Layer 3 - Routing
IPv4 and IPv6 Routing

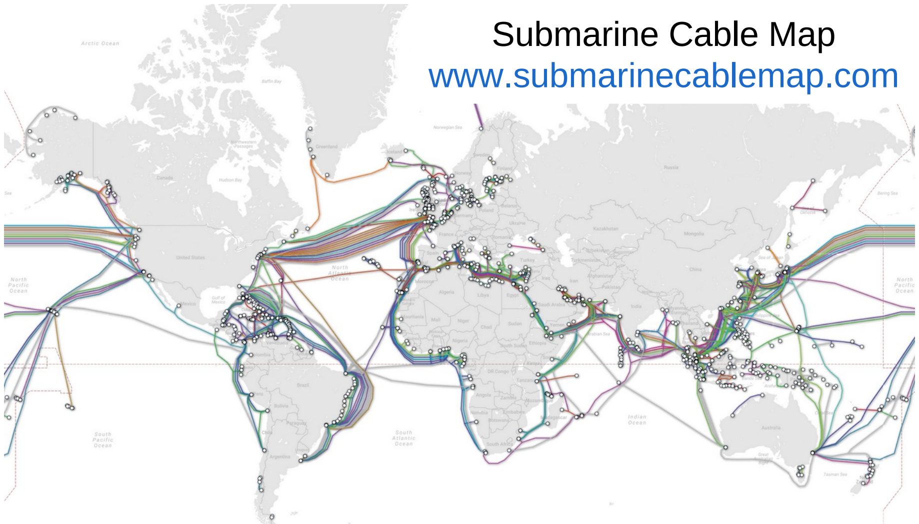
Fundamentos de Redes

**Mestrado Integrado em
Engenharia de Computadores e Telemática
DETI-UA**

universidade de aveiro
deti.ua.pt

Internet Physical Structure (1)

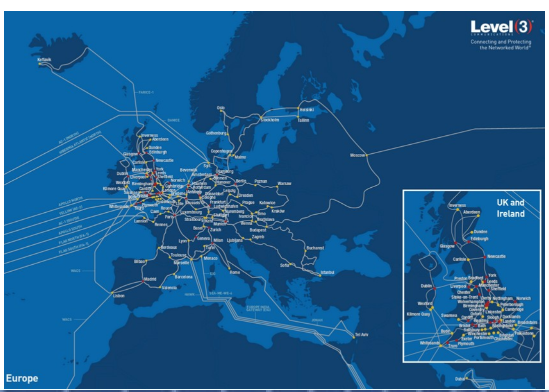
- Submarine Cables and IXs (majority of) information is public.
  - Internet Exchange (IX): Place where ISP exchange networking information/traffic.

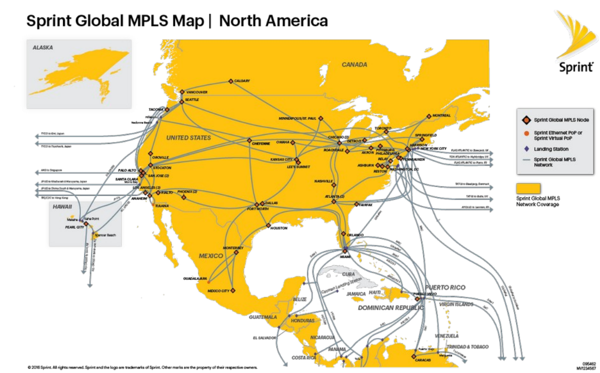


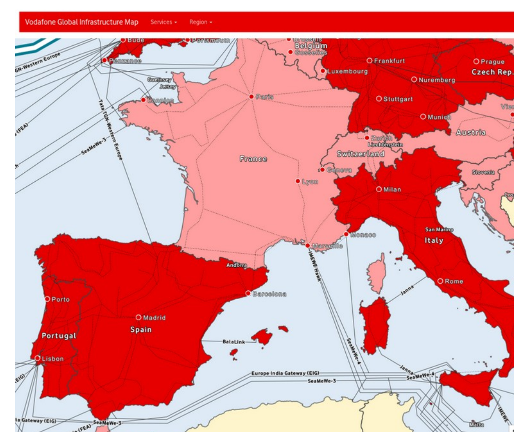
Submarine Cable Map
www.submarinecablemap.com

Internet Exchange Map
www.internetexchangemap.com

- Land Cables information is sparse (sometimes secret) and provider specific.

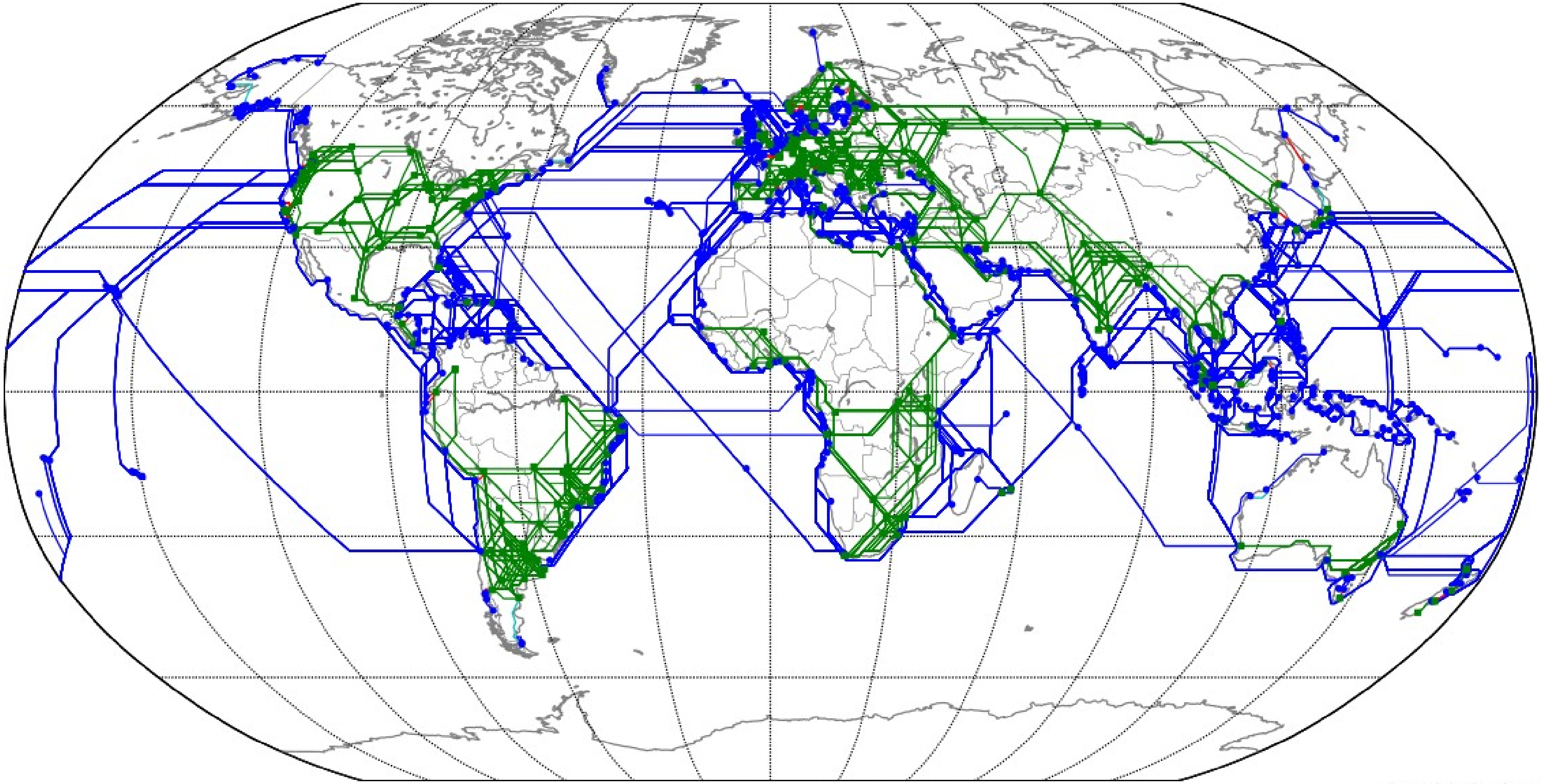






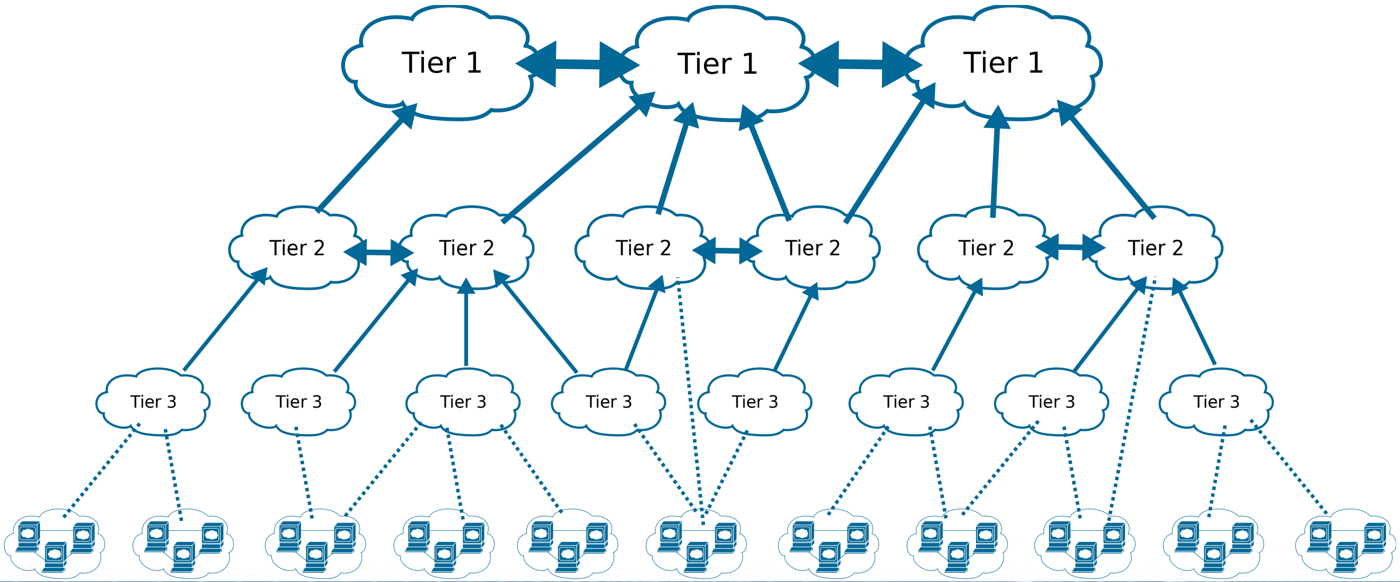
Internet Physical Structure (2)

- It is only possible to have an approximated overview of the Internet physical structure based on extrapolations of publicly available information.



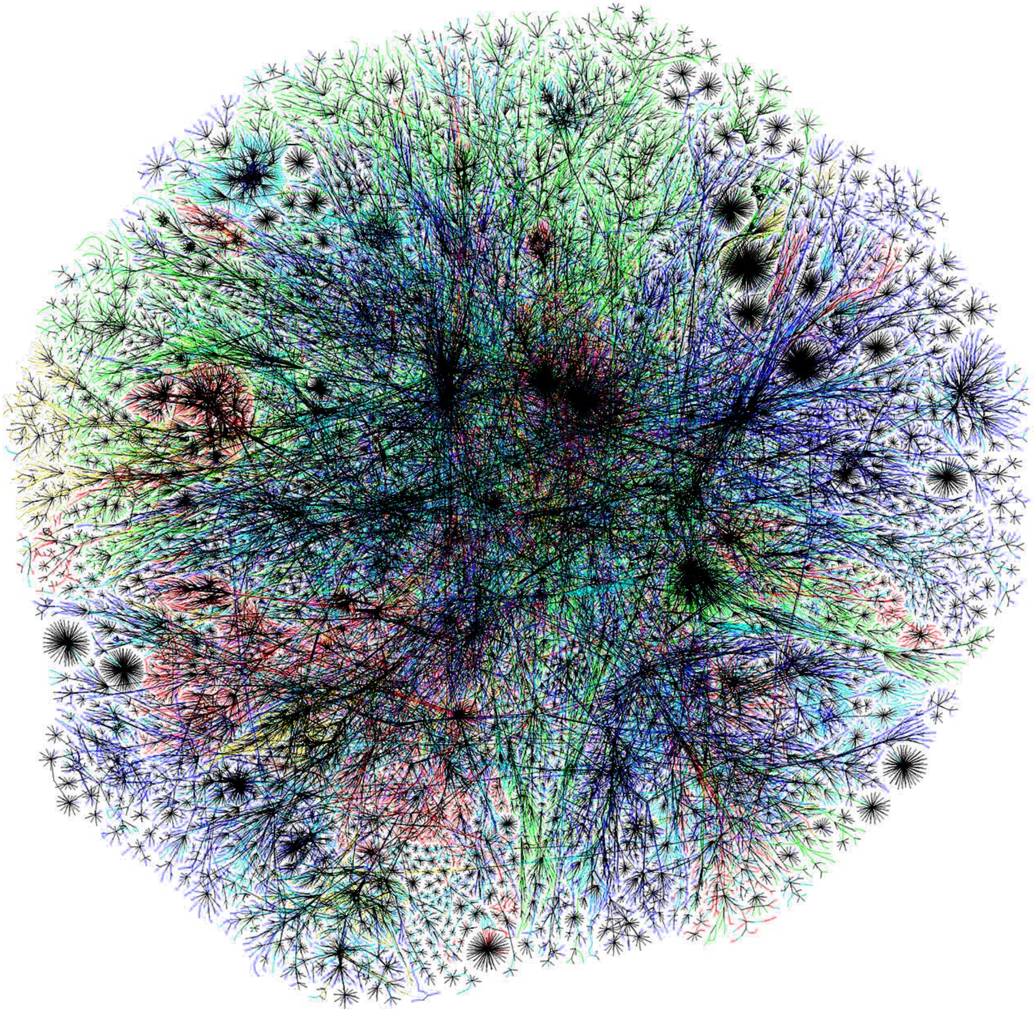
Internet Logical Structure (1)

- ISPs must agree on how there are going to exchange traffic (Peering agreement).
  - One ISP may only transport traffic from/to the other peer internal networks (non-transit), or
  - May transport traffic from/to any network the other peer sends towards it (transit).
- Tier ISP classification depends only how big the ISP is, in terms of geographical scope and how inter-operate with other ISPs.
  - Typically operate large high-capacity networks.
- Tier levels
  - Tier 1 ISPs often do not provide services to end-users, instead they provide Internet transit (transport of traffic from other ISPs networks).
  - Tier 2 ISPs also provide Internet transit at a non-global level, but require Tier1 ISPs to achieve global connectivity. May provide services to end-users.
  - Tier 3 operate at a regional level, provide services to end-users, and require Tier1 and Tier2 ISPs to achieve global connectivity.



