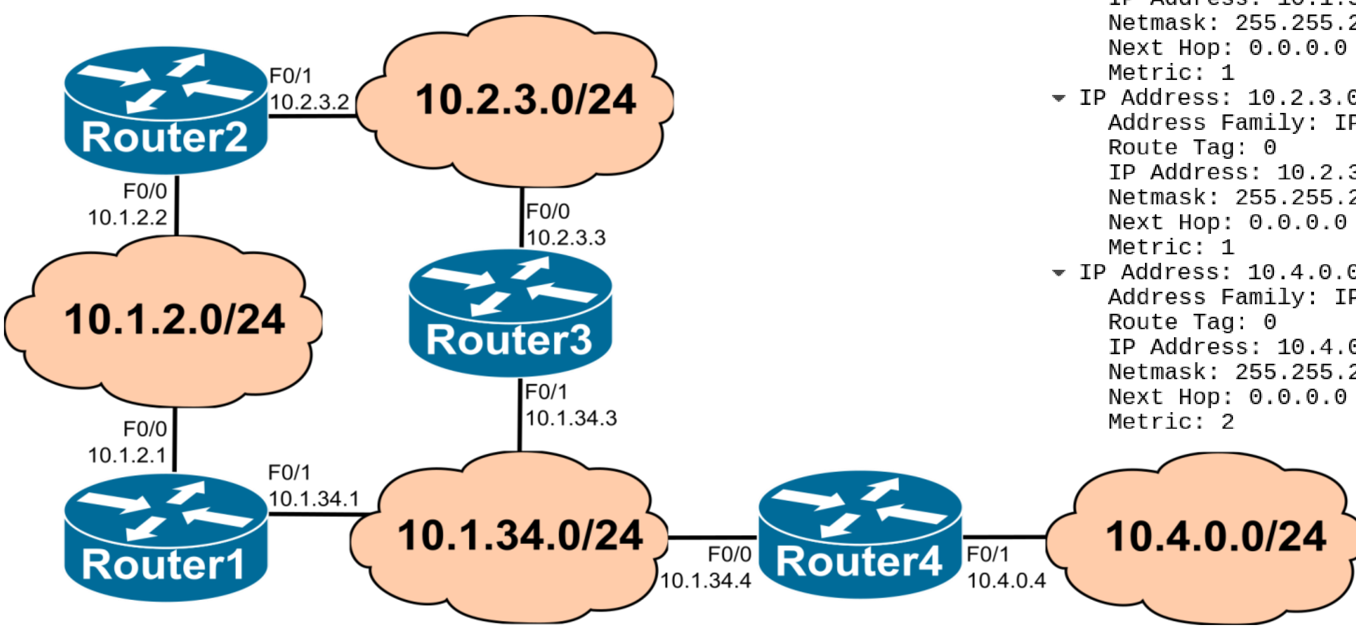
RIPv2 Response messages <u>without split-horizon</u>

- ***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

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
```

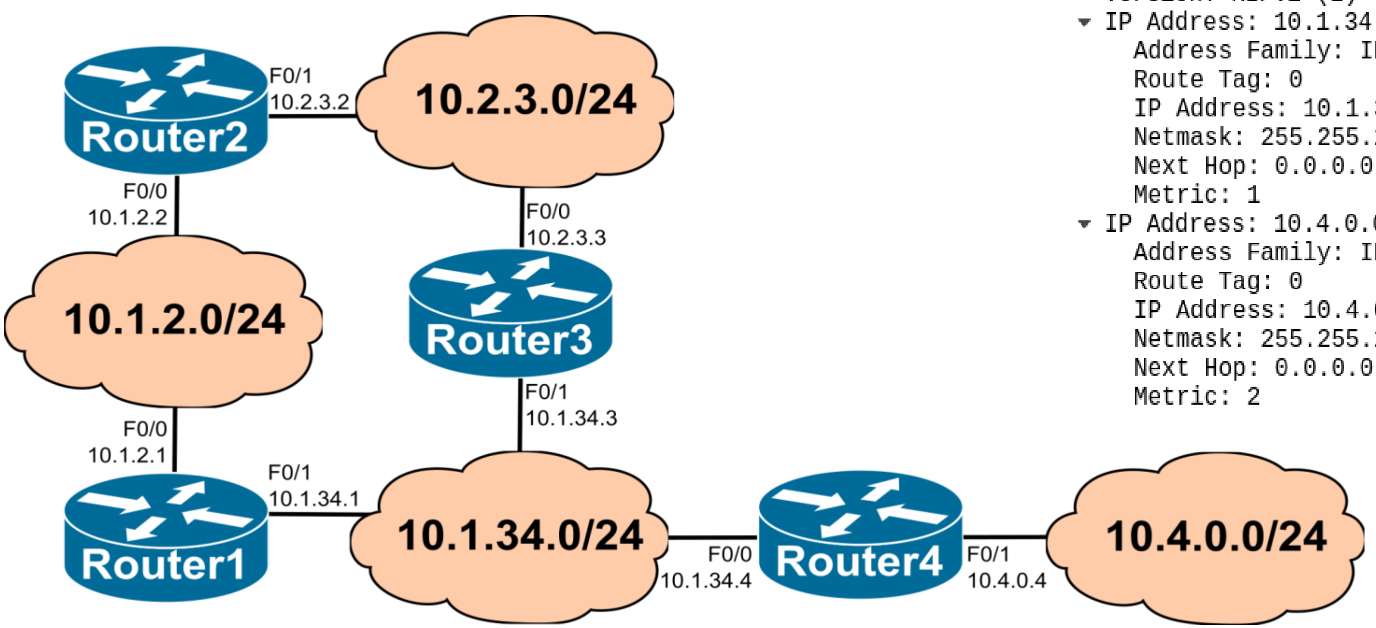


RIPv2 Response messages <u>with split-horizon</u>

- ***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.

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)
      Route Tag: 0
      IP Address: 10.4.0.0
      Netmask: 255.255.255.0
      Next Hop: 0.0.0.0
      Metric: 2
```



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.

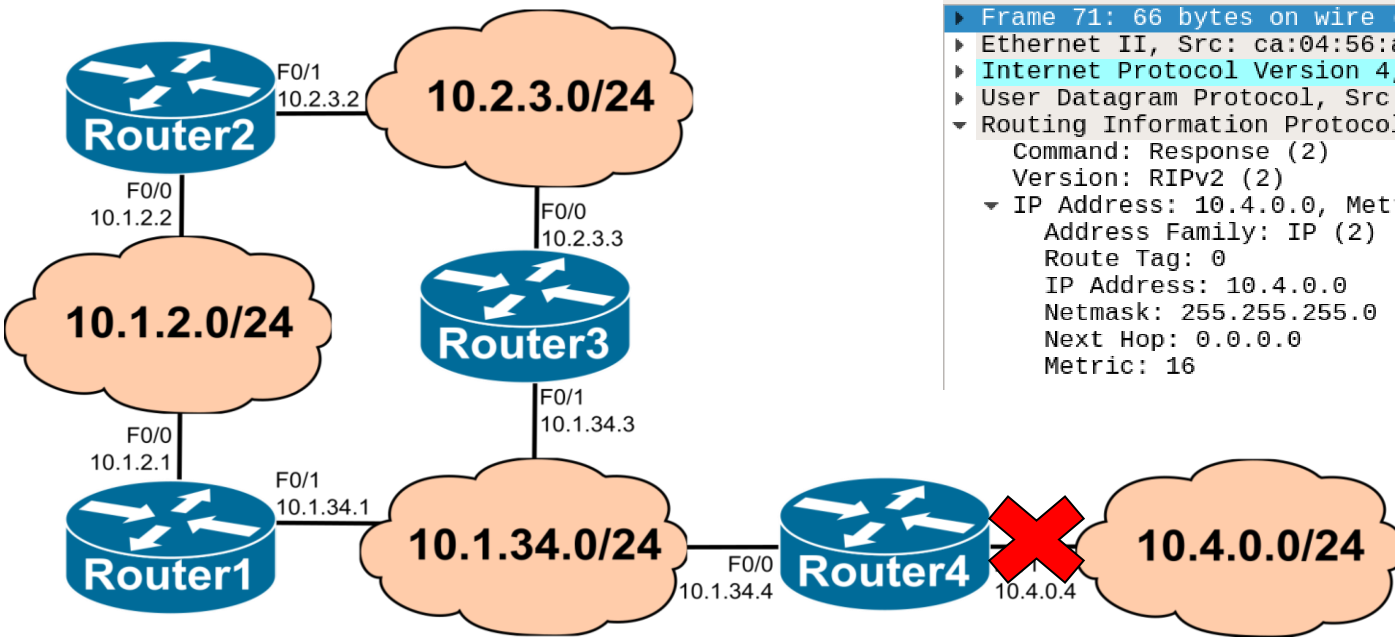
Triggered Updates (Illustration)

- 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

Sent by Router 4 through F0/0

No.	Time	Source	Destination	Protocol
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
86	296.796580	10.1.34.4	224.0.0.9	RIPv2

```
Frame 71: 66 bytes on wire (528 bits), 66 bytes captured (52
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
    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
```