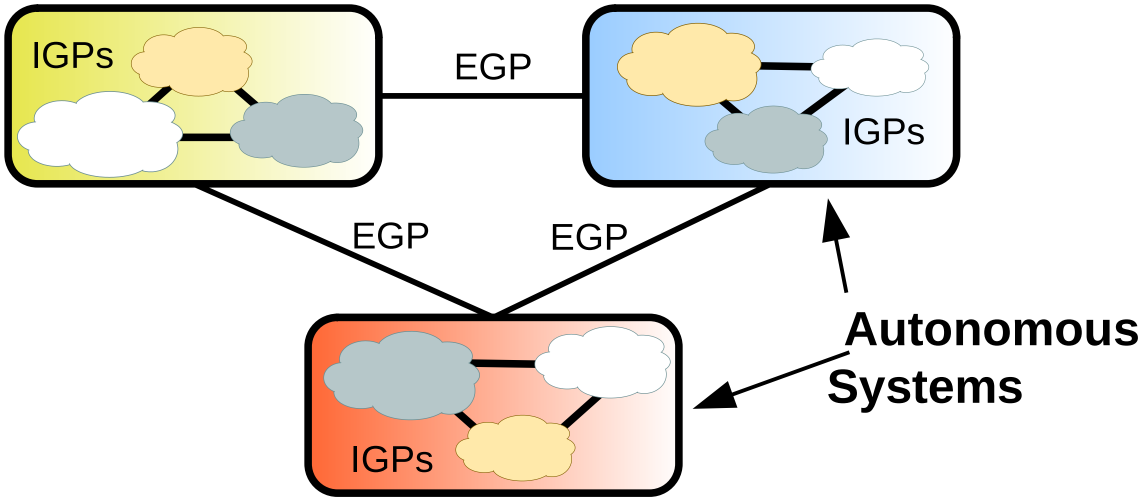
Internet Logical Structure (2)

- Internet complexity:



Autonomous Systems

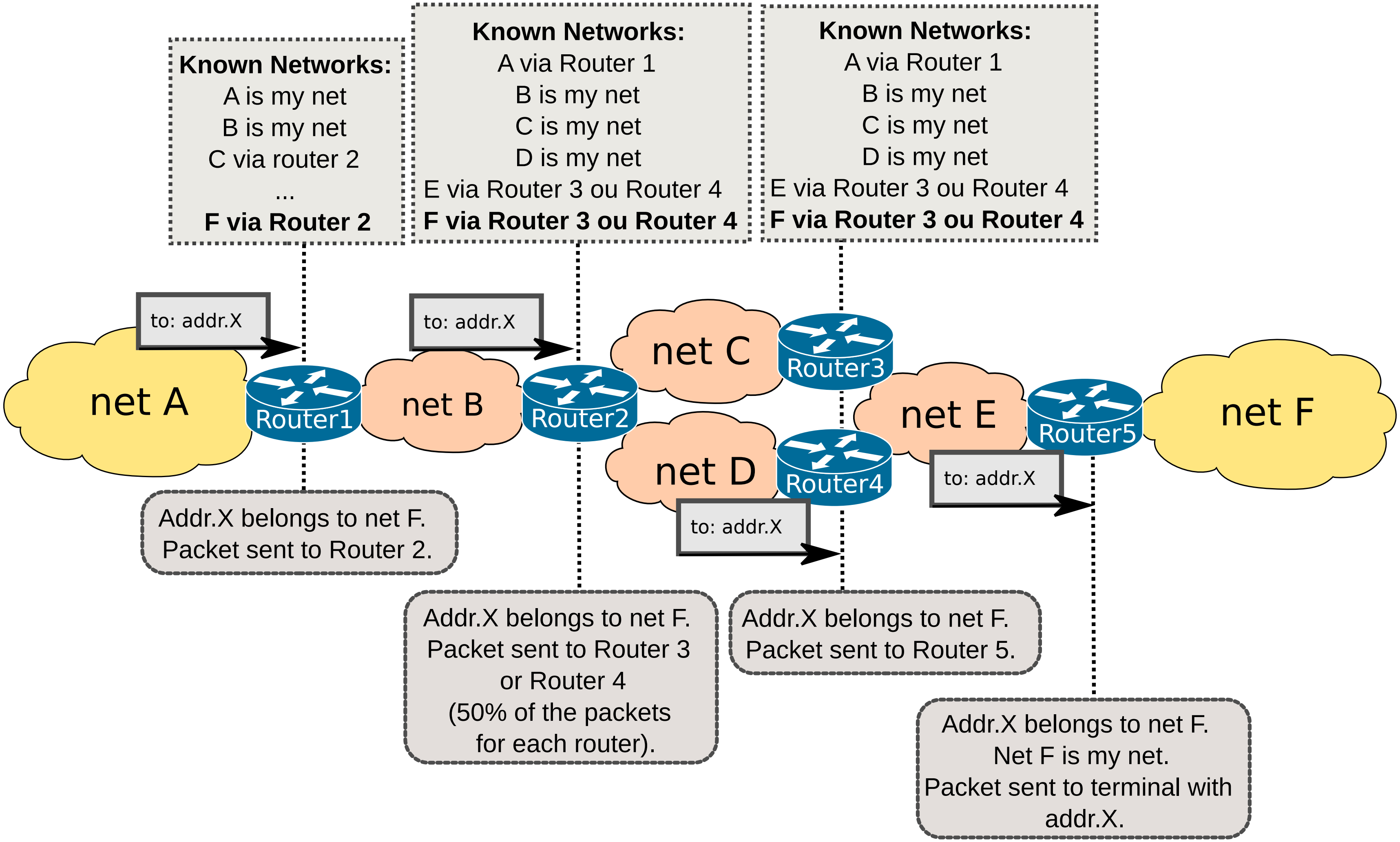
- AS (Autonomous System) – set of routers/networks with a common routing policy and under the same administration.
- Routing inside an AS is performed by IGPs (Interior Gateway Protocols) such as RIPv1, RIPv2, RIPng, OSPFv2, OSPFv3, IS-IS and EIGRP.
  - Called Internal Routing
- Routing between AS is performed by EGPs (Exterior Gateway Protocols) such as BGP and MP-BGP.
- IGPs and EGPs have different objectives:
  - IGPs: optimize routing performance
  - EGPs: optimize routing performance obeying political, economic and security policies.



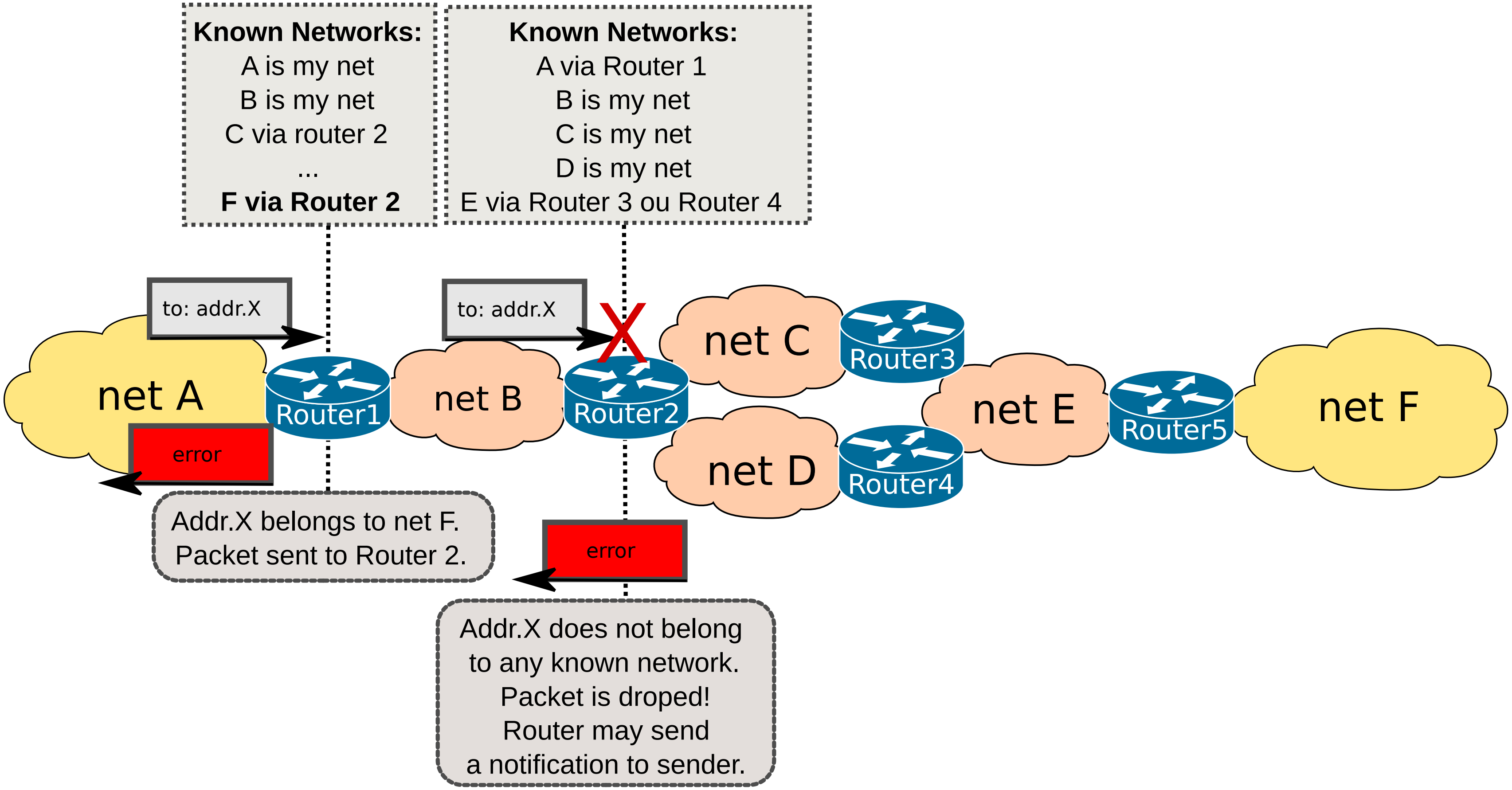
IP Routing Overview (1)

- Routers forward packets toward destination networks.
- Routers must be aware of destination networks to be able to forward packets to them.
- A router knows about the networks directly attached to its interfaces
- For networks not directly connected to one of its interfaces, however, the router must rely on outside information.
- A router can be made aware of remote networks by:
  - **Static routing**: An administrator manually configure the information.
  - **Dynamic routing**: Learns from other routers.
  - **Routing policies**: Manually routing rules that outweigh static/dynamic routing.
- If a packet for un unknown network reaches the router this will drop the packet, and MAY notify the sender about the routing error.

IP Routing Overview (2)



IP Routing Overview (3)



Routing Tables (1)

- Define how a remote network is reachable:
  - Next-hop (identified by its address), and
  - Local interface that provides connection.
- A network may be reachable using more than one path: (next-hop,local interface) pair.
- Mandatory elements
  - **Destination prefix**
  - **Destination mask**
  - **Metric**
    - Could be defined by key tags.
      - e.g., Directly Connected
  - One or both
    - **Next-hop address**
    - **Output interface**
- Optional elements
  - **Administrative distance**
  - **Protocol**
  - **Entry age (last time information received)**
  - **Scope**
  - **Flags**
  - **Source-specific**
- The next path hop (next hop address) may be found using more than one table entry (recursive resolution).
  - e.g., Network A is reachable through address from network B, Network B is reachable through address from network C, …
- The next-hop address may be obtained from external information (configurations or other mechanisms).
  - e.g., Tunnels, Point-to-point connections, etc…
- When an entry uses a next-hop address from an unknown network, that entry is removed.
- All entries obtain by dynamic methods may have an entry age (time since last update/confirmation).
  - After a timeout value without an update/confirmation the entry is removed.

Cisco IOS

```
R     200.1.1.0/24 [120/1] via 200.19.14.10, 00:00:16, FastEthernet0/1
      200.19.14.0/24 is variably subnetted, 2 subnets, 2 masks
C        200.19.14.0/24 is directly connected, FastEthernet0/1
L        200.19.14.4/32 is directly connected, FastEthernet0/1
R     200.38.0.0/24 [120/1] via 200.43.0.8, 00:00:03, FastEthernet1/1
      200.43.0.0/24 is variably subnetted, 2 subnets, 2 masks
C        200.43.0.0/24 is directly connected, FastEthernet1/1
L        200.43.0.1/32 is directly connected, FastEthernet1/1
```

Linux: route -n

```
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
0.0.0.0         193.136.92.1    0.0.0.0         UG    100    0        0 enp5s0f1
169.254.0.0     0.0.0.0         255.255.0.0     U     1000   0        0 enp5s0f1
193.136.92.0    0.0.0.0         255.255.254.0   U     100    0        0 enp5s0f1
```

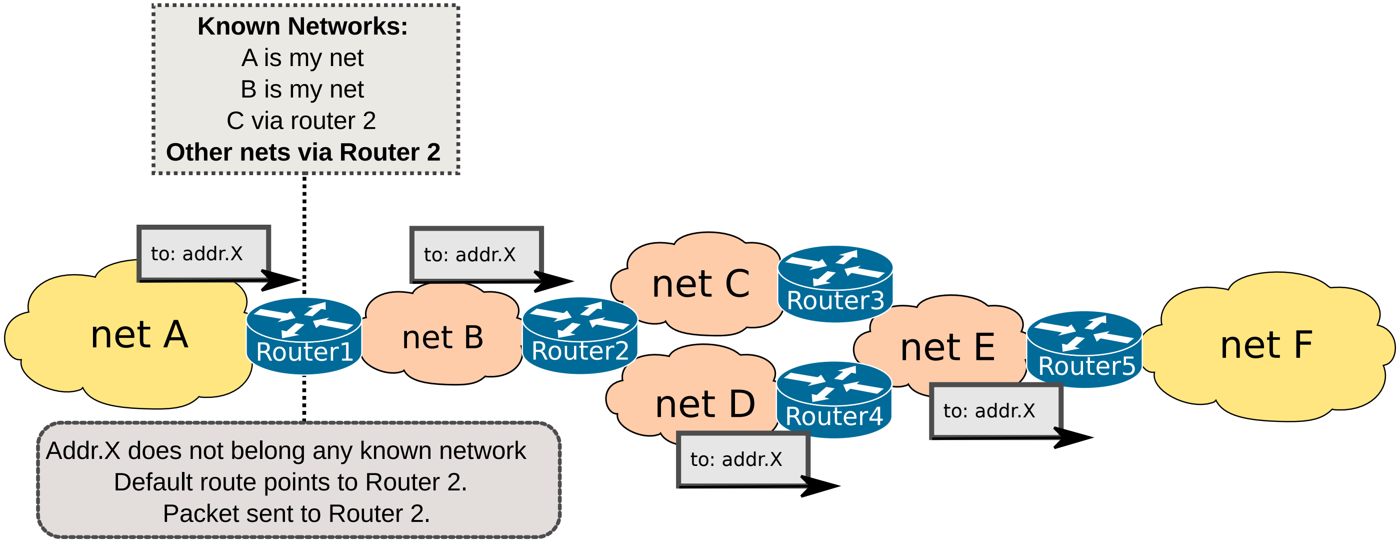
Linux: ip route

```
default via 193.136.92.1 dev enp5s0f1  proto static  metric 100
169.254.0.0/16 dev enp5s0f1  scope link  metric 1000
193.136.92.0/23 dev enp5s0f1  proto kernel  scope link  src 193.136.93.104  metric 100
```

Routing Tables (2)

- An IP address may have multiple matches on a Routing Table:
    - Example: 192.168.1.12
    - Will match:
        - 192.168.1.0/25 via ...
        - 192.168.1.0/24 via ...
        - 192.168.0.0/23 via ...
        - 192.168.0.0/16 via ...
        - …
  - Router will choose entry with the largest network prefix (most specific network).
        - i.e., 192.168.1.0/25 via ...
- Load balancing
  - Routing tables may have more than one path for each network
    - Traffic will be divided by all entries.
    - By packet, flow (TCP session, UDP IPs/port), etc...
        - E.g, packet 1 path 1, packet 2 path 2, packet 3 path 1, …
        - Flow 1 path 1, flow 2 path 2, flow 3 path 3, flow 4 path 1, flow 5 path 2, ...