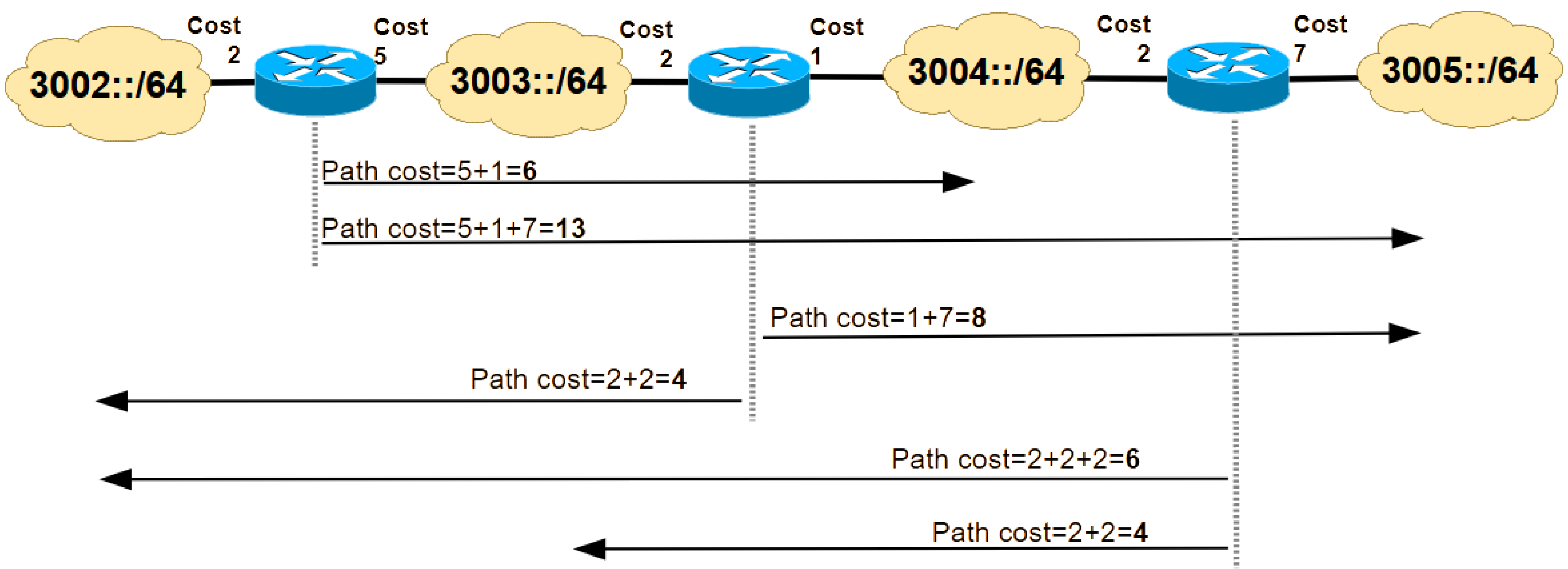


RIP Next Generation (RIPng) for IPv6 Routing

- RIPng (for IPv6) is similar to RIP version 2 (RIPv2):
  - Distance-vector concept, maximum cost is 15 (16 is infinite), split-horizon, 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.
- <u>With the infinity metric value at 16, the costs must be carefully configured to prevent any routing path to cost more than 15</u>.



RIPng Response Messages with Split-Horizon

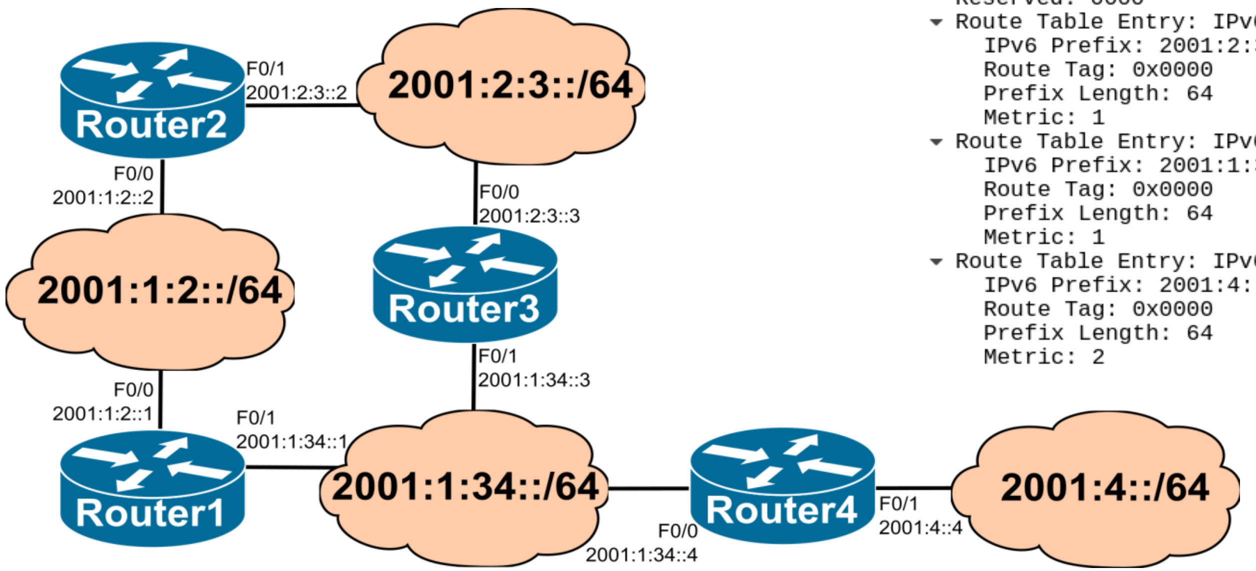
Sent by Router 2 through F0/0

```
▸ 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
  ▾ Route Table Entry: IPv6 Prefix: 2001:2:3::/64 Metric: 1
      IPv6 Prefix: 2001:2:3::
      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

Sent by Router 3 through F0/0

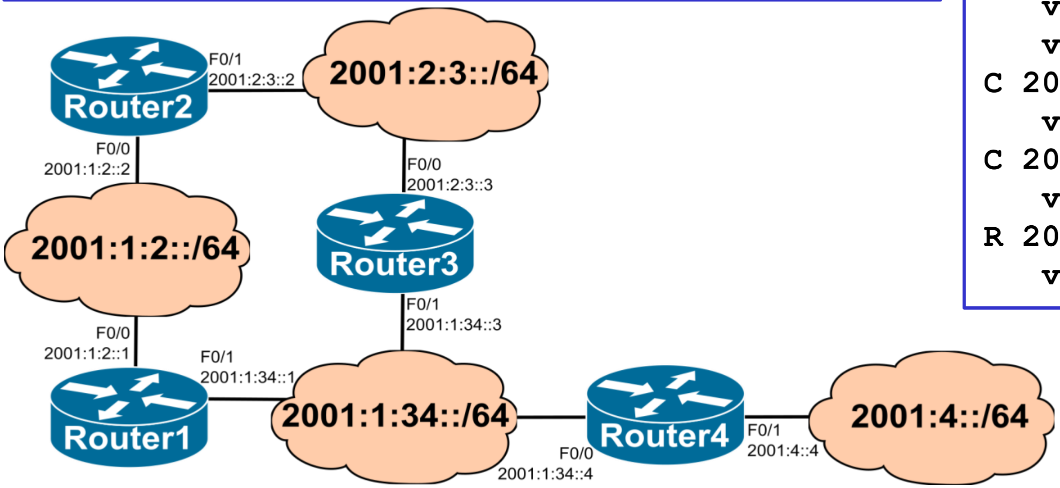
```
▸ Internet Protocol Version 6, Src: fe80::c803:56ff:fe0a:8, Dst: ff02::9
▸ User Datagram Protocol, Src Port: 521, Dst Port: 521
▾ RIPng
    Command: Response (2)
    Version: 1
    Reserved: 0000
  ▾ Route Table Entry: IPv6 Prefix: 2001:2:3::/64 Metric: 1
      IPv6 Prefix: 2001:2:3::
      Route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



IPv6 Routing Tables with RIPng

Assuming the default cost values

Router2
```
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/
```



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
```
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
```

Router4
```
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
```

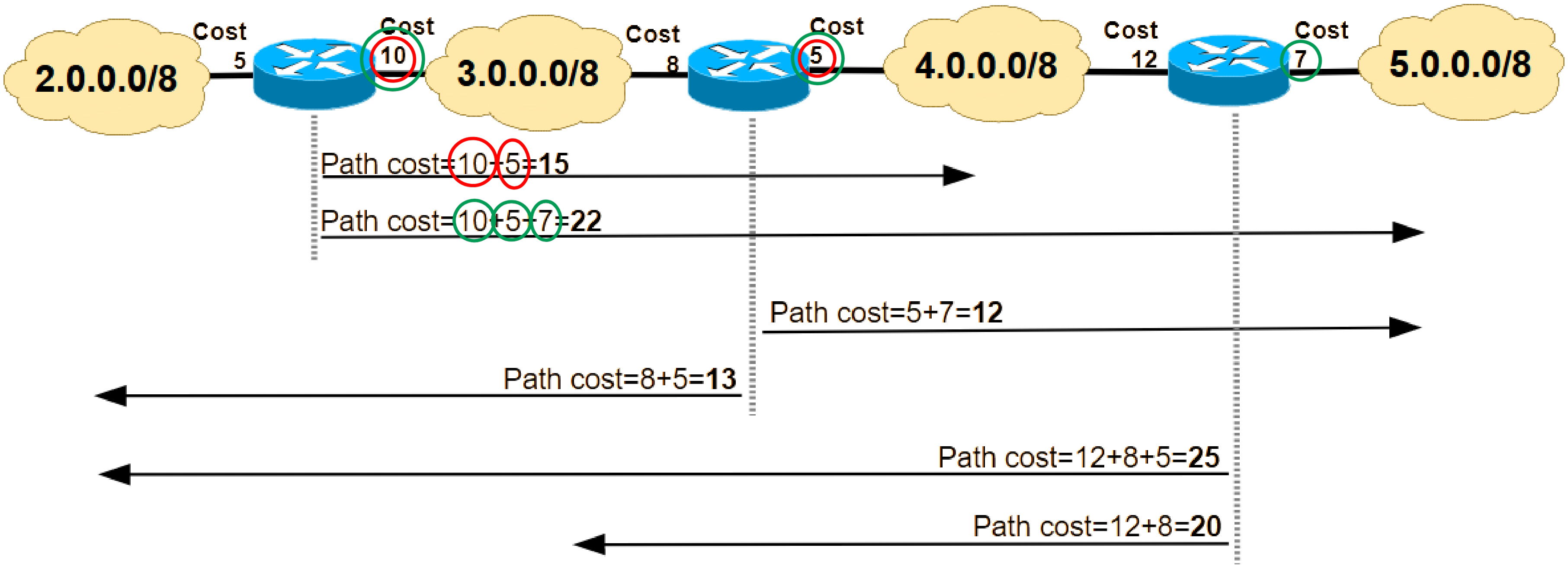
OSPF (Open Shortest Path First) Protocol

- OSPF is a *link-state* routing protocol
  - Each router learns and maintains a database with the complete network topology view, known as the LSDB (Link-State Database)
  - Each router sends its part of the network topology to all routers in the form of LSAs (Link-State Advertisements); these LSAs are sent through a flooding process.
  - LSAs are flooded by the routers on bootstrap (i.e., when they start) or when the topology changes.
- Each router determines the entries of its IP routing table for each remote IP network based on the LSDB:
  - it uses its LSDB to calculate the minimum cost routing paths towards each remote IP network (using Dijkstra's algorithm).
  - and adds to its IP routing table an entry towards each next-hop (i.e., neighbour) router providing the minimum cost.