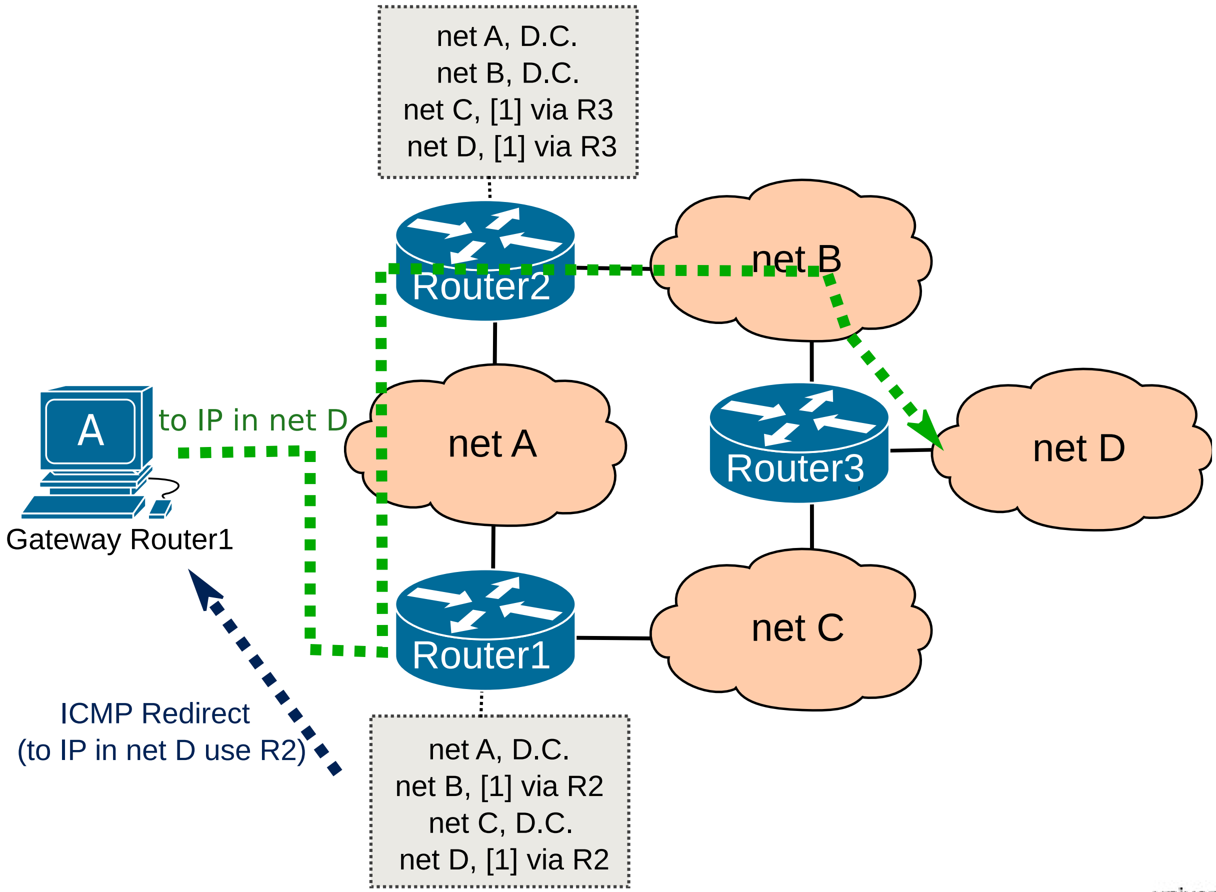
Default Routes



- In some circumstances, a router does not need to recognize the details of remote networks.
- The router can be configured to send all traffic (or all traffic for which there is not a more specific entry in the routing table) to a specific neighbor router.
- This is known as a default route.
- Default routes are either dynamically advertised using routing protocols or statically configured.
- IPv4 default route - 0.0.0.0/0
- IPv6 default route - ::/0

ICMP Redirect

- When a Router has a path defined via the interface from which received a packet:
  - The router sends an **ICMP Redirect** to the sender, defining a new gateway for that specific destination (IP address),
  - The sender updates its routing table for that specific destination address.



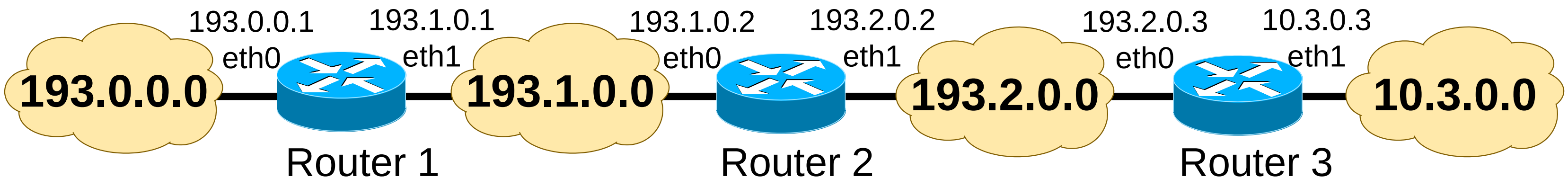
Administrative Distance

- Most routing protocols have metric structures and algorithms that are incompatible with other protocols.
- It is critical that a network using multiple routing protocols be able to seamlessly exchange route information and be able to select the best path across multiple protocols.
- Routers use a value called administrative distance to select the best path when they learn from different routing protocols the same destination (same network prefix and mask length).
- The Protocol/Method with the lowest Administrative Distance is preferred
  - The Administrative Distance value is configurable.
- Example:
  - Static [**1**/0] 192.168.1.0/24 via … ← Chosen!
  - RIP [**120**/1] 192.168.1.0/24 via …
  - OSPF [**110**/1] 192.168.1.0/24 via ...

Static Routing

- Stating routing do not react to network topology changes.
  - If a link fails, the static route is no longer valid if it is configured to use that failed link, so a new static route must be configured.
  - Connectivity may be lost until intervention of an administrator.
- Static routing does not scale well when network grows.
  - Administrative burden to maintain routes may can become excessive.
- Static routes can be used in the following circumstances:
  - When the administrator needs total control over the routes used by the router.
  - When a backup to a dynamically recognized route is necessary.
  - When it is used to reach a network accessible by only one path (a stub network).
    - There is no backup link, so dynamic routing has no advantage.
  - When a router connects to its ISP and needs to have only a default route pointing toward the ISP router, rather than learning many routes from the ISP.
    - Again, a single path of access without backup.
  - When a router is underpowered and does not have the CPU or memory resources necessary to handle a dynamic routing protocol.
  - When it is undesirable to have dynamic routing updates forwarded across low bandwidth links.

Static Routing Examples



- Example 1
  - Router2 do not know networks 193.0.0.0/24 and 10.3.0.0/24
  - Necessary static routes:
    - 193.0.0.0/24 accessible through 193.1.0.1 (eth1, Router1)
    - 10.3.0.0/24 accessible through 193.2.0.3 (eth0,Router3)
- Example 2
  - Router1 do not know networks 193.2.0.0/24 and 10.3.0.0/24
  - Necessary static routes:
    - 193.2.0.0/24 accessible through 193.1.0.2 (eth0, Router2)
    - 10.3.0.0/24 accessible through 193.1.0.2 (eth0, Router2)

      OR
    - Using default route: 0.0.0.0/0 accessible through 193.1.0.2 (eth0, Router2)

Dynamic Routing

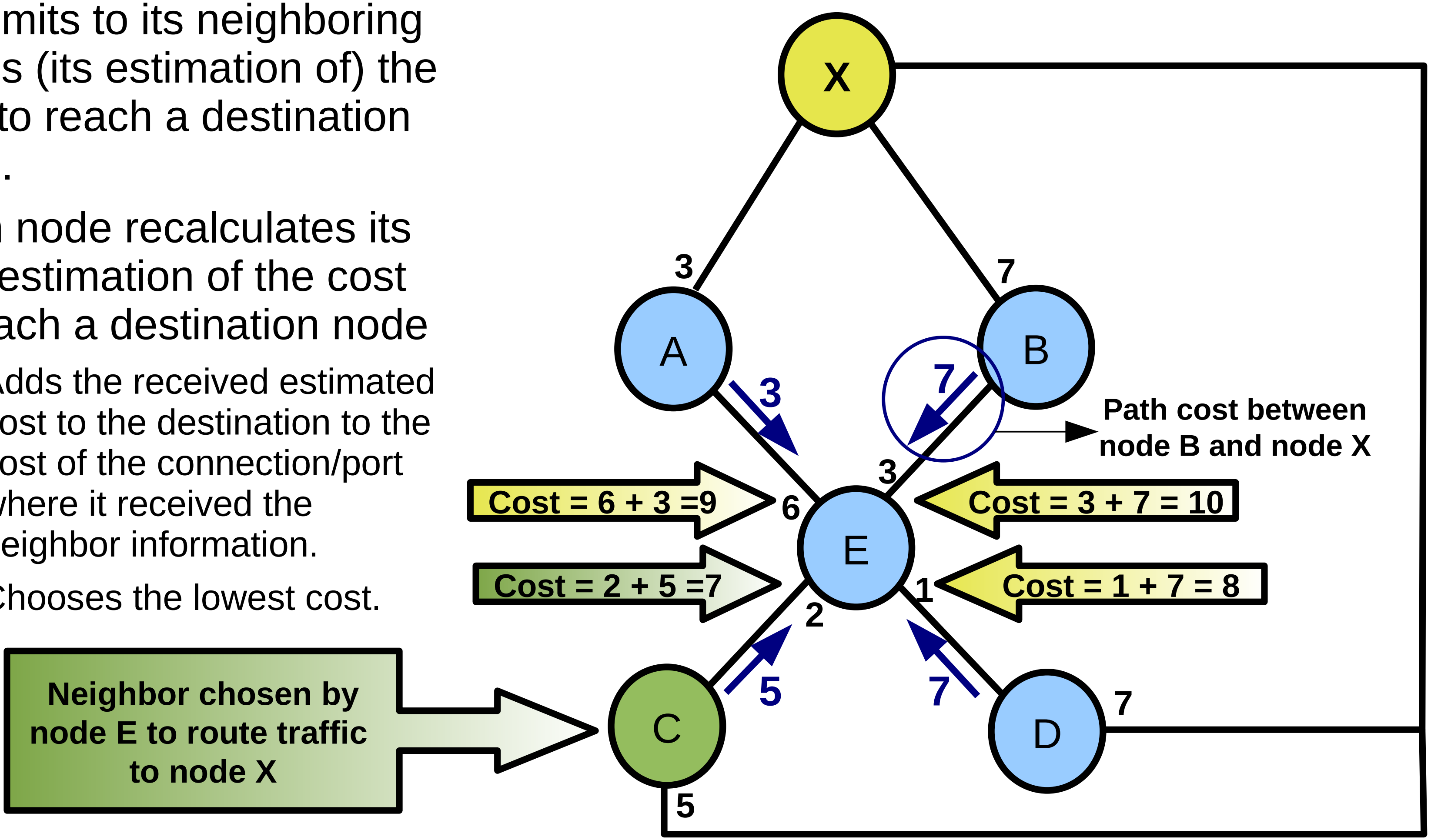
- Dynamic routing allows the network to adjust to changes in the topology automatically, without administrator involvement.
- Routers exchange information about the reachable networks and the state of each network/link.
  - Routers exchange information only with other routers running the same routing protocol.
  - 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.

Distance Vector vs. Link State

- *Distance vector*
  - Each router knows only the distance/cost to all network destinations.
  - Information (distances to all known destinations) is periodically sent by routers to its neighbors
  - Each router determines the lowest cost paths to all destinations based on a distributed and asynchronous version of the Bellman-Ford algorithm
  - Examples: RIP, IGRP, EIGRP
- *Link state*
  - Routers know the complete network topology.
  - Use a centralized algorithm to determine the lowest cost path to all destinations.
  - The required information to build and maintain the network topology in each
  - Examples: OSPF, IS-IS

Distributed and Asynchronous Bellman-Ford Algorithm

- Each node periodically transmits to its neighboring nodes (its estimation of) the cost to reach a destination node.
- Each node recalculates its own estimation of the cost to reach a destination node
  - Adds the received estimated cost to the destination to the cost of the connection/port where it received the neighbor information.
  - Chooses the lowest cost.



RIP (Routing Information Protocol)

- Is a *distance vector* protocol
  - Each router maintains a list of known networks and, for each network, an estiamtion of the cost to reach it – this is called a distance vector.
  - Each router periodically send to its neighboring routers its own distance vector (partially or complete) – announcement/update.
  - Each router uses the distance vector sent by its neighbors to update its own distance vector.
- The path cost to a destination is given by the number of routers/hops in the path.
  - Maximum cost is 15.
  - A cost of 16 is considered infinite (or unattainable destination).
- Each router determines the entries in its own routing table, based on the constructed distance vector.
  - For each destination (network) learned, it adds an entry to that network that uses the path (or paths) with the lowest cost, using as next-hop the neighboring router(s) that announced that network with that lowest cost path.

RIP Version 1

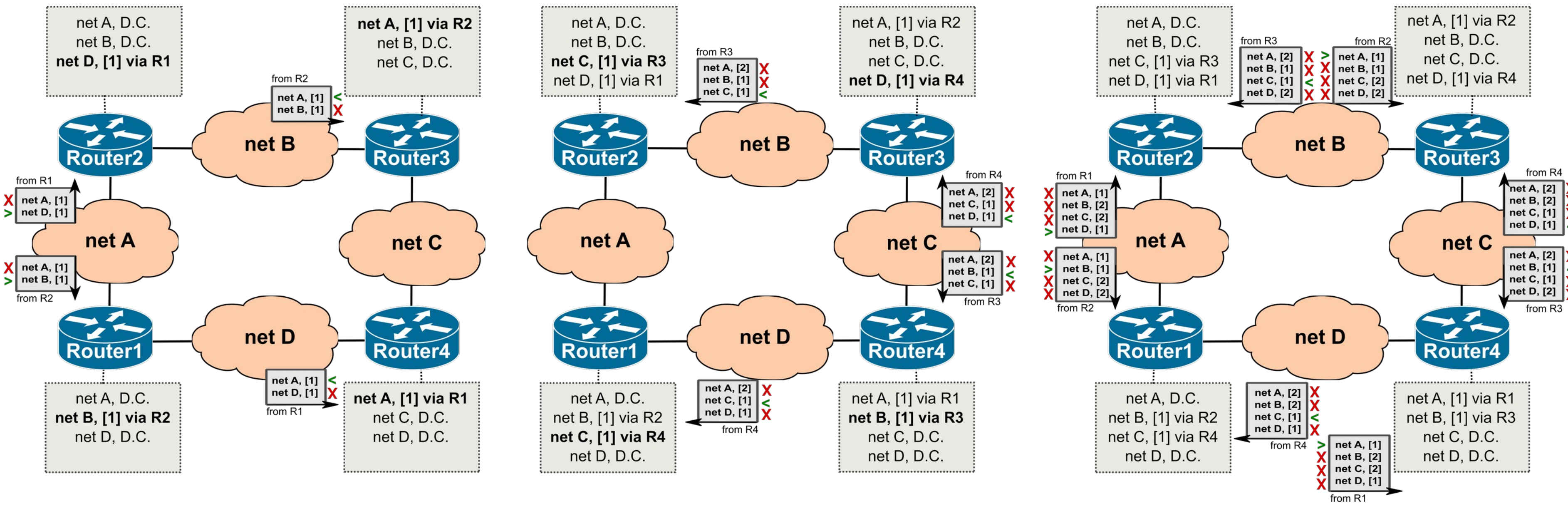
- RIP Version 1 (RIPv1) is a classfull protocol.
  - Does not announces (sub-)networks masks, only network prefixes.
  - Network masks are assumed based on the incoming interface mask.
    - If all networks have the same mask it works perfectly, however, when networks with different masks exist it is problematic.
- RIPv1 uses the broadcast address 255.255.255.255 to send announcements/updates.
  - All network devices must process the packets.
- Does not support authentication.
  - Messages may be forged by an attacker.

RIP Version 2

- RIP Version 2 (RIPv2) is a *classless* protocol.
  - RIPv2 announcements include network prefix and mask.
  - Supports variable length masks.
- RIPv2 used the multicast address 224.0.0.9 to send announcements/updates only to routers running RIPv2.
- RIPv2 supports authentication using message-digest and clear text password.
  - Clear text password authentication should not be used!

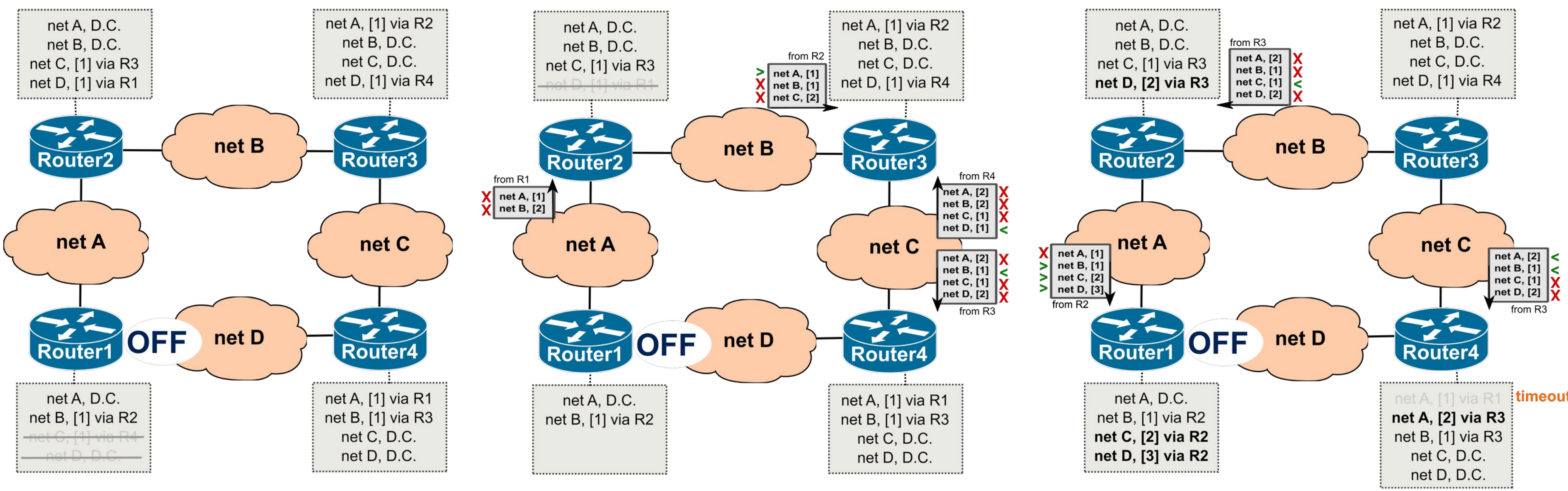
RIP Algorithm (1)

- Assuming that Router1 and Router2 send announcements first.
  - With split-horizon disabled.



RIP Algorithm (2)

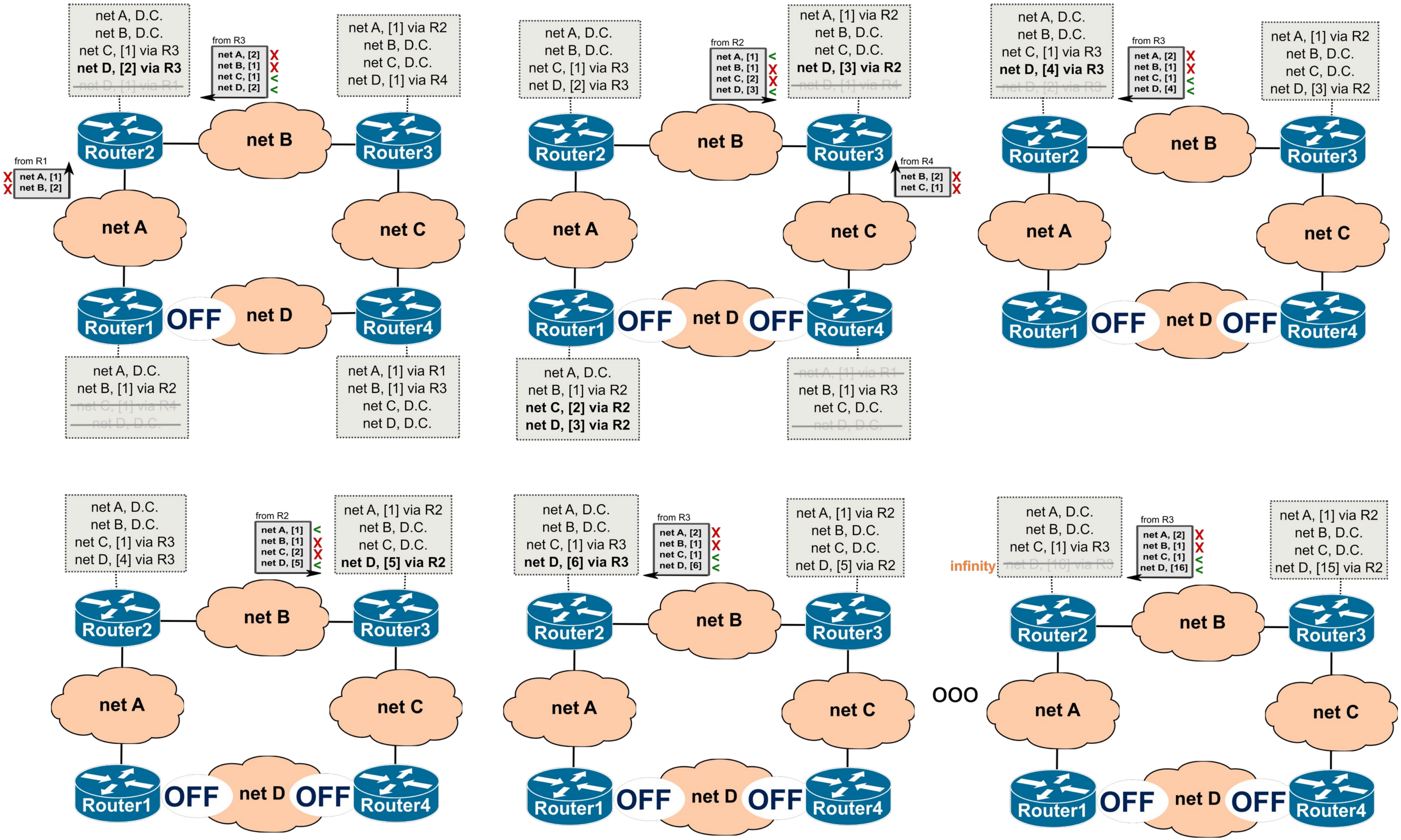
- Assuming Router1 connection to network D goes down.
  - No triggered updates.



universidade de aveiro

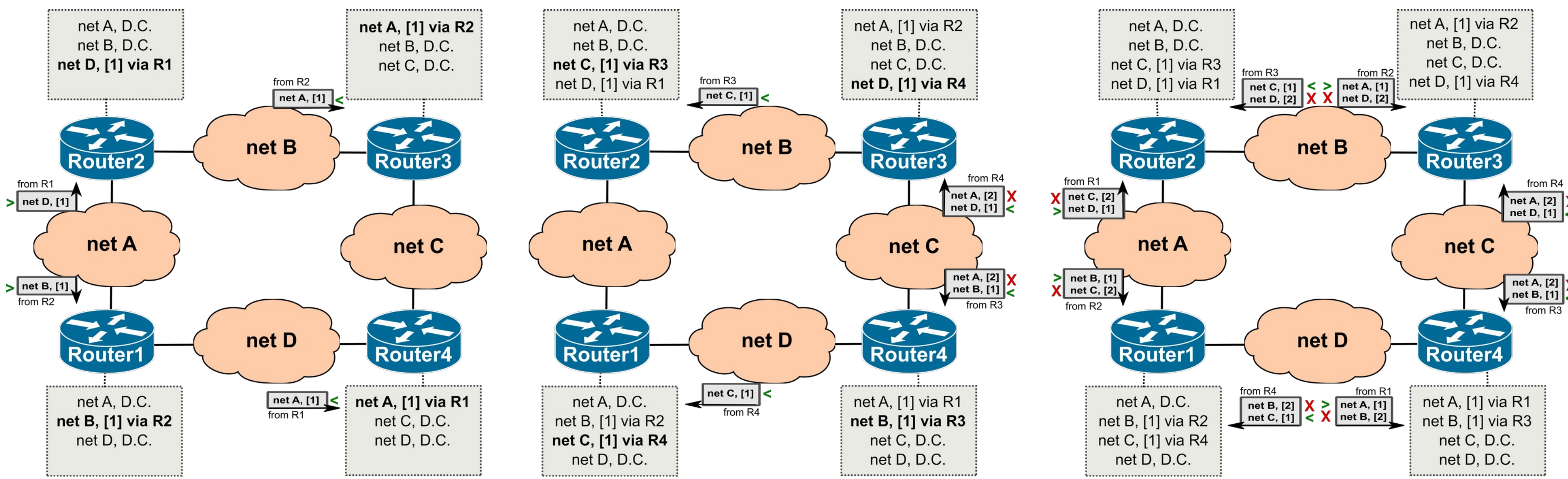
Count to Infinity Problem

- When multiple failures occur before algorithm convergence!



Split-Horizon (1)

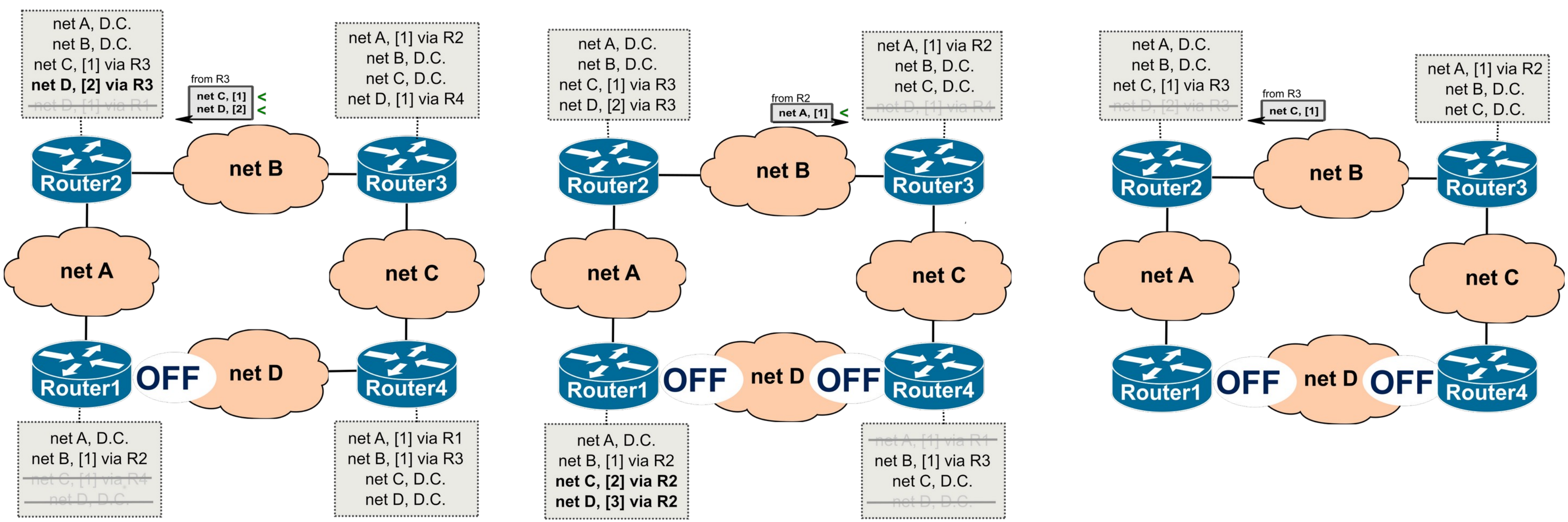
- Solution for the count to infinity problem.
- Each Router, in each interface, announces only the networks in which that interface is not used to provide the best path to that destination.
- Split horizon lowers the convergence time of the routing tables when there is a topology change.
  - RIPv1 e RIPv2 supports it.
- Assuming Router1 and Router2 start sending announcements first:



- In Split horizon with Poisoned Reverse, routers announce all networks but set metric to infinity (16) for networks learned by the interface by which they are sending the announcement.
  - Larger update messages.

Split-Horizon (2)

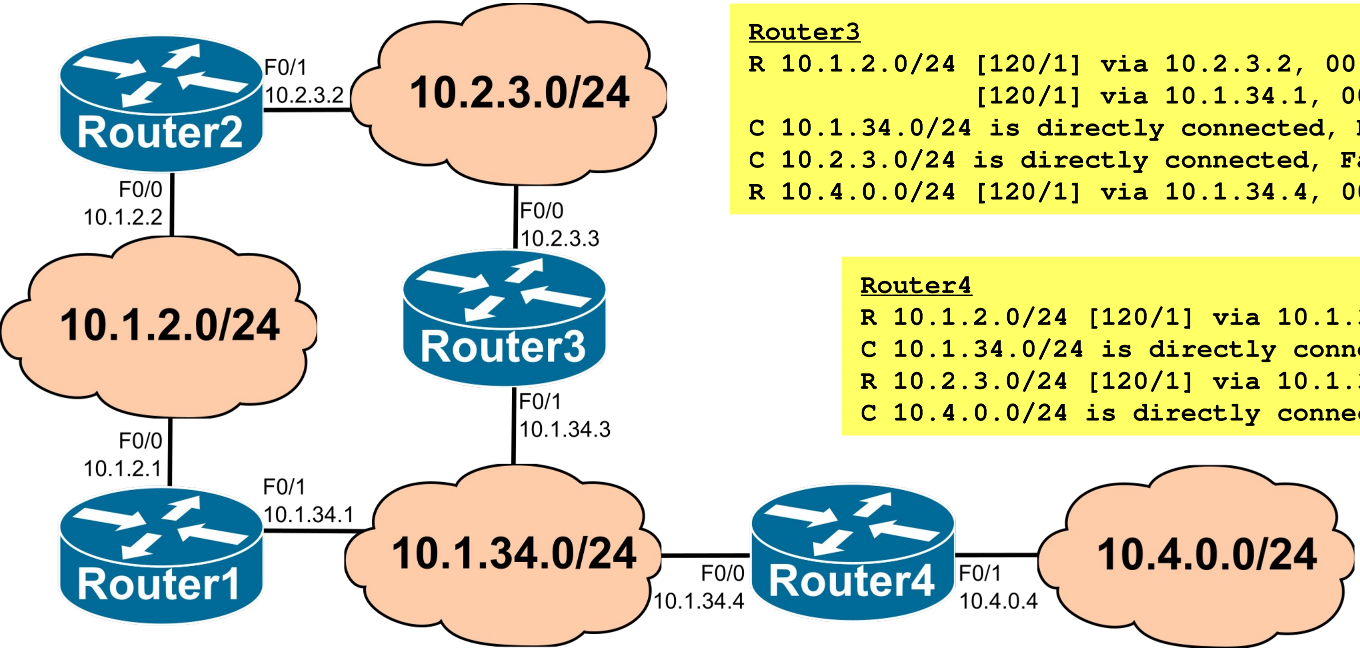
- Solution for the count to infinity problem.
- Prevents any routing loops that involve two routers.
  - It is possible to end up with patterns in which three or more routers are engaged in mutual deception.
- Assuming Router1 and Router4 loose connection to network D almost simultaneously:



universidade de aveiro

Routing Tables with RIP

```
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, 00:00:21, FastEthernet0/1
               [120/1] via 10.1.2.1, 00:00:11, 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, 00:00:21, FastEthernet0/1
              [120/2] via 10.1.2.1, 00:00:11, FastEthernet0/0
```




```
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, 00:00:28, 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, 00:00:24, FastEthernet0/1
```

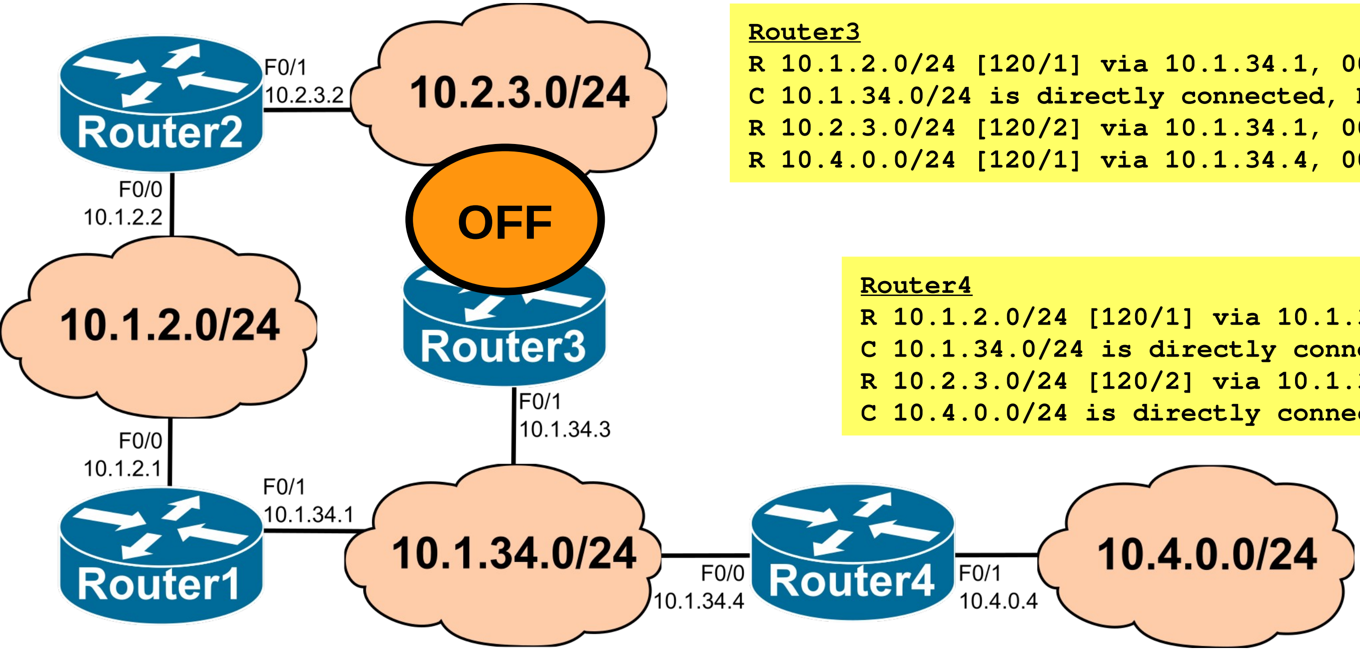
```
Router4
R 10.1.2.0/24 [120/1] via 10.1.34.1, 00:00:18, FastEthernet0/0
C 10.1.34.0/24 is directly connected, FastEthernet0/0
R 10.2.3.0/24 [120/1] via 10.1.34.3, 00:00:29, FastEthernet0/0
C 10.4.0.0/24 is directly connected, FastEthernet0/1
```

```
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, 00:00:11, FastEthernet0/1
              [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
```