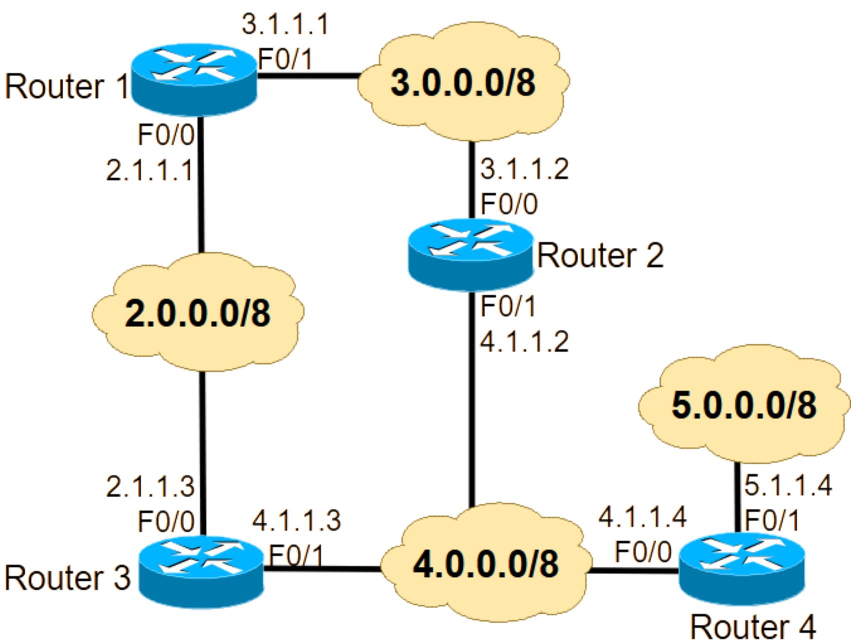
OSPF Path Costs

- Each router interface has an associated OSPF cost.
  - Cost values are administratively configured by the manager
  - Cost values can be defined between 1 and 65535
- The cost of a routing path from a router to a remote IP is the sum of the cost values of the output interfaces of each router along the path.

Depende do sentido do ping!



Routing Tables with OSPF



OSPF costs = 10
(in all interfaces of all routers)

```
Router 1
C  2.0.0.0/8 is directly connected, F0/0
C  3.0.0.0/8 is directly connected, F0/1
O  4.0.0.0/8 [110/20] via 3.1.1.2, 00:01:13, F0/1
             [110/20] via 2.1.1.3, 00:01:31, F0/0
O  5.0.0.0/8 [110/30] via 3.1.1.2, 00:01:10, F0/1
             [110/30] via 2.1.1.3, 00:01:10, F0/0
```

After a failure of interface F0/1 in Router 2:

```
Router 1
C  2.0.0.0/8 is directly connected, F0/0
C  3.0.0.0/8 is directly connected, F0/1
O  4.0.0.0/8 [110/20] via 2.1.1.3, 00:01:18, F0/0
O  5.0.0.0/8 [110/30] via 2.1.1.3, 00:01:00, F0/0

Router 2
O  2.0.0.0/8 [110/20] via 3.1.1.1, 00:01:13, F0/0
C  3.0.0.0/8 is directly connected, F0/0
O  4.0.0.0/8 [110/30] via 3.1.1.1, 00:01:13, F0/0
O  5.0.0.0/8 [110/40] via 3.1.1.1, 00:01:10, F0/0
```

After configuration of interface F0/1 in Router 2 with a OSPF cost = 5:

```
Router 1
C  2.0.0.0/8 is directly connected, F0/0
C  3.0.0.0/8 is directly connected, F0/1
O  4.0.0.0/8 [110/15] via 3.1.1.2, 00:01:13, F0/1
O  5.0.0.0/8 [110/25] via 3.1.1.2, 00:01:10, F0/1
```

OSPF Protocol Principles

- OSPF must guarantee that the LSDB is synchronized in all routers
  - Otherwise, the routing tables on the different OSPF routers can become inconsistent between them
- For that:
  1. Each router must have a unique Router ID (RID)
  2. Each network (either stub or transit) must have a unique Network ID (NID)
  3. Each OSPF router must reliably know what are its OSPF neighbour routers on each of its interfaces (OSPF adjacencies)
  4. Each OSPF router must guarantee that when new LSAs are sent by it through an interface, they are effectively received by all OSPF neighbour routers on that interface.
- Router and Network IDs are in the form of IP addresses
- OSPF runs directly over IP and splits tasks 3. and 4. in different types of messages

OSPF Router ID (RID)

- The Router ID (RID) is:
  - an IP address configured by the administrator (the administrator must guarantee that the IDs of all routers are unique).
  - the highest IPv4 address among all router interfaces at the instant of the OSPF activation (if administrator does not configure router IDs).
- If the IP address of a router interface is being used as the RID, the physical interface fails, and the router (or OSPF process) restarts, the RID changes.
  - This change in router ID makes it more difficult for network administrators to troubleshoot and manage OSPF.
  - Administratively configured RIDs guarantees that they stay the same, regardless of the state of the physical interfaces.
- Each router is responsible to generate its own LSA and send it, through flooding, to all other routers

OSPF Adjacencies

- Each router first establishes neighbour adjacencies through the Hello protocol.
- Each router sends/receives Hello messages to/from its neighbouring routers on each interface running OSPF.
  - The destination address is typically a multicast address (224.0.0.5).
  - It is possible to define unicast OSPF relations.
- The routers exchange Hello messages subject to protocol-specific parameters, such as checking whether the neighbour is in the same area, using the same hello interval, and so on.
  - Routers declare the neighbour up when the exchange is complete.
- On each LAN, routers elect one router as the Designated Router (DR) and another as the Backup Designated Router (BDR).
  - All other routers on the LAN form full adjacencies with these two routers and send LSAs only to them.

DR (Designated Router) and BDR (Backup Designated Router) Election

- On each LAN:
  - The first router to activate the Hello protocol becomes the DR.
  - The second router to activate the Hello protocol becomes the BDR.
- If multiple routers activate the Hello protocol simultaneously:
  - The DR and the BDR are the routers with the highest and second highest priorities, respectively (by default, all routers have the same priority, but different priority values can be administratively configured).
  - In case of tie, the DR and the BDR are the routers with the highest and second highest RIDs (Router IDs), respectively.
- When the DR fails, the BDR assumes the role of DR.
  - The BDR does not perform any DR functions when the DR is active.
  - The choice of the new BDR is again the router high the highest priority or the highest RID in case of equal priorities.
- After the election, the DR and BDR maintain that role, independently of which routers join the OSPF process afterwards.
- <u>The elected DR becomes the responsible to generate the LSA of the LAN IP network and its Network ID (NID) is the IP address of the DR interface.</u>

OSPF Link-State Database (LSDB)

- The OSPF LSDB is organized in two tables.
  - Router Link States – Routers related information table.
    - Each router is identified by its RID (Router ID).
  - Net Link States – Networks related information table.
    - Each IP network is identified by its NID (Network ID).

```
           OSPF Router with ID (20.20.20.1) (Process ID 1)

                     Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
20.20.20.1      20.20.20.1      40          0x8000000A 0x00E7FB 2
30.30.30.2      30.30.30.2      69          0x80000006 0x002906 2
30.30.30.3      30.30.30.3      41          0x80000007 0x00283D 2

                     Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
10.10.10.3      30.30.30.3      41          0x80000001 0x00051C
20.20.20.2      30.30.30.2      70          0x80000001 0x00A164
30.30.30.3      30.30.30.3      154         0x80000001 0x00A91C
```

OSPF LSDB: Router Link State entry

- The Router Link State of a router contains the information about the networks directly connected to the router.
  - In Stub networks, the Link ID is the IP network address, and the Link Data is the network mask

```
LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 20.20.20.1
Advertising Router: 20.20.20.1
LS Seq Number: 8000000A
Checksum: 0xE7FB
Length: 48
Number of Links: 2

  Link connected to: a Transit Network
   (Link ID) Designated Router address: 20.20.20.2
   (Link Data) Router Interface address: 20.20.20.1
    Number of TOS metrics: 0
     TOS 0 Metrics: 1

  Link connected to: a Transit Network
   (Link ID) Designated Router address: 10.10.10.3
   (Link Data) Router Interface address: 10.10.10.1
    Number of TOS metrics: 0
     TOS 0 Metrics: 1
```

RID (Router ID)

Number of connected IP networks

Network Type (Transit of Stub)

OSPF interface cost

NID (Network ID)

IP address of the interface

OSPF LSDB: Net Link State entry

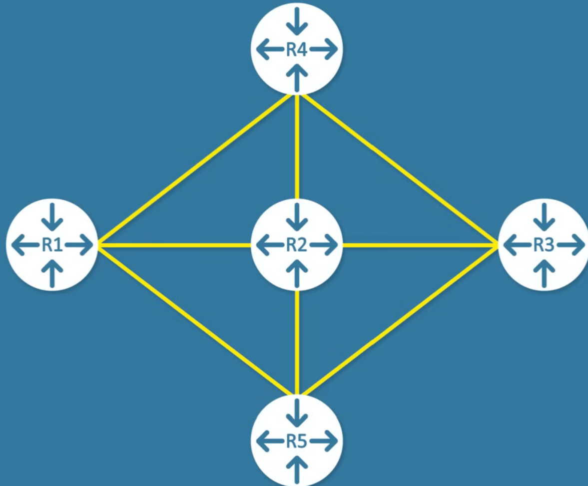
- The net link state of a transit network identifies the routers directly connected to the network.
  - Stub networks do not have net link state entries.

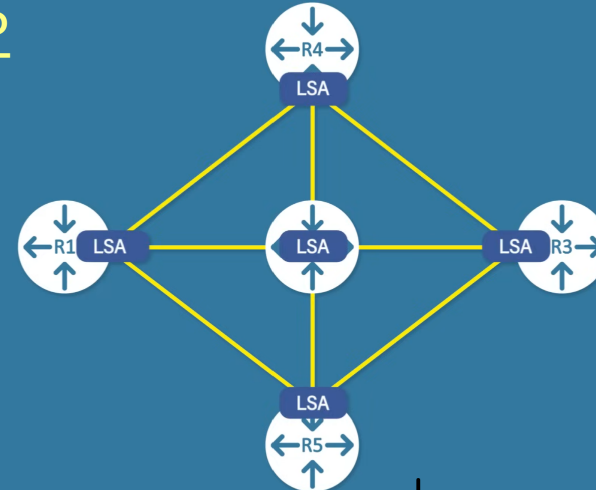
```
LS age: 483
Options: (No TOS-capability, DC)
LS Type: Network Links
Link State ID: 10.10.10.3 (address of Designated Router)
Advertising Router: 30.30.30.3
LS Seq Number: 80000001
Checksum: 0x51C
Length: 32
Network Mask: /24
          Attached Router: 30.30.30.3
          Attached Router: 20.20.20.1
```

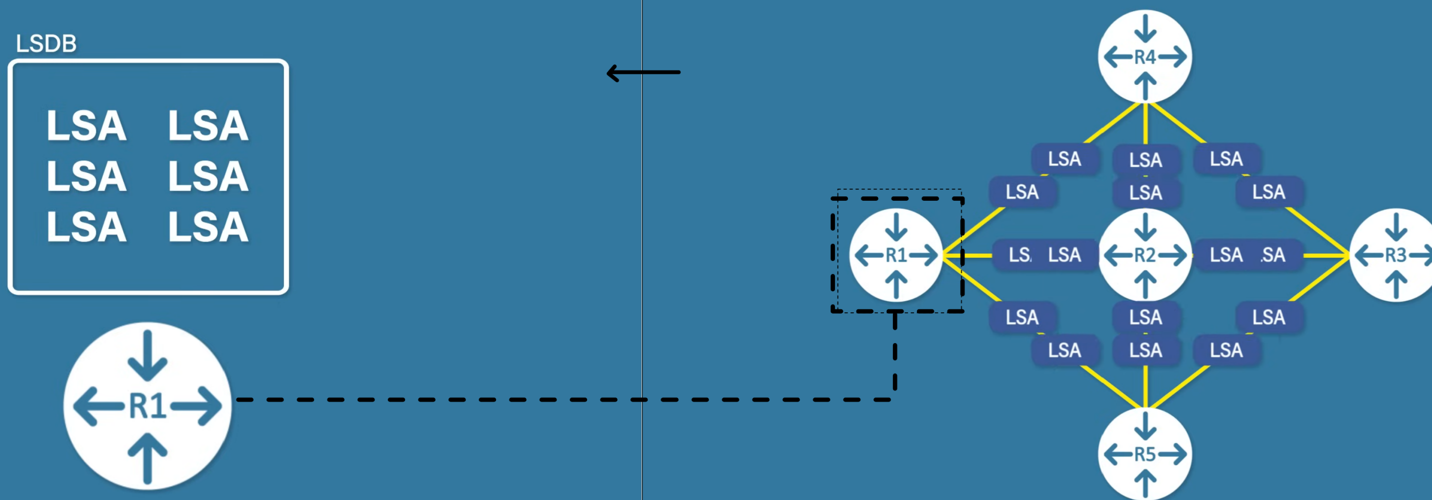
NID (Network ID)