Routing Tables with RIP

```
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, FastEthernet0/11
```




```
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, 00:00:29, 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
```

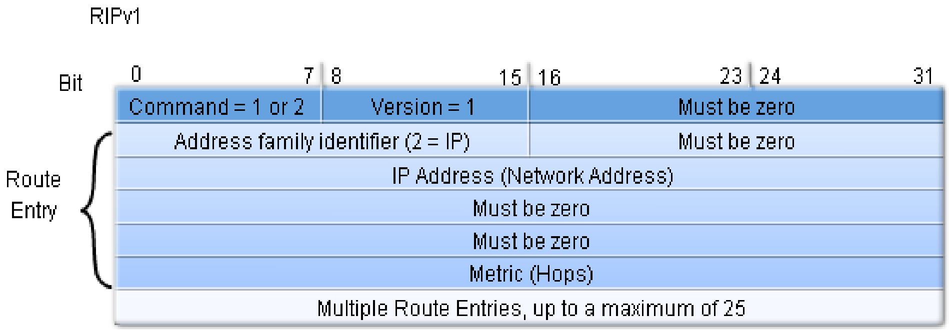
RIP Message Types

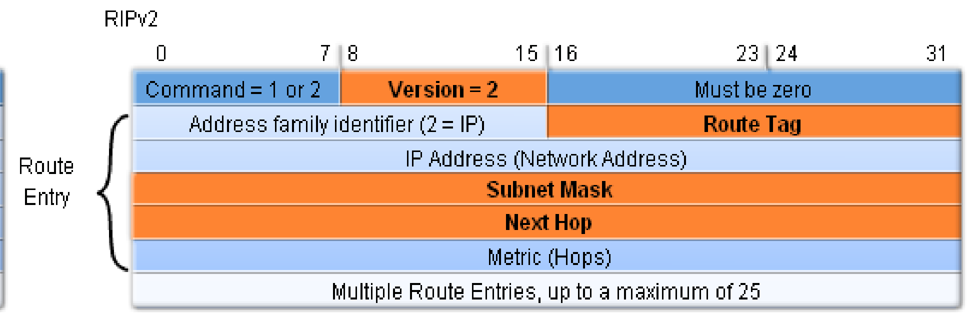
- RIP Response
  - Distance vector announcement/update message.
    - Contains the distance vector.
  - It is sent:
    - 1 – Periodically (~30 seconds by default, there is a random component).
    - 2 – Optionally, when some information changes (triggered updates).
    - 3 – In response to a RIP Request.

    - In cases 1 and 2:
      - In RIPv1, is sent to the broadcast address.
      - In RIPv2, is sent to the multicast address 224.0.0.9 (Routers com RIP).
    - In case 3, it is sent only (unicast) to the router that sent the RIP Request.
- RIP Request (Optional)
  - Sent by a router that was recently started (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 neighbor distance vector.

RIPv1 vs. RIPv2 Responses (1)

- New RIPv2 message fields in Response packets:
  - Subnet mask
    - Supports variable length masks.
    - Makes RIPv2 *classless* protocol.
  - Route tag
    - Attribute assigned to a specifc network that must be reserved a re-announced.
    - Provides a method to separate internal (to the RIP domain) and external networks.
  - Next hop
    - Address to which the packets must be routed.
    - 0.0.0.0 indicates that th epackets ,must be routed to the router that sent the RIP message.





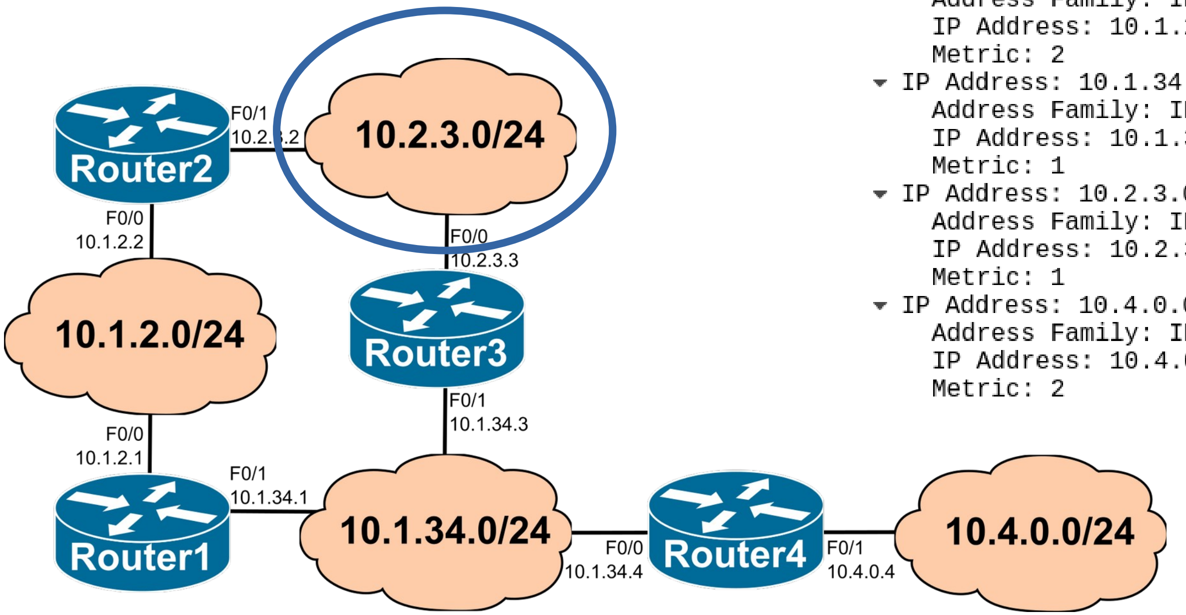
RIPv1 Messages (Example)

Sent by Router3 with Split-Horizon

```
▸ Internet Protocol Version 4, Src: 10.2.3.3, Dst: 255.255.255.255
▸ User Datagram Protocol, Src Port: 520, Dst Port: 520
▾ Routing Information Protocol
    Command: Response (2)
    Version: RIPv1 (1)
  ▾ IP Address: 10.1.34.0, Metric: 1
      Address Family: IP (2)
      IP Address: 10.1.34.0
      Metric: 1
  ▾ IP Address: 10.4.0.0, Metric: 2
      Address Family: IP (2)
      IP Address: 10.4.0.0
      Metric: 2
```

Sent by Router3 without Split-Horizon

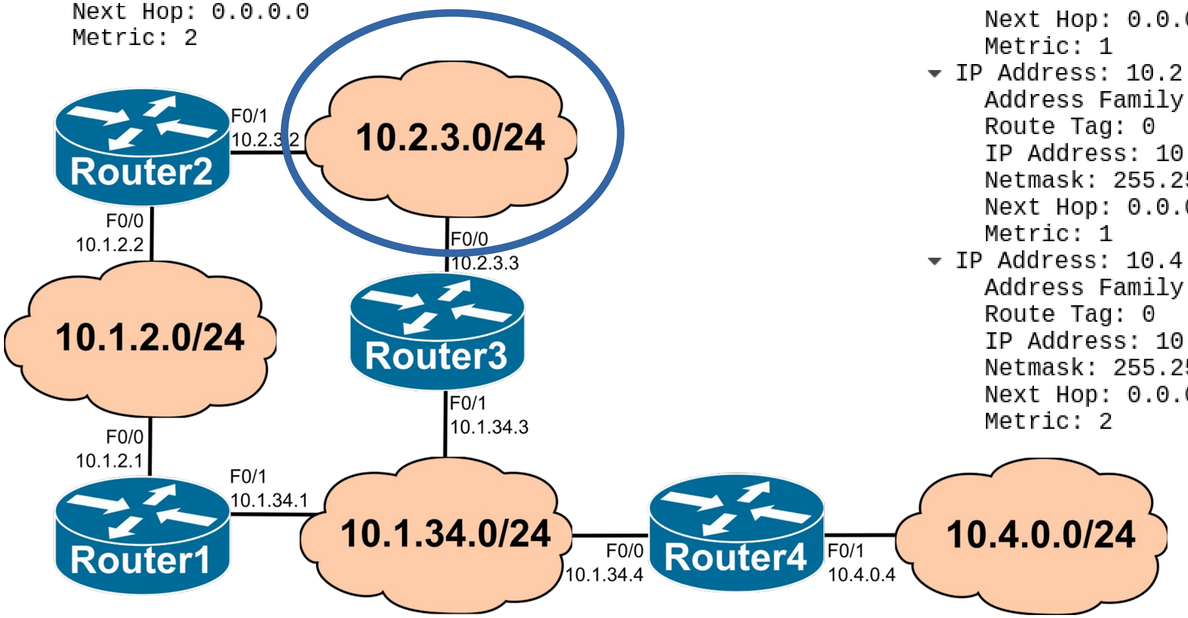
```
▸ Internet Protocol Version 4, Src: 10.2.3.3, Dst: 255.255.255.255
▸ User Datagram Protocol, Src Port: 520, Dst Port: 520
▾ Routing Information Protocol
    Command: Response (2)
    Version: RIPv1 (1)
  ▾ IP Address: 10.1.2.0, Metric: 2
      Address Family: IP (2)
      IP Address: 10.1.2.0
      Metric: 2
  ▾ IP Address: 10.1.34.0, Metric: 1
      Address Family: IP (2)
      IP Address: 10.1.34.0
      Metric: 1
  ▾ IP Address: 10.2.3.0, Metric: 1
      Address Family: IP (2)
      IP Address: 10.2.3.0
      Metric: 1
  ▾ IP Address: 10.4.0.0, Metric: 2
      Address Family: IP (2)
      IP Address: 10.4.0.0
      Metric: 2
```



RIPv2 Messages (Example)

Sent by Router3 with Split-Horizon

```
▸ 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
```

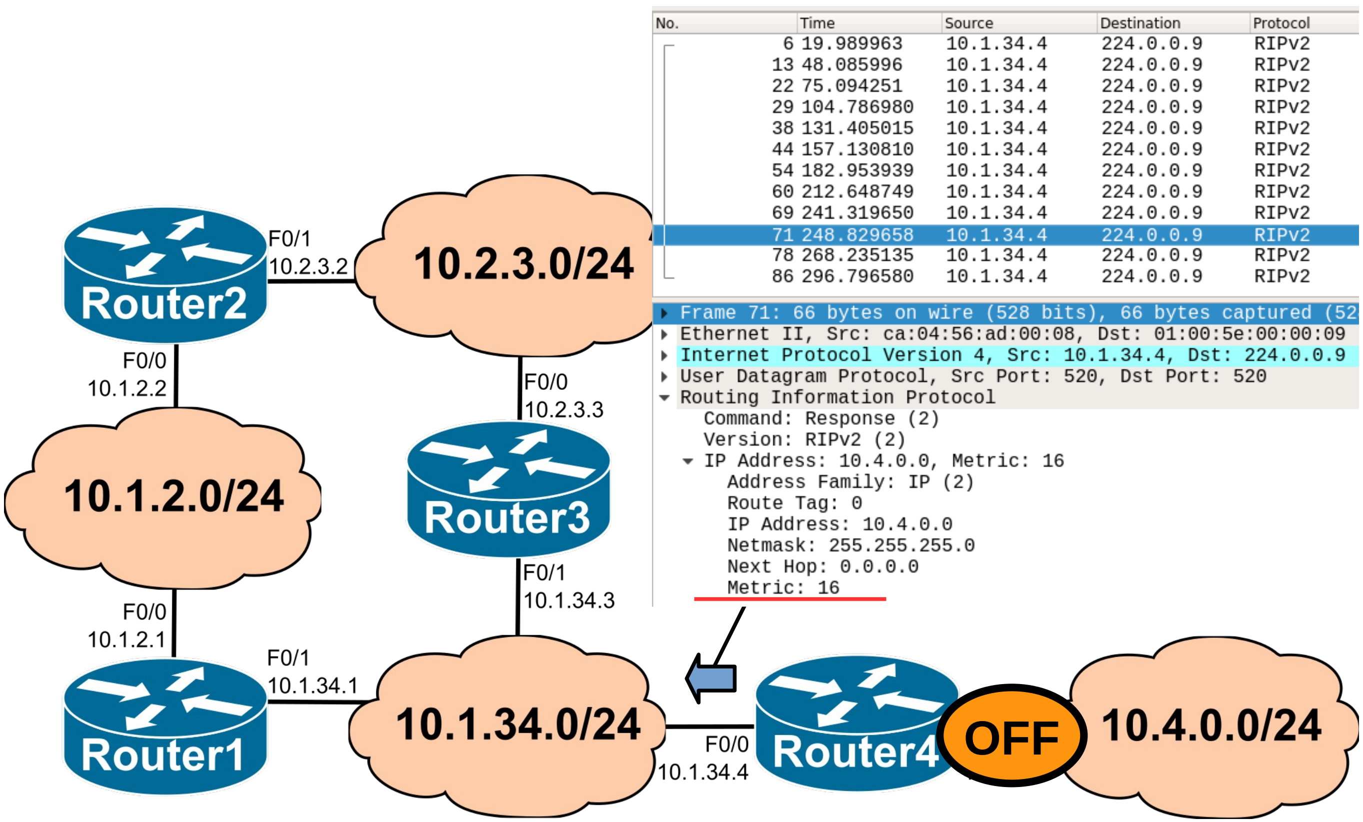


Sent by Router3 without Split-Horizon

```
▸ 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
```

Triggered Updates

- Prevents any routing loops that involve more than two routers.
- Whenever a router changes the metric for a route, it is required to send update messages almost immediately, even if it is not yet time for one of the regular update message.
- Neighboring routers update routing tables faster and overall convergence is faster.
  - Including entries that were removed by timeout!