List of RIDs of all routers connected to the IP network

Resumo OSFP

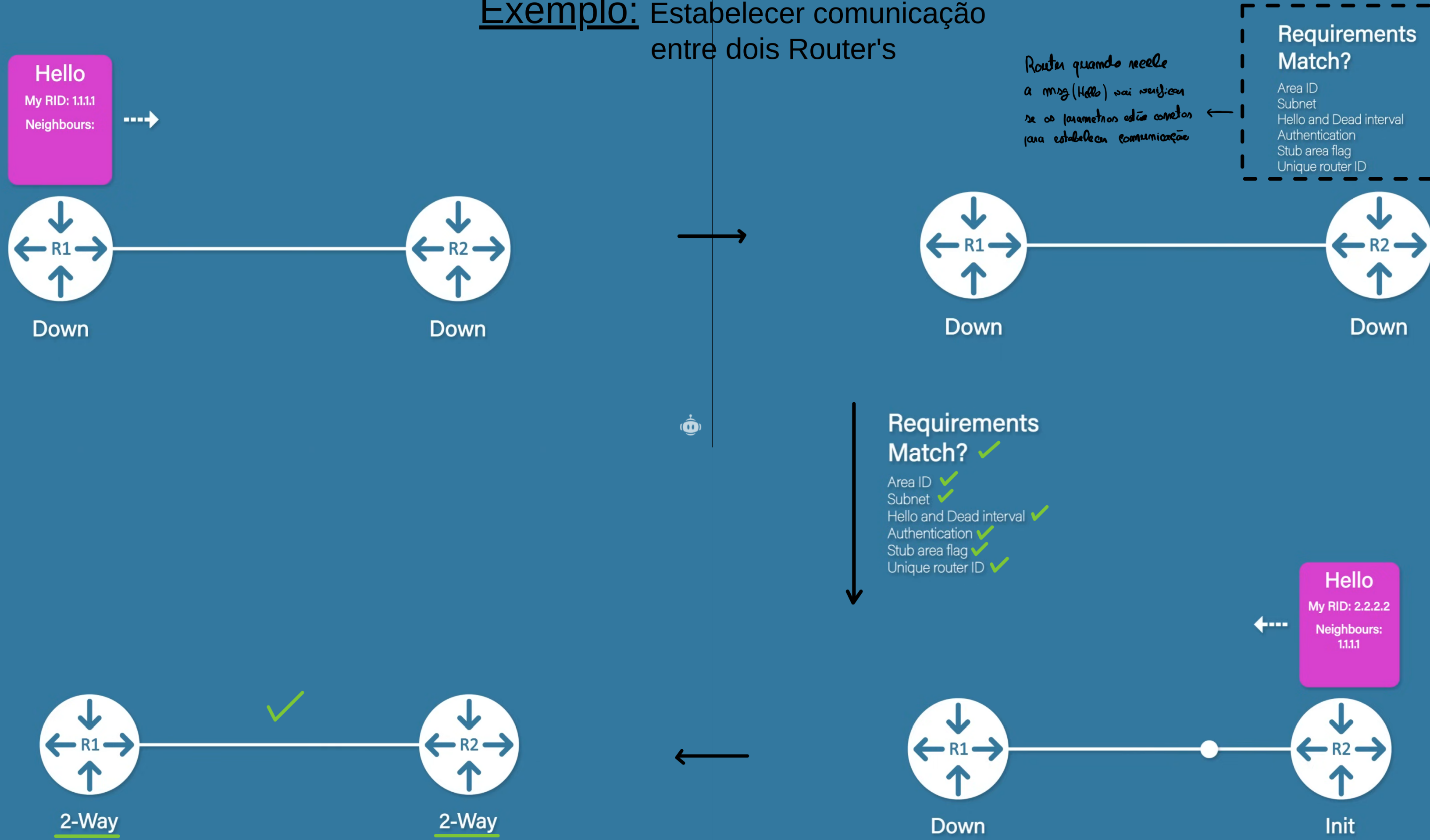

R4
R1
R2
R3
R5


R4
LSA
R1
LSA
LSA
LSA
R3
LSA
R5

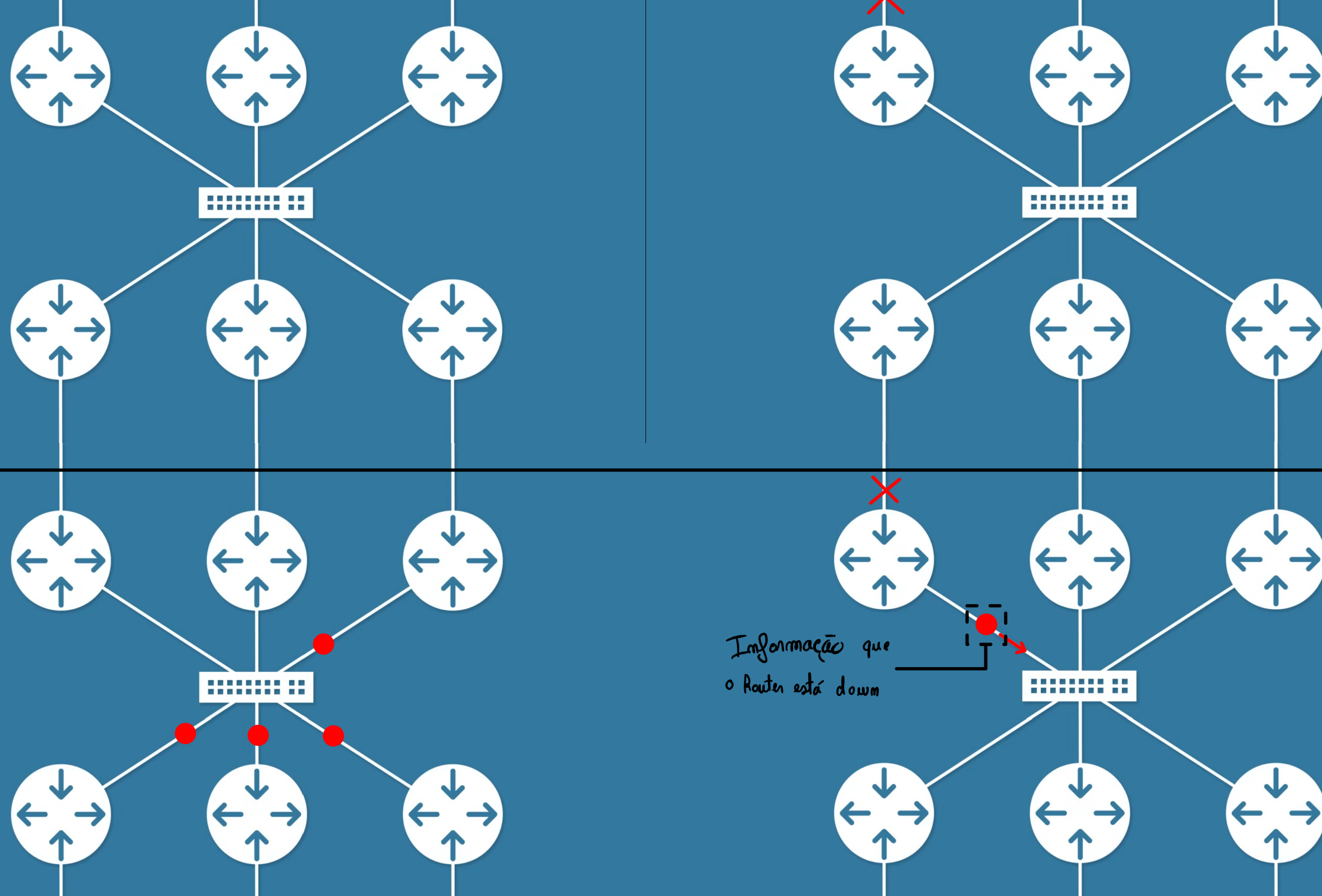

LSDB
LSA LSA
LSA LSA
LSA LSA
R1
R4
R1
R2
R3
R5

<u>Exemplo:</u> Estabelecer comunicação entre dois Router's

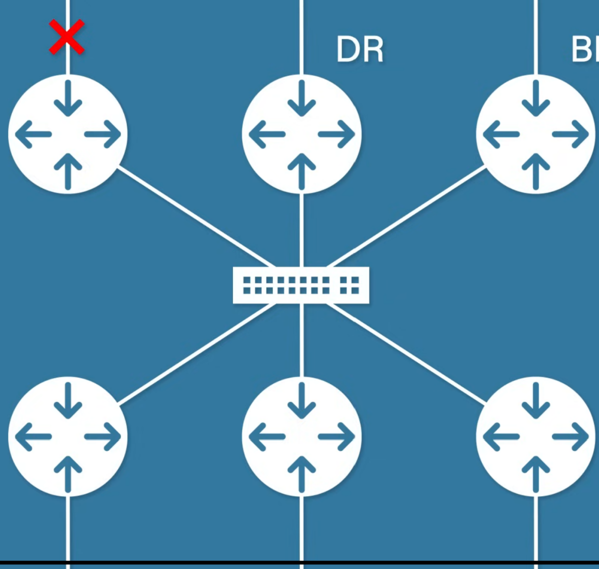
Router quando recebe a msg (Hello) vai verificar se os parametros estão corretos para estabelecer comunicação

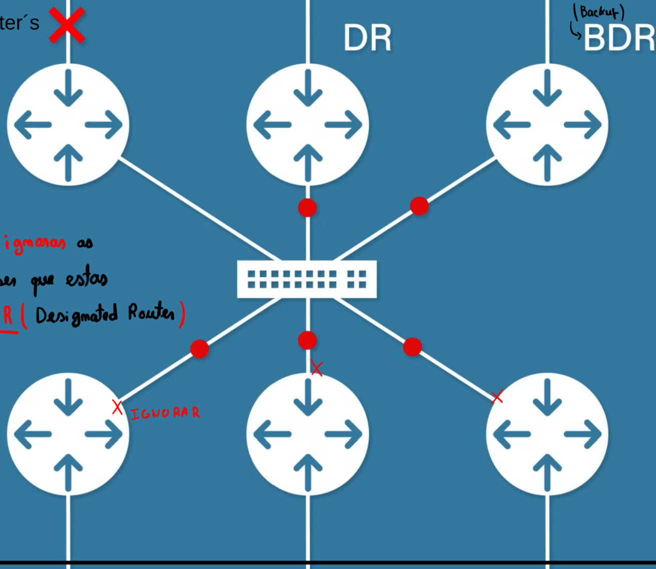


Exemplo: Atualização entre Router´s

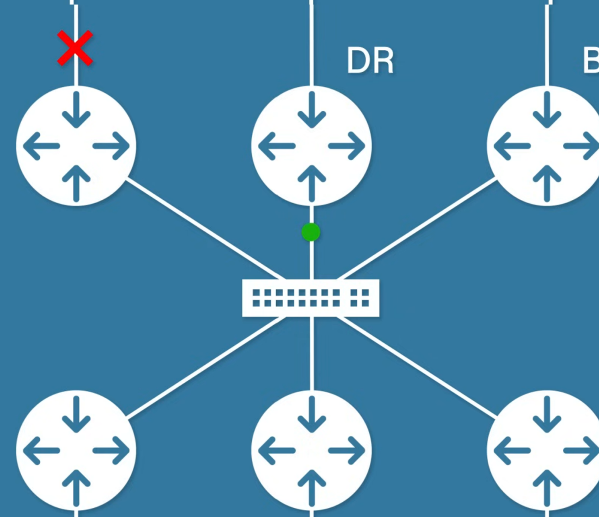


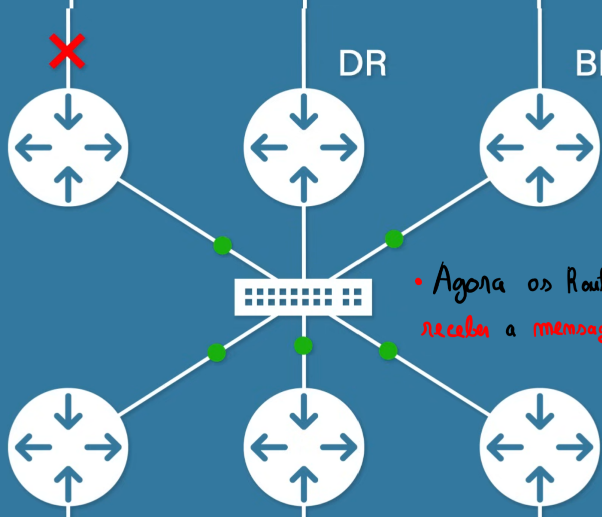
Exemplo: Atualização entre Router´s Com DR e BDR



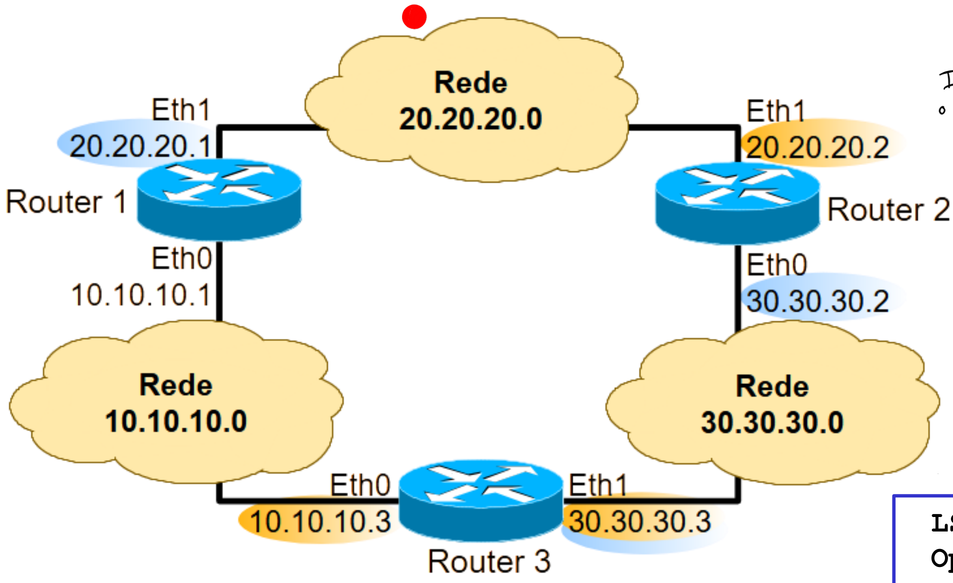


- Os routers vão ignorar as mensagens, a não ser que estas sejam enviadas pelo DR (Designated Router)





- Agora os Routers vão receber a mensagem



Informação que o Router está down

OSPF LSDB Example

Router Link State of router 20.20.20.1:

```
LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 20.20.20.1
Advertising Router: 20.20.20.1
LS Seq Number: 8000000A
Checksum: 0xE7FB
Length: 48
Number of Links: 2

  Link connected to: a Transit Network
   (Link ID) Designated Router address: 20.20.20.2
   (Link Data) Router Interface address: 20.20.20.1
    Number of TOS metrics: 0
     TOS 0 Metrics: 1

  Link connected to: a Transit Network
   (Link ID) Designated Router address: 10.10.10.3
   (Link Data) Router Interface address: 10.10.10.1
    Number of TOS metrics: 0
     TOS 0 Metrics: 1
```

Net Link State of network 10.10.10.3:

```
LS age: 483
Options: (No TOS-capability, DC)
LS Type: Network Links
Link State ID: 10.10.10.3 (address of …
Advertising Router: 30.30.30.3
LS Seq Number: 80000001
Checksum: 0x51C
Length: 32
Network Mask: /24
          Attached Router: 30.30.30.3
          Attached Router: 20.20.20.1
```

OSPF Messages

- Building adjacencies with neighbour routers:
  - Hello protocol
  - For reliability, the Hello messages sent by each router include the Router ID of the neighbour routers
    - In this way, each router knowns it is known by all neighbour routers
- Learning and maintaining the same LSDB in all routers:
  - Database Description (DBD) - Send a summary of the current LSDB.
  - Link-State Request (LSR) - Request specific LSAs from another router.
  - Link-State Update (LSU) - Send specifically LSAs.
  - Link-State Acknowledgement (LSAck) - Acknowledge the reception of LSUs.
  - LSAck messages provide reliable LSA transfer
    - If the sender of a LSU does not receive a LSAck from the receiver, it repeats sending the LSU until the LSAck is received

OSPF Example

OSPF activated on Router 1

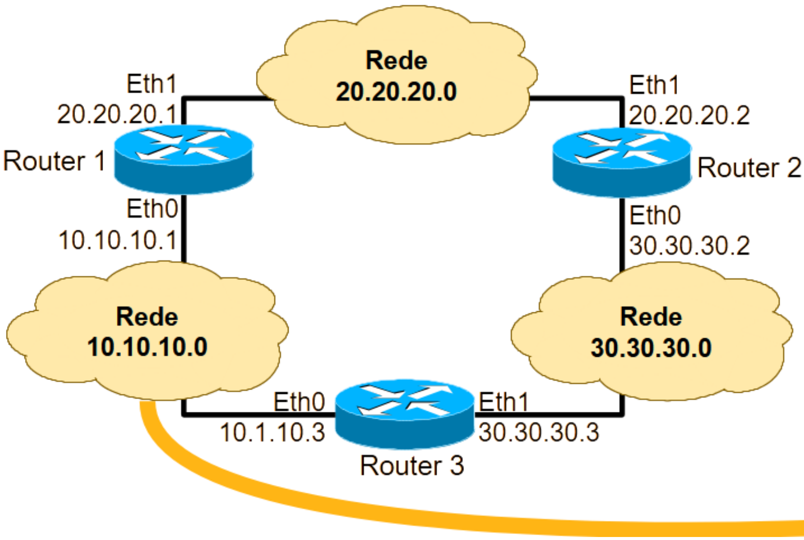
OSPF activated on Router 2

OSPF activated on Router 3

Time	Source	Destination	Protocol	Info
0.000000	10.10.10.1	224.0.0.5	OSPF	Hello Packet
10.002318	10.10.10.1	224.0.0.5	OSPF	Hello Packet
20.003116	10.10.10.1	224.0.0.5	OSPF	Hello Packet

80.000000	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.683033	10.10.10.3	224.0.0.5	OSPF	LS Update
83.715683	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.717864	10.10.10.1	10.10.10.3	OSPF	Hello Packet
83.726166	10.10.10.3	10.10.10.1	OSPF	DB Descr.
83.726258	10.10.10.3	10.10.10.1	OSPF	Hello Packet
83.728433	10.10.10.1	10.10.10.3	OSPF	DB Descr.
83.732590	10.10.10.3	10.10.10.1	OSPF	DB Descr.
83.734733	10.10.10.1	10.10.10.3	OSPF	DB Descr.
83.738942	10.10.10.3	10.10.10.1	OSPF	LS Request
83.741083	10.10.10.1	10.10.10.3	OSPF	LS Update
84.240362	10.10.10.3	224.0.0.5	OSPF	LS Update
86.245792	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
86.380876	10.10.10.1	224.0.0.5	OSPF	Hello Packet
86.741036	10.10.10.1	224.0.0.5	OSPF	LS Acknowledge
93.721376	10.10.10.3	224.0.0.5	OSPF	Hello Packet
96.380005	10.10.10.1	224.0.0.5	OSPF	Hello Packet

213.780338	10.10.10.3	224.0.0.5	OSPF	Hello Packet
216.542473	10.10.10.1	224.0.0.5	OSPF	Hello Packet
216.568852	10.10.10.1	224.0.0.5	OSPF	LS Update
217.048427	10.10.10.1	224.0.0.5	OSPF	LS Update
217.084909	10.10.10.1	224.0.0.5	OSPF	LS Update
219.067748	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
219.650308	10.10.10.1	224.0.0.5	OSPF	LS Update
222.150349	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
223.779492	10.10.10.3	224.0.0.5	OSPF	Hello Packet
224.284149	10.10.10.3	224.0.0.5	OSPF	LS Update
224.789598	10.10.10.1	224.0.0.5	OSPF	LS Update
224.789775	10.10.10.3	224.0.0.5	OSPF	LS Update
226.545718	10.10.10.1	224.0.0.5	OSPF	Hello Packet
226.785254	10.10.10.1	224.0.0.5	OSPF	LS Acknowledge
227.294756	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
233.779863	10.10.10.3	224.0.0.5	OSPF	Hello Packet
236.544658	10.10.10.1	224.0.0.5	OSPF	Hello Packet

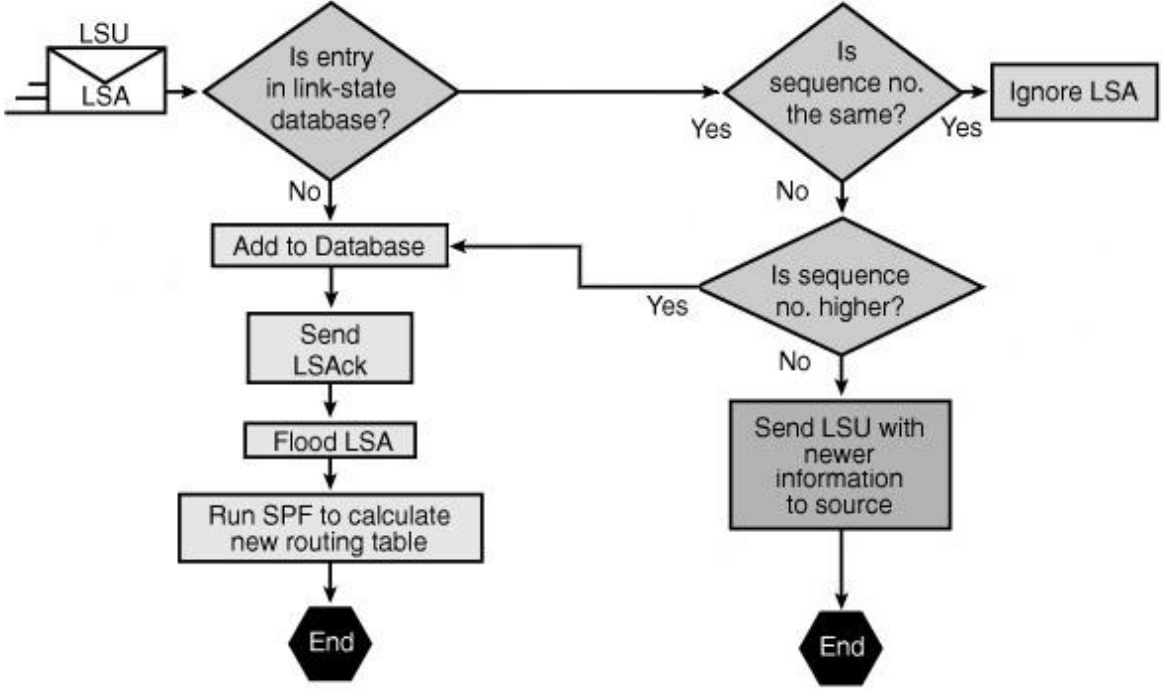


OSPF: flooding process on each LAN

1. When a router notices a change in a link state of its responsibility (either its Router Link State or the Net Link State of the networks that it is the Designated Router) it sends an LSU message, which includes the updated LSA entry with the sequence number incremented, to 224.0.0.6 (the DR multicast address).
   - This message is received by the DR and the BDR.
   - On point-to-point links, the LSU is sent to 224.0.0.5.
   - An LSU message might contain several distinct LSAs.
2. The DR receives the LSU, processes it, acknowledges the receipt of the change and floods the LSU to all routers on the LAN using the OSPF multicast address 224.0.0.5.
3. After receiving the LSU, each router responds to the DR with an LSAck.
   - For reliability, each LSA is acknowledged separately.
4. Each router updates its LSDB using the LSU that includes the changed LSA and updates the IP routing table if necessary.