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
```


← Small time gap!

RIPv2 Message Authentication

With Keyed Message Digest (MD5)

```
▸ 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)
  ▾ Authentication: Keyed Message Digest
      Authentication type: Keyed Message Digest (3)
      Digest Offset: 64
      Key ID: 1
      Auth Data Len: 20
      Seq num: 0
      Zero adding:
    ▾ Authentication Data Trailer
        Authentication Data: 7f7d4fc23f02a76b9986f517f3b6a8cf
  ▾ 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
```

With Clear Text Authentication
<u>Useless</u>!

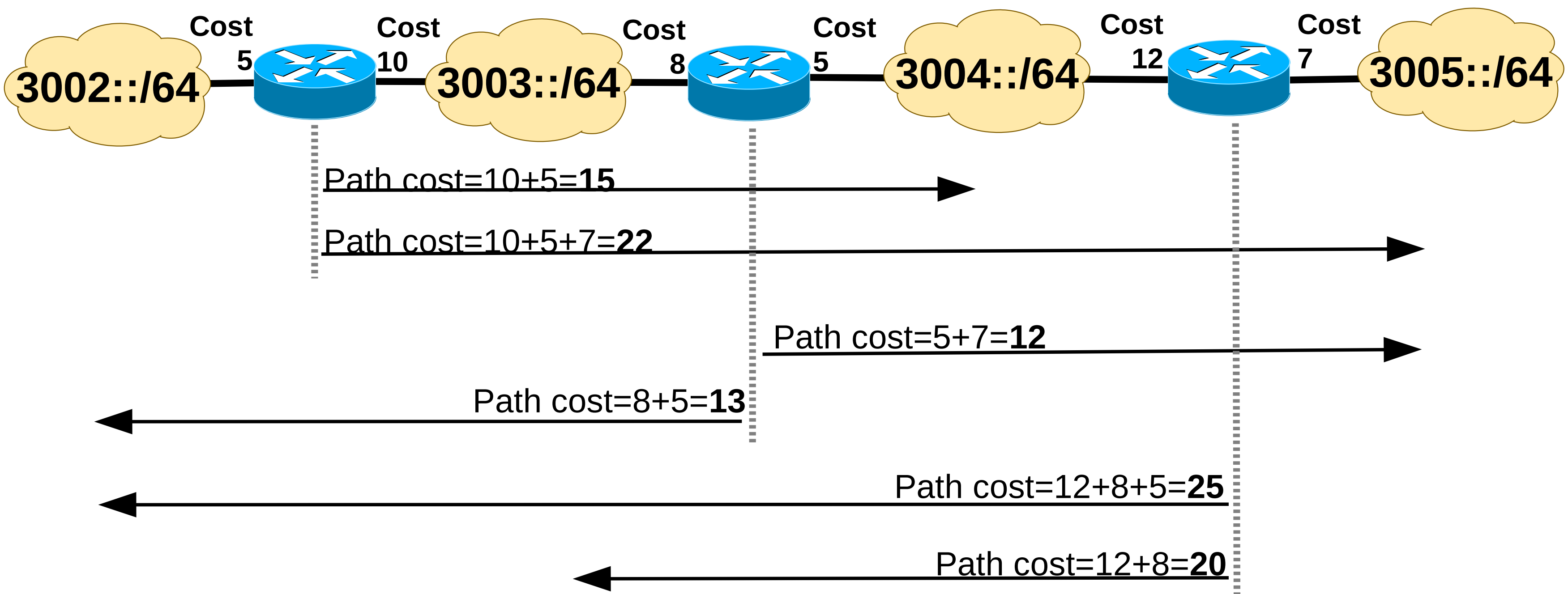
```
▸ Internet Protocol Version 4, Src: 10.2.3.3, Dst: 224.0.0.9
▸ User Datagram Protocol, Src Port: 520, Dst Port: 520
▾ Routing Information Protocol
    Command: Response (2)
    Version: RIPv2 (2)
  ▾ Authentication: Simple Password
      Authentication type: Simple Password (2)
      Password: labcom
  ▾ 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: 0.0.0.0
      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: 10.4.0.0, 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
```

RIPng for IPv6 Routing

- Similar to IPv4 RIPv2:
  - Distance-vector concept, radius of 15 hops, infinity metric is 16, split-horizon, triggered update.
- Differences between RIPv2 and RIPng
  - Uses IPv6 for transport.
    - Uses link-local addresses (not the global ones).
  - IPv6 prefix, next-hop IPv6 link-local address.
  - Uses multicast group address FF02::9 (all-RIP-routers) as the destination address for RIP updates.
  - Routers always add the cost of the interface to the metric received.
    - Metric is sum of “output interfaces” costs to destination and not number of hops.
    - If all costs are 1, metric is number of “output interfaces” to destination.
  - Allows for node/interface costs other than 1.
    - Cisco calls it “cost offset” per interface (out or in direction).
    - Cost to network is given by the <u>sum of all output interfaces costs</u> along the path.
    - With the infinity metric value at 16, this require careful configurations.
  - Routers always announce directed connected networks.
  - in IOS Cisco
    - Activation per interface, named process, more than one active process.

RIPng Path Costs

- Each router link/interface has an associated RIPng cost.
- The total cost between a router and a network is given by the sum of all RIPng costs of the (routers) output interfaces along the path.
  - Routers to access directly connect networks never use RIPng paths.

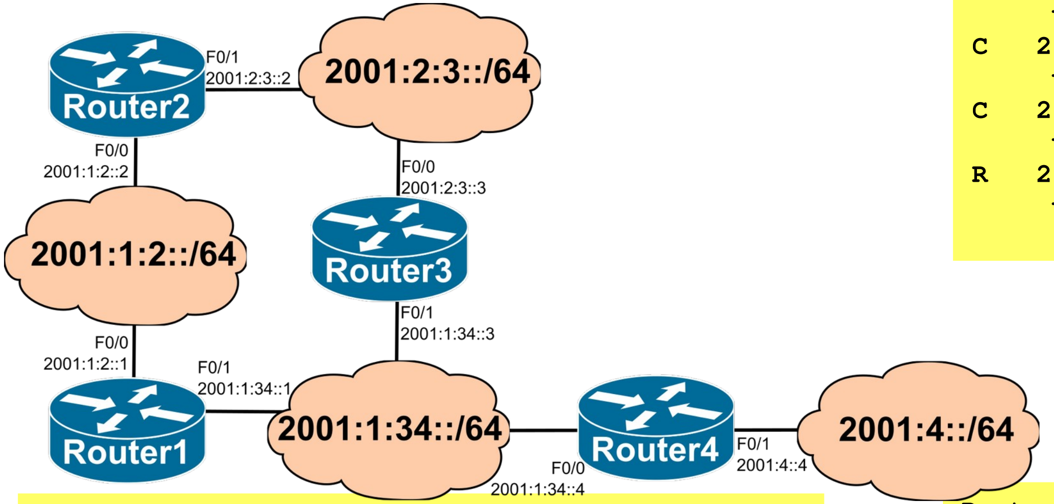


IPv6 Routing Tables with RIPng

```
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, FastEthernet0/1
```

Assuming all interfaces with cost 1.

```
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
```



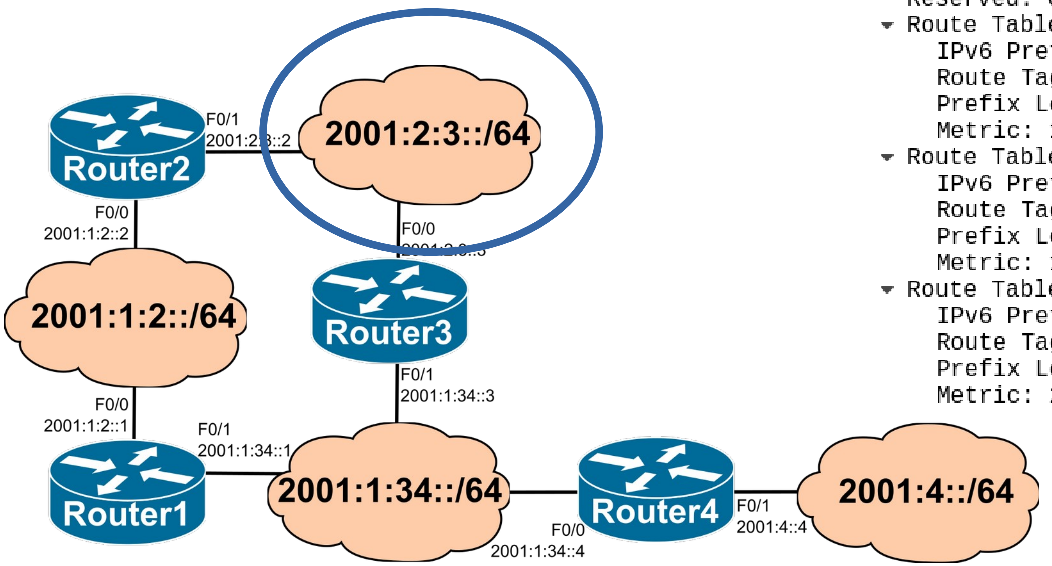

```
Router1
C    2001:1:2::/64 [0/0]
      via FastEthernet0/0, directly connected
C    2001:1:34::/64 [0/0]
      via FastEthernet0/1, directly connected
R    2001:2:3::/64 [120/2]
      via FE80::C802:54FF:FEF5:8, FastEthernet0/0
      via FE80::C803:56FF:FE0A:6, FastEthernet0/1
R    2001:4::/64 [120/2]
      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
```

RIPng Messages (Example)

Sent by Router2 with Split-Horizon

```
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
```



Sent by Router3 with Split-Horizon

```
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
```

Open Shortest Path First (OSPF) Protocol

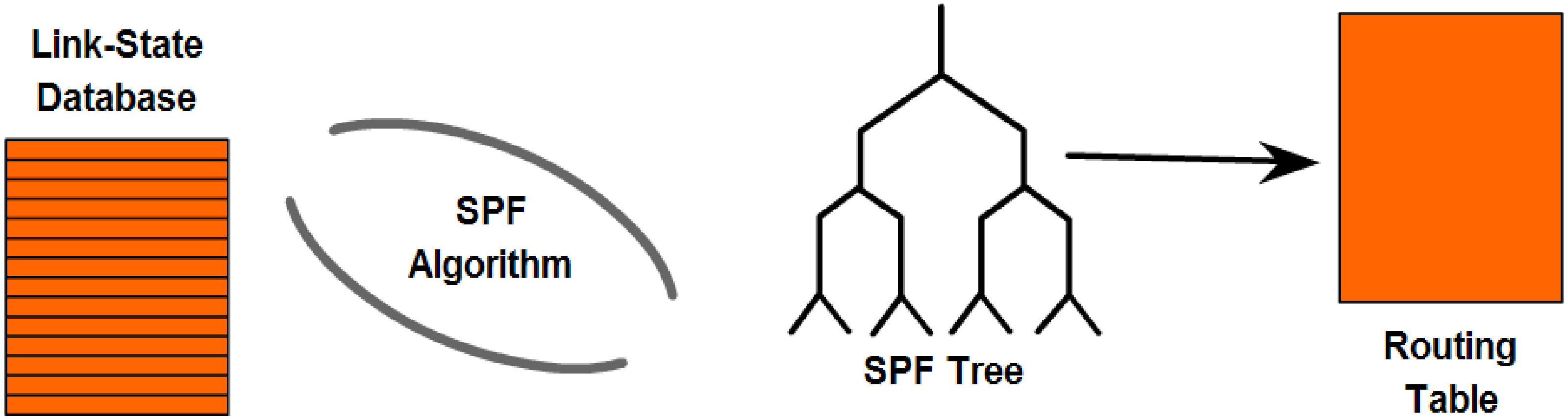
- OSPF is an open-standard protocol based primarily on RFC 2328.
- OSPF is a link-state routing protocol
  - Respond quickly to network changes,
  - Send triggered updates when a network change occurs,
  - Send periodic updates, known as link-state refresh, at long time intervals, such as every 30 minutes.
- Routers running OSPF collect routing information from all other routers in the network (or from within a defined area of the network)
- And then each router independently calculates its best paths to all destinations in the network, using Dijkstra’s (SPF) algorithm.

OSPF Necessary Routing Information

- For all the routers in the network to make consistent routing decisions, each link-state router must keep a record of the following information:
  - Its immediate neighbor routers
    - If the router loses contact with a neighbor router, within a few seconds it invalidates all paths through that router and recalculates its paths through the network.
    - For OSPF, adjacency information about neighbors is stored in the OSPF neighbor table, also known as an adjacency database.
  - All the other routers in the network, or in its area of the network, and their attached networks
    - The router recognizes other routers and networks through LSAs, which are flooded through the network.
    - LSAs are stored in a topology table or database (which is also called an LSDB).
  - The best paths to each destination
    - Each router independently calculates the best paths to each destination in the network using Dijkstra's (SPF) algorithm.
    - All paths are kept in the LSDB.
    - The best paths are then offered to the routing table (also called the forwarding database).
    - Packets arriving at the router are forwarded based on the information held in the routing table.

Link-State Protocol Operation

- Link-state routing protocols generate routing updates only when a change occurs in the network topology.
- When a link changes state, the device that detected the change creates a Link-State Advertisement (LSA) concerning that link.
  - LSA propagates to neighbor devices using a special multicast address.
- Each router stores the LSA, forwards the LSA to neighboring devices and updates its Link-State DataBase (LSDB).
- Link-state routers find the best paths to a destination by applying Dijkstra's algorithm, also known as SPF, against the LSDB to build the SPF tree.
- Each router selects the best paths from their SPF tree and places them in their routing table.



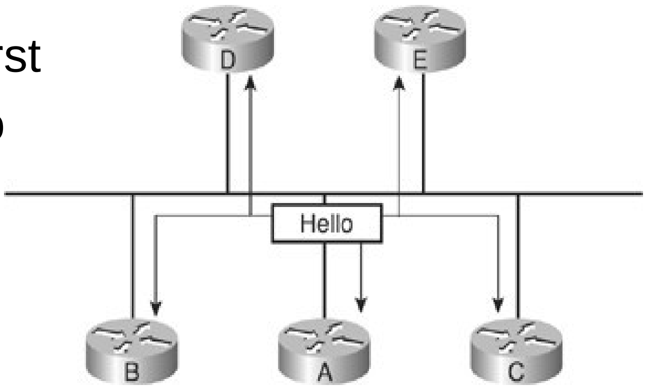
Link-State Advertisement (LSA)

- LSAs report the state of routers and the links between routers.
- Link-state information must be synchronized between routers.
- LSAs have the following characteristics:
  - LSAs are reliable. There is a method for acknowledging their delivery.
  - LSAs are flooded throughout the area (or throughout the domain if there is only one area).
  - LSAs have a sequence number and a set lifetime, so each router recognizes that it has the most current version of the LSA.
  - LSAs are periodically refreshed to confirm topology information before they age out of the LSDB.

OSPF Router ID (RID)

- The Router ID identifies the router and is:
  - The highest IPv4 address of all router interfaces at the moment of the OSPF process activation.
  - A value administratively defined.
- If a physical interface address is being used as the router ID, and that physical interface fails, and the router (or OSPF process) is restarted, the router ID will change.
  - This change in router ID makes it more difficult for network administrators to troubleshoot and manage OSPF.
- Administratively defining the RID or using loopback interfaces for the router ID forces the router ID to stay the same, regardless of the state of the physical interfaces.

OSPF Adjacencies



- A router running a link-state routing protocol must first establish neighbor adjacencies, by exchanging hello packets with the neighboring routers
- The router sends and receives Hello packets to and from its neighboring routers.
    - The destination address is typically a multicast address.
    - It is possible to define unicast OSPF relations.
- The routers exchange hello packets subject to protocol-specific parameters, such as checking whether the neighbor is in the same area, using the same hello interval, and so on.
    - Routers declare the neighbor up when the exchange is complete.
- Two OSPF routers on a point-to-point serial link, usually encapsulated in High-Level Data Link Control (HDLC) or Point-to-Point Protocol (PPP), form a full adjacency with each other.
- However, OSPF routers on broadcast networks, such as LAN links, 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 pass LSAs only to them.

DR and BDR Election

- The first OSPF router to boot becomes the Designated Router (DR).
- The second router to boot becomes the Backup Designated Router (BDR).
- If multiple routers boot simultaneously,
  - The DR it will be the router with the highest priority. The BDR the second.
    - The OSPF priority is a administratively defined parameter.
  - In case of tie, it will be chosen the router with the highest Router ID (RID).
- When the DR fails, the BDR assumes the role of DR.
  - The BDR does not perform any DR functions when the DR is operating.
  - The choice of the new BDR is done according to some criteria of the initial election.
- After the election, the DR and BDR maintain that role, independently of which routers join the OSPF process.
- The ID of an OSPF Network is the IP address of the network's Designated Router (DR) interface.

OSPF LS Database

- The OSPF database (LSDB) is organized in two tables.
  - Router Link States – Routers related information table.
    - The routers are identified by theirs RID.
  - Net Link States – Networks/Links related information table.
    - Networks are identified by their 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 LS Database Tables (1)

- Router Link States
  - For each router, it contains the information about the networks directly connected to that router.

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

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

  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
```

OSPF LS Database Tables (2)

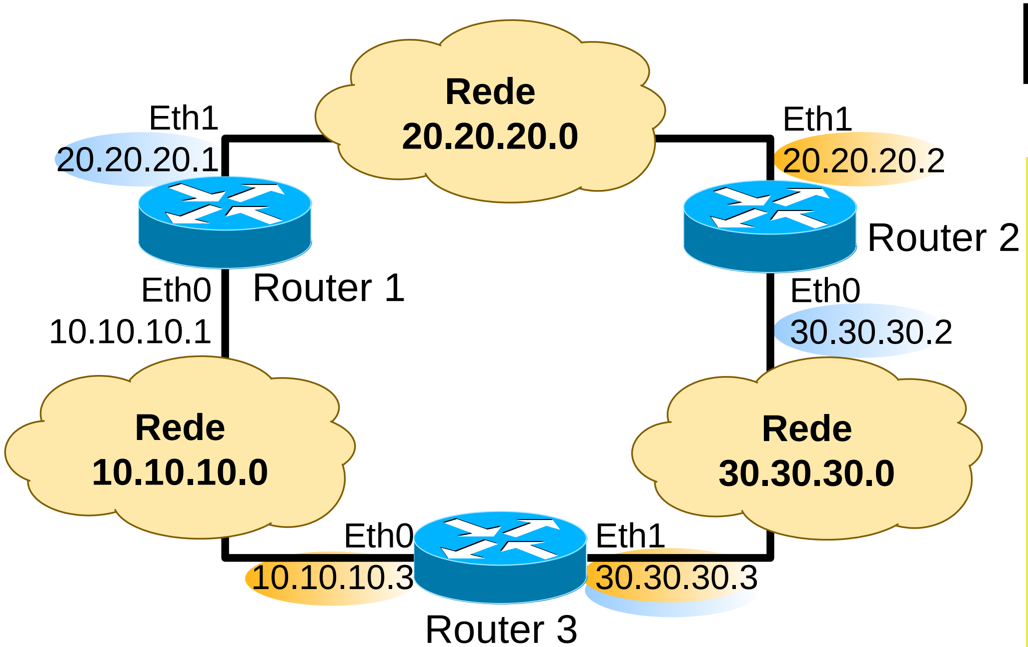
- Network Link States
  - For each network, it contains the information about the routers directly attached to that network.

```
Routing Bit Set on this LSA
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
```

Network ID

Attached routers (RID)

OSPF LSDatabase Example




```
 Routing Bit Set on this LSA
  LS age: 208
  Options: (No TOS-capability, DC)
  LS Type: Network Links
  Link State ID: 20.20.20.2 (address of Designated
Router)
  Advertising Router: 30.30.30.2
  LS Seq Number: 80000001
  Checksum: 0xA164
  Length: 32
  Network Mask: /24
Attached Router: 30.30.30.2
Attached Router: 20.20.20.1
```

Network 20.20.20.0's Network Link State

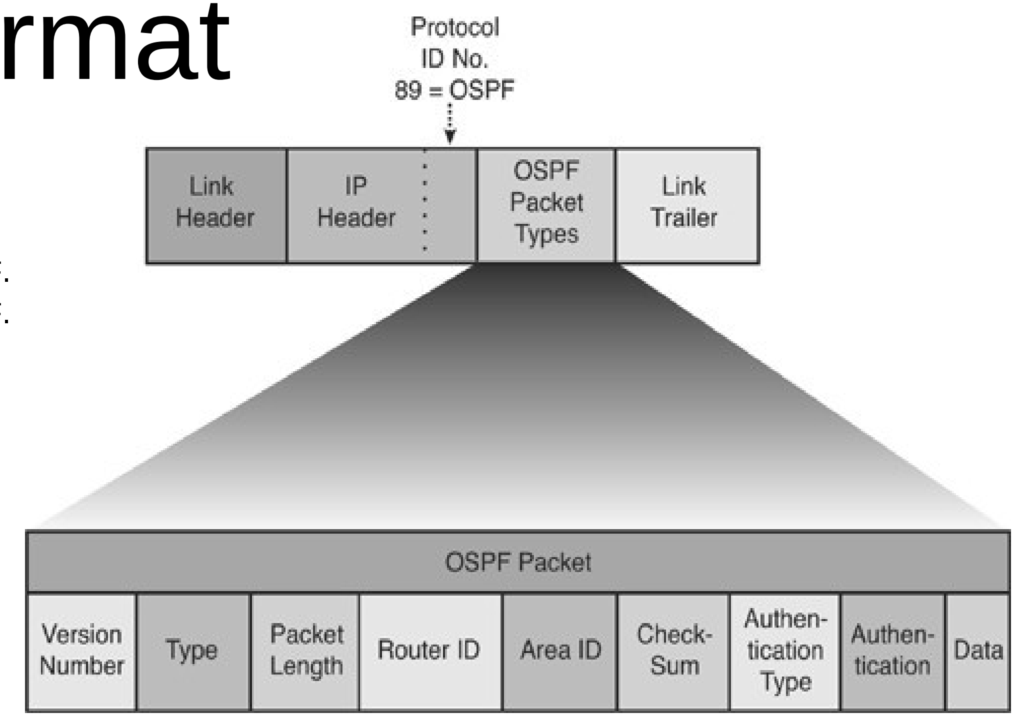
```
 LS age: 321
  Options: (No TOS-capability, DC)
  LS Type: Router Links
  Link State ID: 20.20.20.1
  Advertising Router: 20.20.20.1
  LS Seq Number: 8000000A
  Checksum: 0xE7FB
  Length: 48
  Number of Links: 2

    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
```

Router 1's Router Link State

OSPF Packets

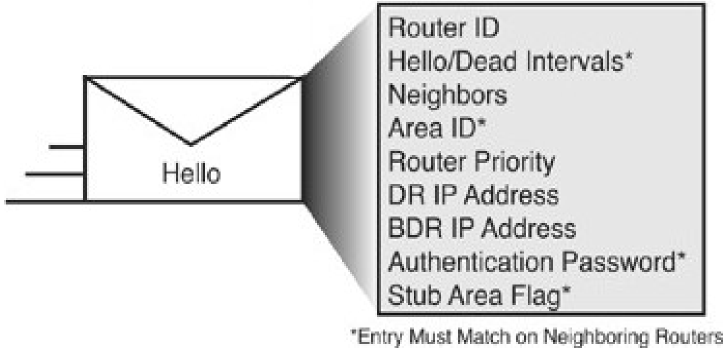
- Hello - Discovers neighbors and builds adjacencies between them.
- Database Description (DBD) - Checks for database synchronization between routers.
- Link-State Request (LSR) - Requests specific link-state records from another router.
- Link-State Update (LSU) - Sends specifically requested link-state records.
- LSAck - Acknowledges the other packet types.

OSPF Packet Format



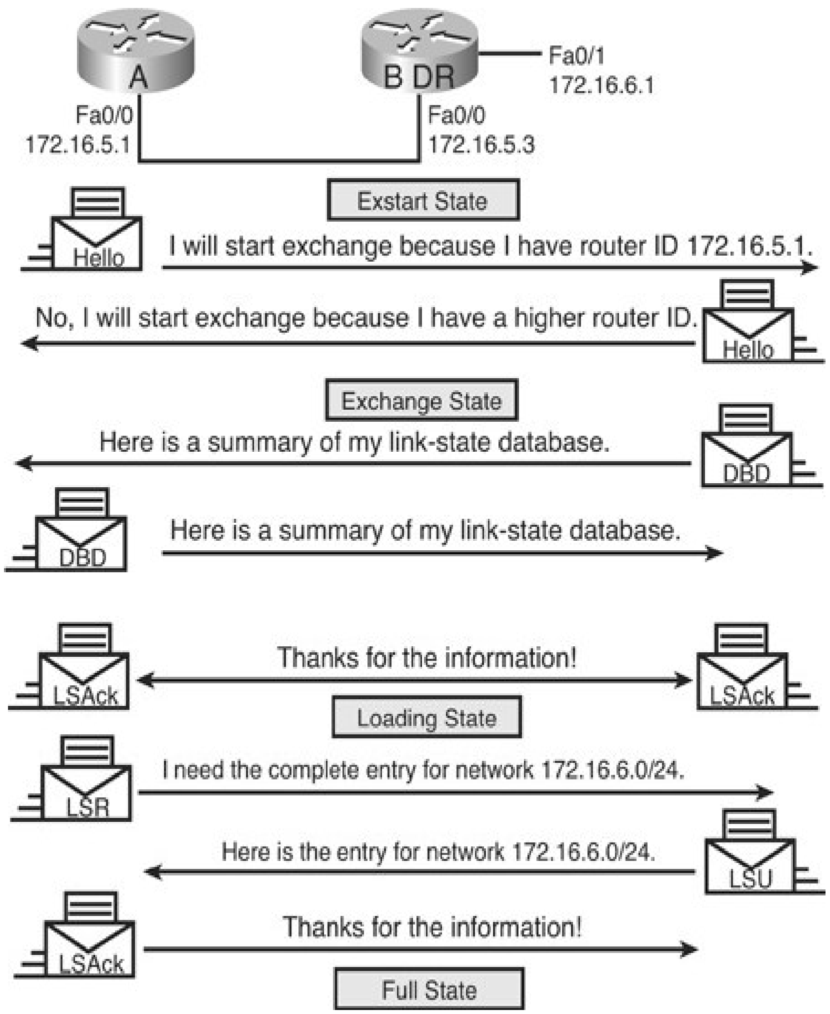
- Version Number
  - Set to 2 for OSPF Version 2, the IPv4 version of OSPF.
  - Set to 3 for OSPF Version 3, the IPv6 version of OSPF.
- Type
  - Differentiates the five OSPF packet types.
- Packet Length
  - The length of the OSPF packet in bytes.
- Router ID
  - Defines which router is the packet's source.
- Area ID
  - Defines the area in which the packet originated.
- Checksum
  - Used for packet header error detection to ensure that the OSPF packet was not corrupted during transmission.
- Authentication Type
  - An option in OSPF that describes either no authentication, clear-text passwords, or encrypted message digest 5 (MD5) for router authentication.
- Authentication
  - Used with authentication type.
- Data, contains different information, depending on the OSPF packet type:
  - For the Hello packet - Contains a list of known neighbors.
  - For the DBD packet - Contains a summary of the LSDB, which includes all known router Ids and their last sequence number, among several other fields.
  - For the LSR packet - Contains the type of LSU needed and the router ID of the router that has the needed LSU.
  - For the LSU packet - Contains the full LSA entries. Multiple LSA entries can fit in one OSPF update packet.
  - For the LSAck packet - This data field is empty.

OSPF Hello Packets



- An hello packet contains the following information:
  - Router ID
    - A 32-bit number that uniquely identifies the router.
  - Hello and dead intervals
    - The hello interval specifies how often, in seconds, a router sends hello packets (10 seconds is the default on multiaccess networks).
    - The dead interval is the amount of time in seconds that a router waits to hear from a neighbor before declaring the neighbor router out of service (the dead interval is four times the hello interval by default).
    - These timers must be the same on neighboring routers; otherwise an adjacency will not be established.
  - Neighbors
    - The Neighbors field lists the adjacent routers with which this router has established bidirectional communication.
    - Bidirectional communication is indicated when the router sees itself listed in the Neighbors field of the hello packet from the neighbor.
  - Area ID
    - To communicate, two routers must share a common segment, and their interfaces must belong to the same OSPF area on that segment.
    - These routers will all have the same link-state information for that area.
  - Router priority
    - An 8-bit number that indicates a router's priority. Priority is used when electing a DR and BDR.
  - DR and BDR IP addresses
    - If known, the IP addresses of the DR and BDR for the specific multiaccess network.
  - Authentication password
    - If router authentication is enabled, two routers must exchange the same password.
    - Authentication is not required, but if it is enabled, all peer routers must have the same password.
  - Stub area flag
    - A stub area is a special area.
    - The stub area technique reduces routing updates by replacing them with a default route.
    - Two neighboring routers must agree on the stub area flag in the hello packets.
- Hello Interval, Dead Interval, Area ID, Authentication Password and Stub Area Flag fields must match on neighboring routers for them to establish an adjacency.

Discovering the Network Routes



- A master and slave relationship is created between each router and its adjacent DR and BDR.
  - Only the DR exchanges and synchronizes link-state information with the routers to which it has established adjacencies.
- The master and slave routers exchange one or more DBD packets.
  - A DBD includes information about the LSA entry header that appears in the router's LSDB.
  - The entries can be about a link or about a network.
  - Each LSA entry header includes information about the link-state type, the address of the advertising router, the link's cost, and the sequence number.
  - The router uses the sequence number to determine the "newness" of the received link-state information.
- It acknowledges the receipt of the DBD using the LSAck packet.
  - It compares the information it received with the information it has in its own LSDB.
- If the DBD has a more current link-state entry, the router sends an LSR to the other router.
- The other router responds with the complete information about the requested entry in an LSU packet.
- Again, when the router receives an LSU, it sends an LSAck.
- The router adds the new link-state entries to its LSDB.

OSPF Example

OSPF activated on Router 1

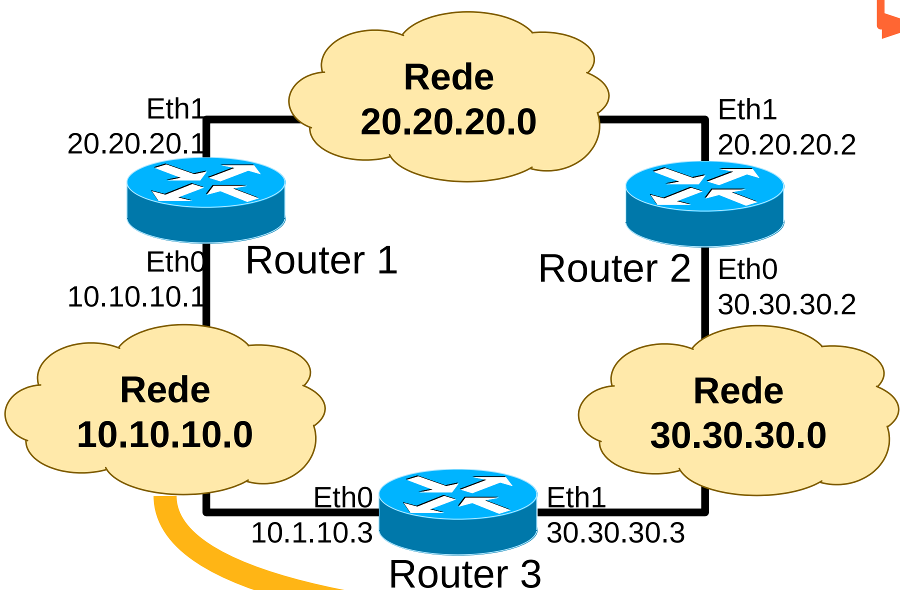
OSPF activated on Router 3

OSPF activated on Router 2

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

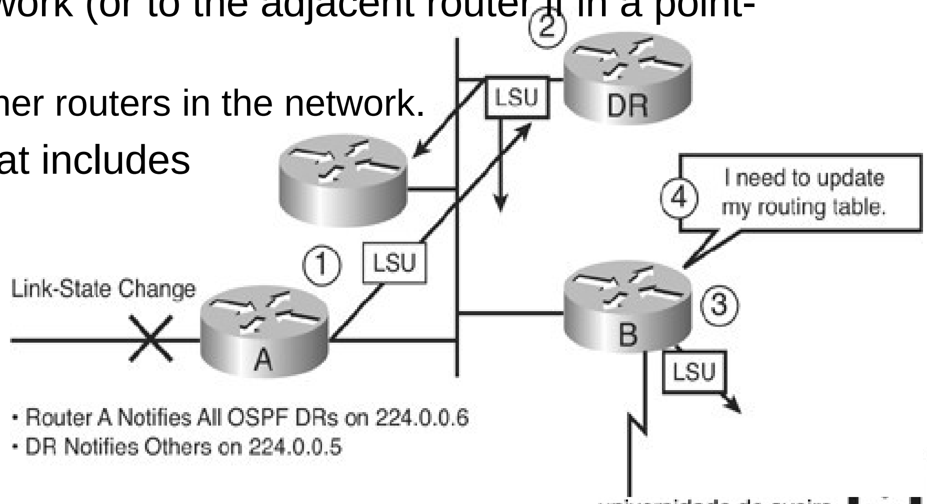
80.000000	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.683033	10.10.10.3	224.0.0.5	OSPF	LS Update
83.715683	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.717864	10.10.10.1	10.10.10.3	OSPF	Hello Packet
83.726166	10.10.10.3	10.10.10.1	OSPF	DB Descr.
83.726258	10.10.10.3	10.10.10.1	OSPF	Hello Packet
83.728433	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