5. If it is the DR flooding the LSU to the LAN, the process starts on step 2.

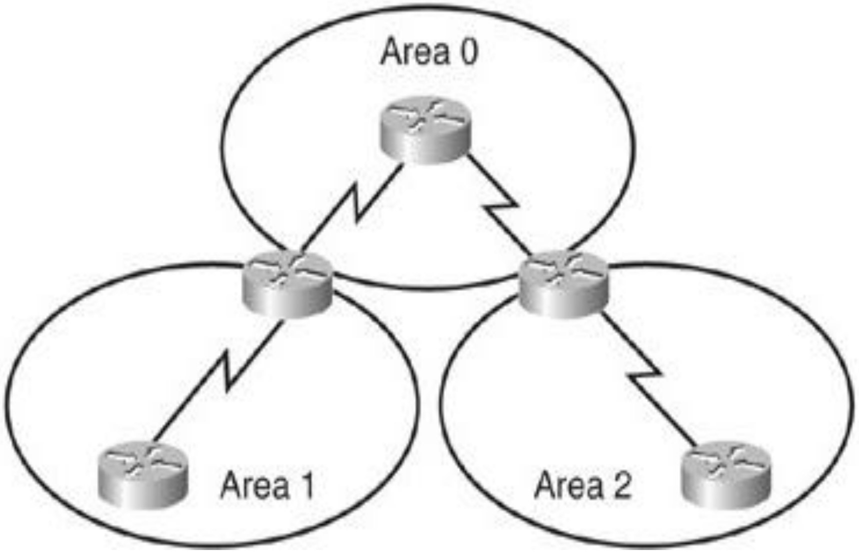
OSPF: processing incoming LSUs



- When a router receives an LSU from one incoming interface:
  - If the LSA entry does not exist, the router adds the entry to its LSDB, sends back a LSAck and floods the information to the other interfaces.
  - If the entry exists and the received LSA has the same sequence number, the router ignores the LSA entry.
  - If the entry exists and the LSA contains newer information (i.e., it has a higher sequence number), the router adds the entry to its LSDB, sends back an LSAck and floods the information to the other interfaces.
  - If the entry already exists and the LSA contains older information (i.e., it has a lower sequence number), it sends an LSU to the sender with its newer information.

OSPF Hierarchical Routing (I)

- In large networks, the resulting web is highly complex, and the number of potential paths to each destination is large. Shortest path calculations can be very complex and can take significant time.
  - **Large LSDB -** Because the LSDB covers the topology of the entire network, each router must maintain an entry for every router and every network.
  - **Frequent shortest path calculations -** In a large network, changes are inevitable, so the routers spend many CPU cycles recalculating the shortest paths and updating the IP routing table.
  - **Large IP routing tables -** OSPF does not perform route summarization by default. If the routes are not summarized, the routing tables can become very large depending on the size of the network.
- Link-state routing protocols usually reduce the complexity of the shortest path calculations by partitioning the network into areas.

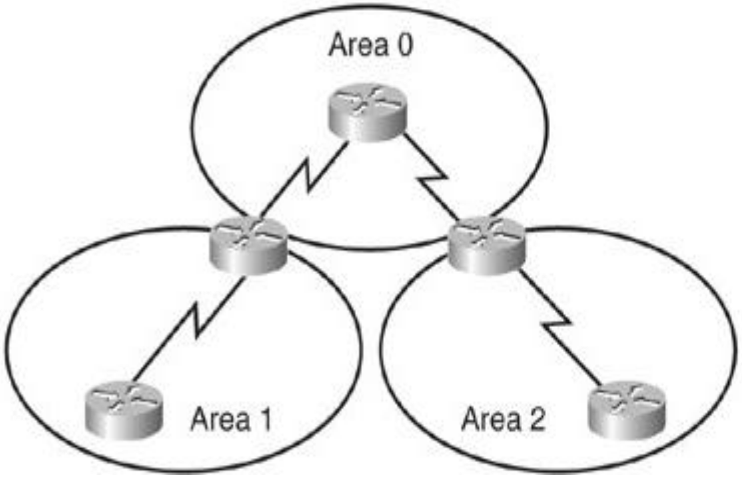


OSPF Hierarchical Routing (II)

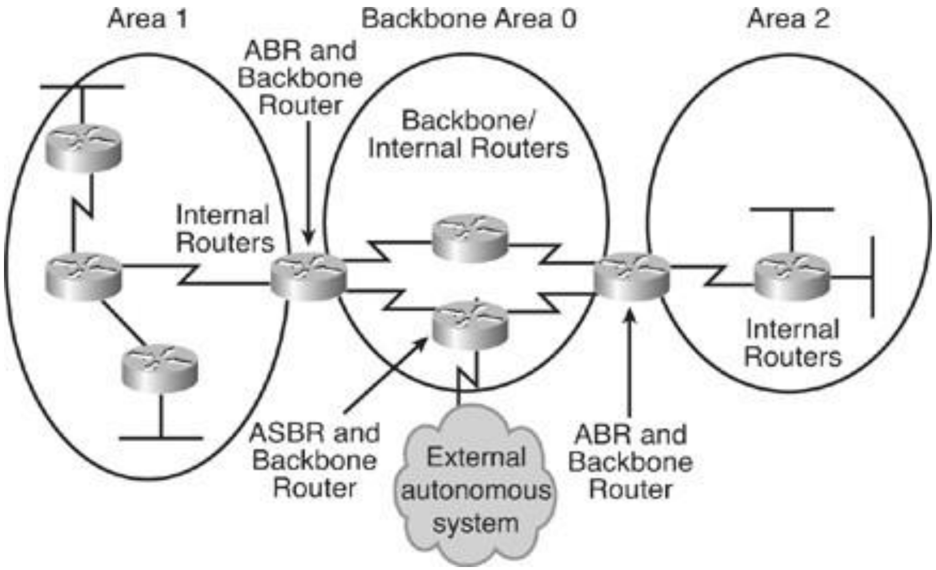
- Using multiple OSPF areas has several important advantages:
  - Reduced frequency of shortest path calculations.
    - Detailed route information only exists within each area
    - It is not necessary to flood all link-state changes to all other areas.
    - Only routers that are affected by the change need to recalculate the shortest paths and the impact of the change is localized within the area.
  - Reduced updates overhead.
    - Rather than send an update about each router and network within an area, a router can advertise a single summarized route or a small number of routes between areas, thereby reducing the overhead associated with updates when they cross areas.
  - Smaller routing tables.
    - Routers can be configured to summarize the routes into one or more summary addresses.
    - Advertising these summaries reduces the number of messages propagated between areas but keeps all networks reachable.

OSPF Two-Layer Area Hierarchy

- Backbone area (Area 0)
  - The OSPF area whose primary function is the fast and efficient forwarding of IP packets between the other areas (or from/to other networks outside the Autonomous System).
  - Generally, end users are not attached to the networks of the backbone area.
  - Hierarchical networking defines Area 0 as the core to which all other areas connect.
- Regular (non backbone) area
  - An OSPF area whose primary function is to connect client users and resources to the Autonomous System.
  - Regular areas are usually set up along functional or geographic groupings.
  - By default, a regular area cannot be used as transit between areas (IP packet between different regular areas must be routed through the backbone area.

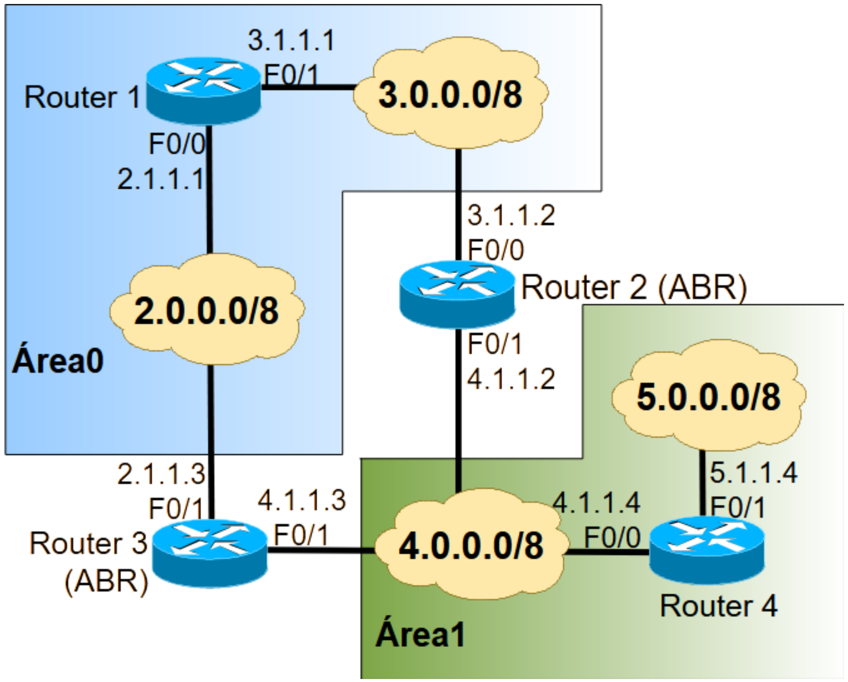


OSPF Router Types



- Internal router
  - Router with all interfaces connected to the same area
- Area Border Router (ABR)
  - Router with interfaces connected to multiple areas
  - ABRs maintain a separate LSDB for each connected area.
  - ABRs connect Area 0 to a non backbone area and are exit points for the area
- Autonomous System Boundary Router (ASBR)
  - Router with at least one interface connected to a different routing domain (another autonomous system or a domain using another routing protocol).
  - ASBRs can redistribute external routes into the OSPF domain and vice versa.
- A router can be simultaneously an ABR and an ASBR
  - For example, if a router is connected to Area 0 and Area 1, and connected to a different routing domain.

OSPF Hierarchical Routing Example



OSPF costs = 10
(in all interfaces of all routers)

Advertising Router: 3.1.1.1	Advertising Router: 4.1.1.2
Number of Links: 2	Number of Links: 1
Router Interface address: 3.1.1.1	Router Interface address: 3.1.1.2
TOS 0 Metrics: 10	TOS 0 Metrics: 10
Router Interface address: 2.1.1.1	Advertising Router: 4.1.1.3
TOS 0 Metrics: 10	Number of Links: 1
	Router Interface address: 2.1.1.3
	TOS 0 Metrics: 10

Router Link States from Router 1

Link State ID: 4.0.0.0	Link State ID: 5.0.0.0
Advertising Router: 4.1.1.2	Advertising Router: 4.1.1.2
Network Mask: /8	Network Mask: /8
TOS: 0 Metric: 10	TOS: 0 Metric: 20
Link State ID: 4.0.0.0	Link State ID: 5.0.0.0
Advertising Router: 4.1.1.3	Advertising Router: 4.1.1.3
Network Mask: /8	Network Mask: /8
TOS: 0 Metric: 10	TOS: 0 Metric: 20

Summary Net Link States from Router 1

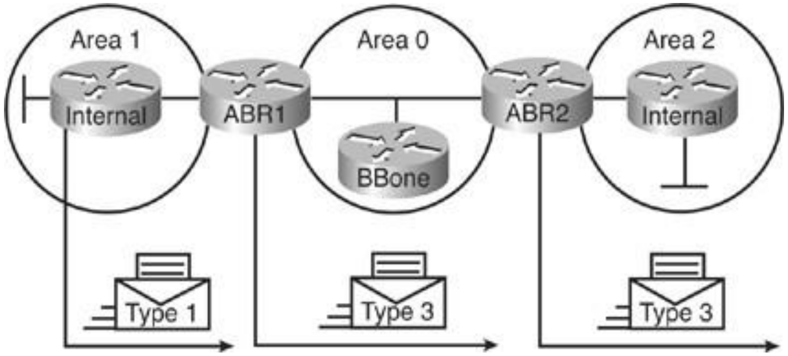
Link State ID: 2.1.1.3	Link State ID: 3.1.1.2
Network Mask: /8	Network Mask: /8
Attached Router: 3.1.1.1	Attached Router: 3.1.1.1
Attached Router: 4.1.1.3	Attached Router: 4.1.1.2

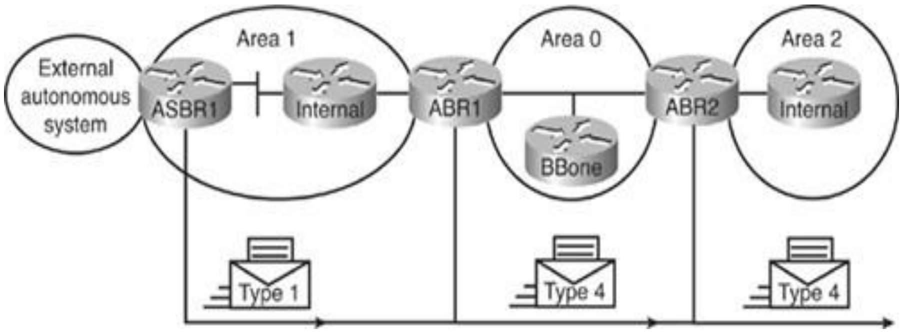
Net Link States from Router 1

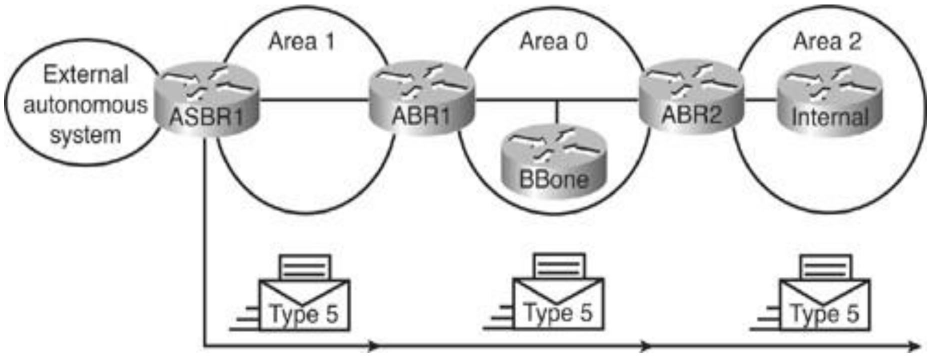
Another LSA type: Summary Net Link State

OSPF relevant LSA types

- Type 1 (Router LSA) - Routers generate their LSAs for each area to which they belong.
- Type 2 (Network LSA) - DRs generate network LSAs for their transit networks.
- Types 3 and 4 (Summary LSA) - ABRs generate summary net link states:
  - Type 3 describes routes to the networks of the other areas.
  - Type 4 describes routes to outside networks announced by ASBRs.
- Type 5 (AS external LSA) - ASBRs generate autonomous system external link advertisements. External link advertisements describe routes to destinations external to the autonomous system.



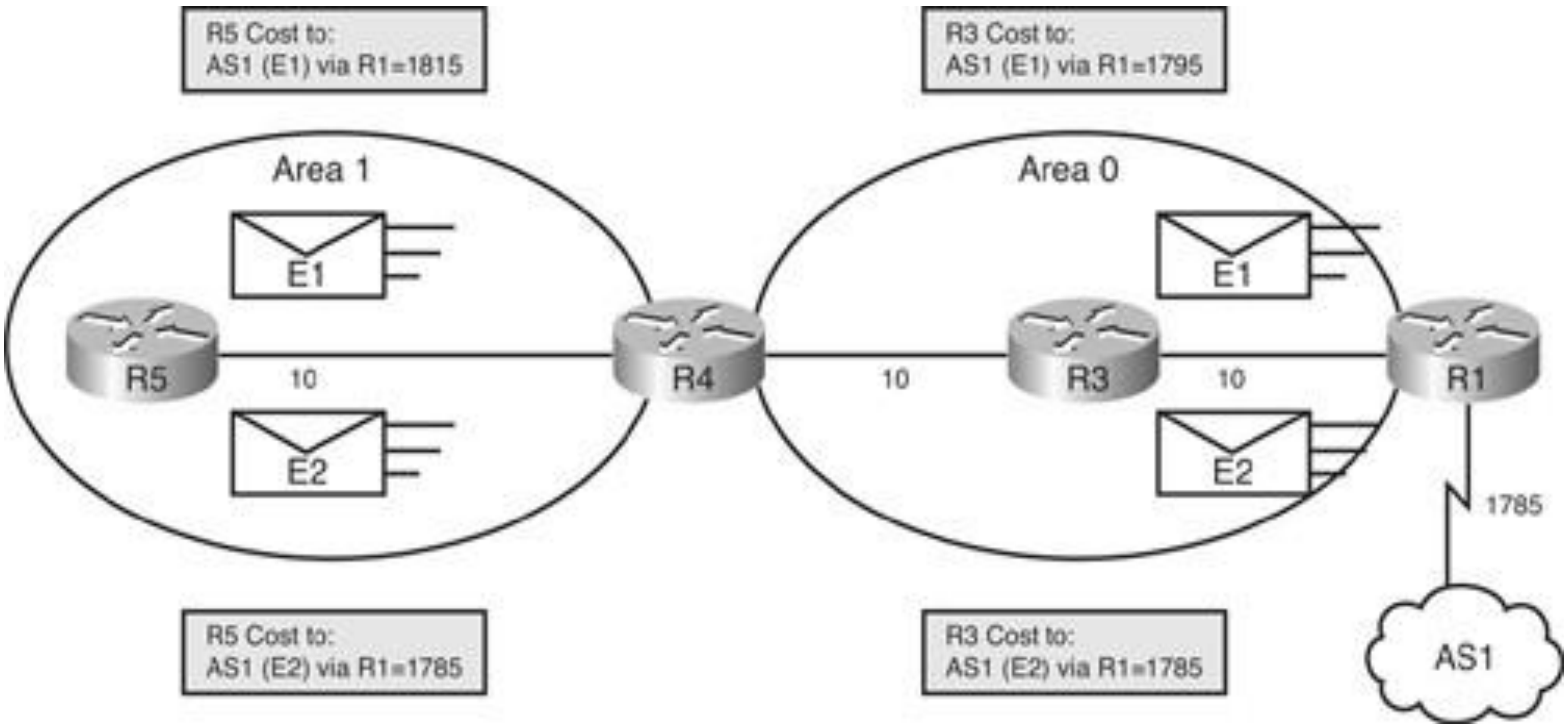




Types of OSPF Routes

- OSPF intra-area (router LSAs and net LSAs)
  - Networks in one of the areas of the router (advertised by Router and Net LSAs).
- Inter-area (summary LSAs)
  - Networks outside the router areas of the router but within the OSPF autonomous system (advertised by summary LSAs).
- Type 2 external routes (E2)
  - Networks outside the OSPF domain (advertised by external LSAs).
  - The cost of OSPF E2 routes is always the external cost only.
  - Use this type if only one ASBR is advertising an external route.
- Type 1 external routes (E1)
  - Networks outside the OSPF domain (advertised by external LSAs).
  - Routers calculate the path cost by adding the external cost to the internal path cost.
  - Use this type when multiple ASBRs are advertising the same outside networks (to avoid suboptimal routing).
  - Always preferred over Type 2 external routes (E2).

OSPF route path cost illustration with external routes of type E1 and E2



Routing - OSPFv3 → Suporta apenas IPv6

- Based on OSPFv2, with enhancements:
  - Uses IPv6 for transport
  - Distributes IPv6 prefixes
  - Uses multicast group addresses FF02::5 (OSPF IGP) and FF02::6 (OSPF IGP Designated Routers)
  - Runs over a link rather than a subnet
  - Multiple instances per link
  - Topology not IPv6-specific
    - Router ID, Area ID, Link ID remain a 4 bytes number
    - Neighbour routers are always identified by Router ID (4 bytes)
    - With an additional table with mapping between IPv6 prefixes and Link IDs
  - Uses link-local addresses as IPv6 source addresses

OSPFv3 - LSA Types

- Link LSA (Type 8)
  - Informs neighbour routers of link local address
  - Informs neighbour routers of IPv6 prefixes on link
- Intra-Area Prefix LSA (Type 9)
  - Associates IPv6 prefixes with a network or router
- Flooding scope for LSAs has been generalized
  - Three flooding scopes for LSAs
    - Link-local
    - Area
    - AS
- LSA Type encoding expanded to 16 bits
  - Includes flooding scope

Route Redistribution

- Domains with difference routing protocols can exchange routes.
- This is called route redistribution.
  - One-way redistribution – Redistributes only the networks learned from one routing protocol into the other routing protocol.
    - Uses a default or static route so that devices in that other part of the network can reach the first part of the network
  - Two-way redistribution – Redistributes routes between the two routing processes in both directions
- Static routes can also be redistributed.