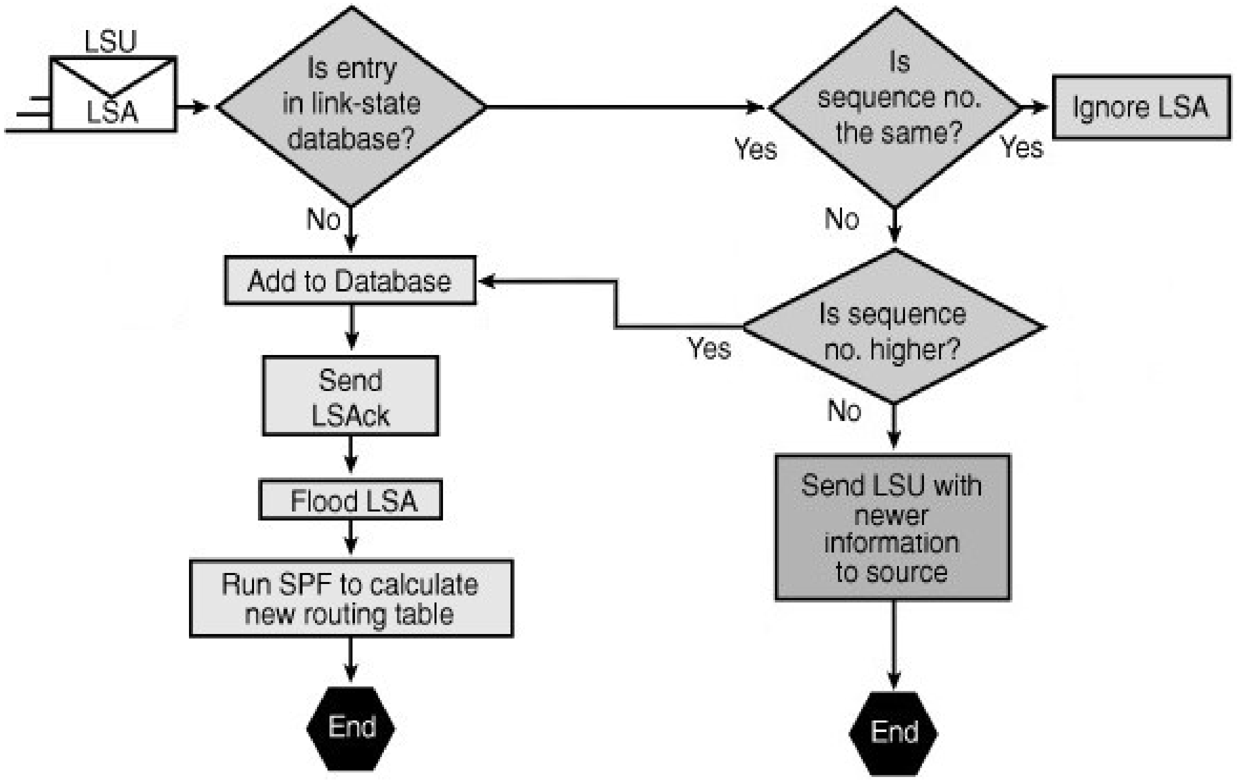


Maintaining Routing Information

- Flooding process:
  - A router notices a change in a link state and multicasts an LSU packet, which includes the updated LSA entry with the sequence number incremented, to 224.0.0.6.
    - This address goes to all OSPF DRs and BDRs.
    - On point-to-point links, the LSU is multicast to 224.0.0.5.)
    - An LSU packet might contain several distinct LSAs.
  - The DR receives the LSU, processes it, acknowledges the receipt of the change and floods the LSU to other routers on the network using the OSPF multicast address 224.0.0.5.
    - After receiving the LSU, each router responds to the DR with an LSAck.
    - To make the flooding procedure reliable, each LSA must be acknowledged separately.
  - If a router is connected to other networks, it floods the LSU to those other networks by forwarding the LSU to the DR of the other network (or to the adjacent router if in a point-to-point network).
    - That DR, in turn, multicasts the LSU to the other routers in the network.
  - The router updates its LSDB using the LSU that includes the changed LSA.
  - It then recomputes the SPF algorithm against the updated database after a short delay and updates the routing table as necessary.



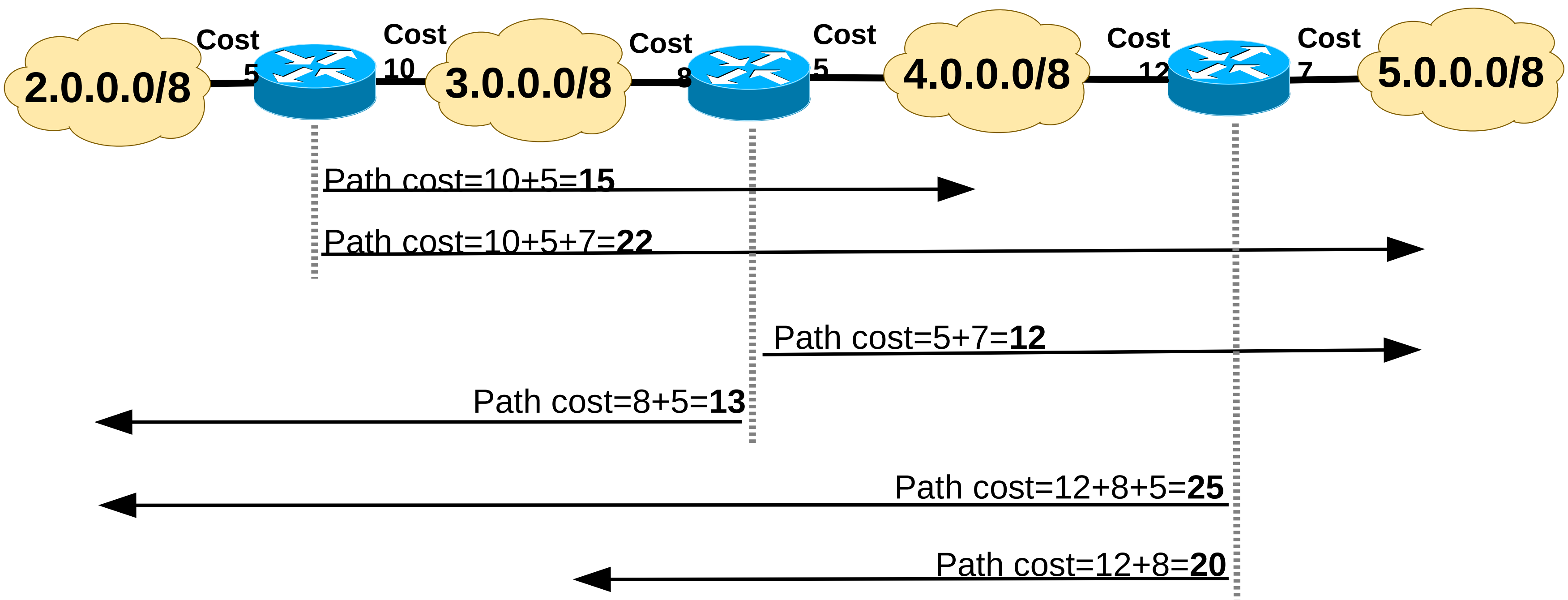
LSA Operation



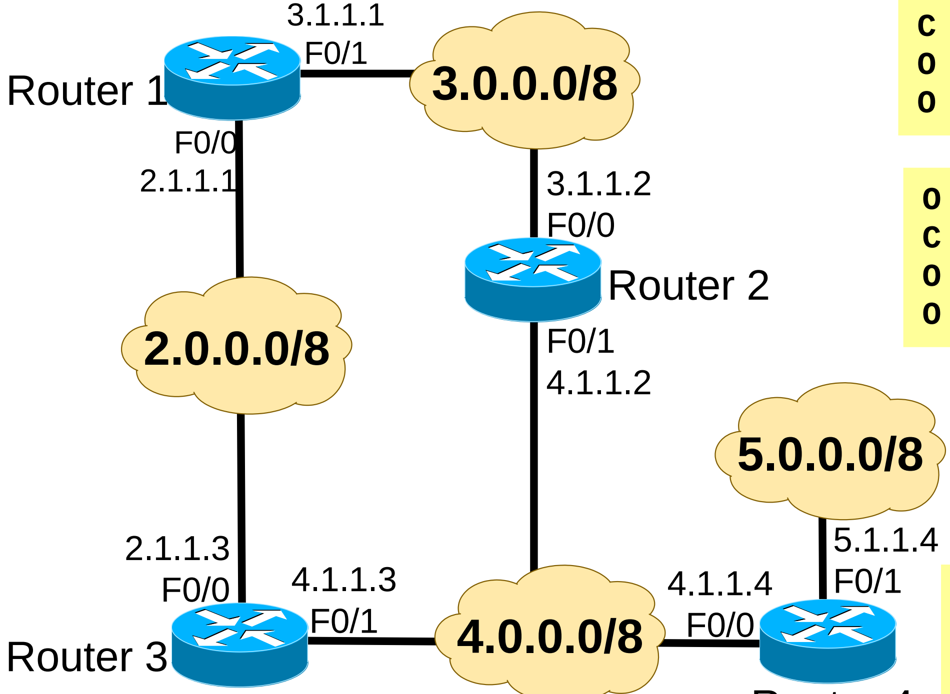
- When each router receives the LSU:
  - If the LSA entry does not already exist, the router adds the entry to its LSDB, sends back a link-state acknowledgment (LSAck), floods the information to other routers, runs SPF, and updates its routing table.
  - If the entry already exists and the received LSA has the same sequence number, the router ignores the LSA entry.
  - If the entry already exists but the LSA includes newer information (it has a higher sequence number), the router adds the entry to its LSDB, sends back an LSAck, floods the information to other routers, runs SPF, and updates its routing table.
  - If the entry already exists but the LSA includes older information, it sends an LSU to the sender with its newer information.

OSPF Path Costs

- Each router link/interface has an associated OSPF cost.
- The total cost between a router and a network is given by the sum of all OSPF costs of the (routers) output interfaces along the path.
  - Routers to access directly connect networks never use OSPF paths.



OSPF Example



All interfaces cost = 10

```
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
```

```
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
```

Router 1 and Router 2 after disconnecting the F0/1 at Router2

```
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
```

Router1, now with the cost of Router2 F0/1 interface equal to 5

```
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
```

Router 1 initial table

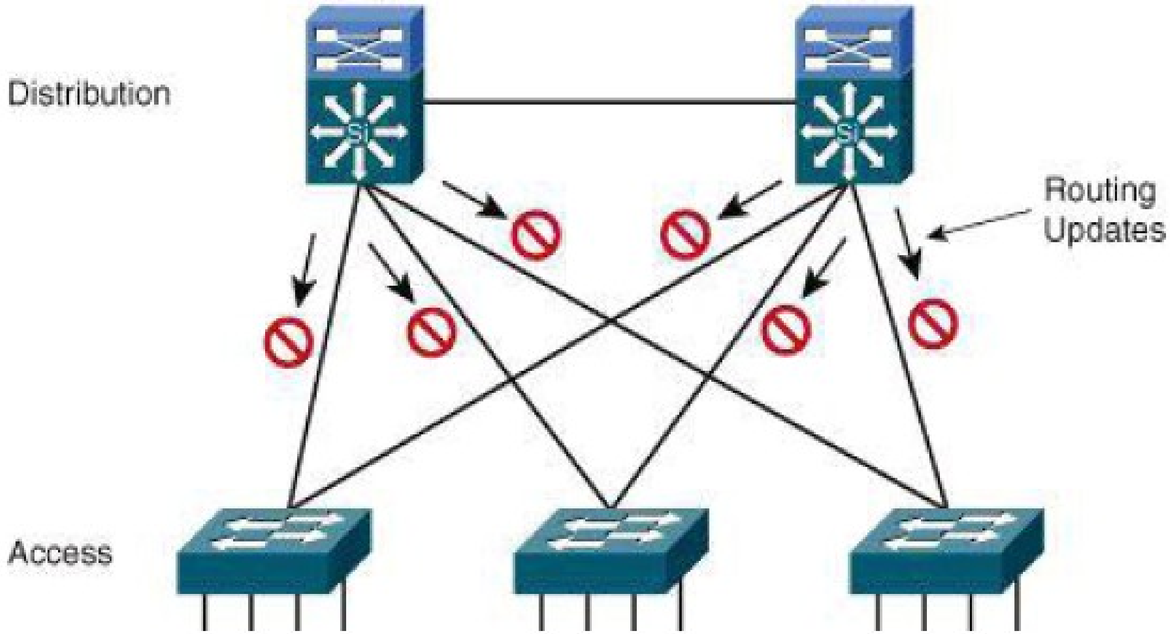
IPv6 Routing - OSPFv3

- 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
    - Neighbors 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 neighbors of link local address
  - Informs neighbors 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

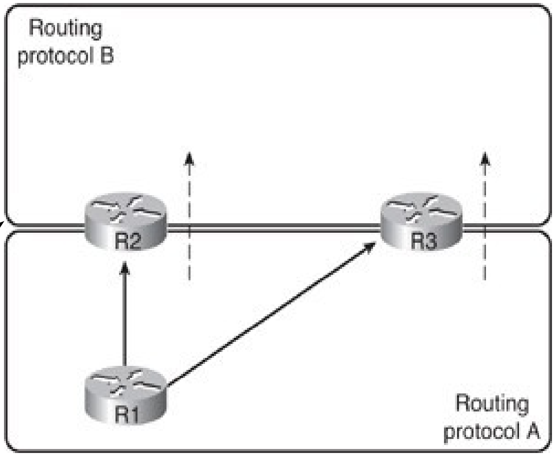
Passive Interfaces on Access Layer

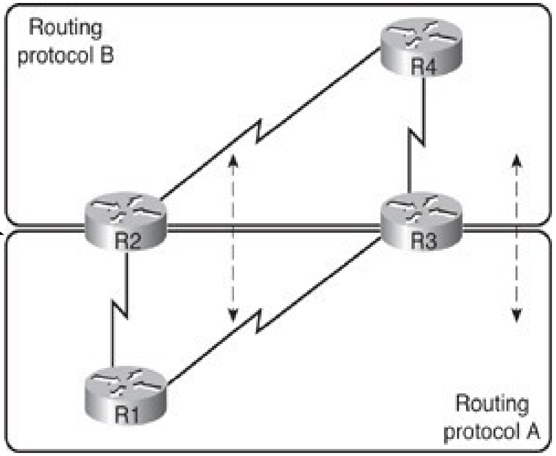


- As a recommended practice, limit unnecessary L3 routing peer adjacencies by configuring the ports toward Layer 2 access switches as passive.
  - Suppress the advertising of routing updates.
  - If a distribution switch does not receive L3 routing updates from a potential peer on a specific interface, it does not form a neighbor adjacency with the potential peer across that interface.

Route Redistribution

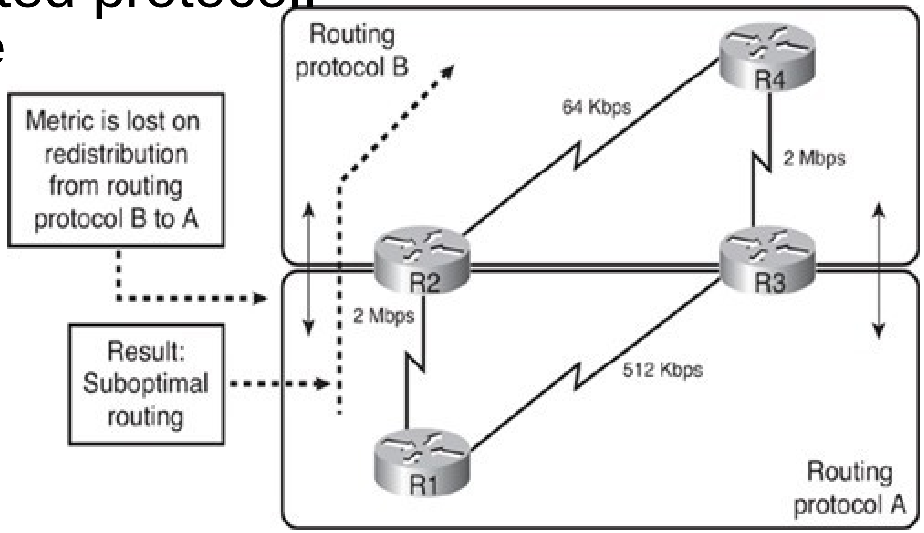
- Domains with difference routing protocols can exchange routes.
  - This is called route redistribution.
    - One-way redistribution - Redistributes only the networks learned from one routing protocol into the other routing protocol.
      - – Uses a default or static route so that devices in that 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.





Redistribution Issues

- Lost metric from redistributed protocol.
  - It is not possible to achieve an optimal overall routing.



- Preventing Routing Loops in a Redistribution Environment.
  - Safest way to perform redistribution is to redistribute routes in only one direction, on only one boundary router within the network.
    - However, that this results in a single point of failure in the network.
  - If redistribution must be done in both directions or on multiple boundary routers, the redistribution should be tuned to avoid problems such as suboptimal routing and routing loops.

Redistribution Techniques

- Redistribute a default route from the core autonomous system into the edge autonomous system, and redistribute routes from the edge routing protocols into the core routing protocol.
  - This technique helps prevent route feedback, suboptimal routing, and routing loops.
- Redistribute multiple static routes about the core autonomous system networks into the edge autonomous system, and redistribute routes from the edge routing protocols into the core routing protocol.
  - This method works if there is only one redistribution point; multiple redistribution points might cause route feedback.
- Redistribute routes from the core autonomous system into the edge autonomous system with filtering to block out inappropriate routes.
  - For example, when there are multiple boundary routers, routes redistributed from the edge autonomous system at one boundary router should not be redistributed back into the edge autonomous system from the core at another redistribution point.
- Redistribute all routes from the core autonomous system into the edge autonomous system, and from the edge autonomous system into the core autonomous system, and then modify the administrative distance associated with redistributed routes so that they are not the selected routes when multiple routes exist for the same destination.