الجـــــمـــهـــوريـــــــة الجـــــــزائـــريـــة الديمـــقـــراطـيـــة الـــشـــعـــبـيـــة
People's Democratic Republic of Algeria

وزارة التـــــعـــــليـــــم العـــــالي والبـــــحث العـــــلـــــمـــــي
Ministry of Higher Education and Scientific Research

المركز الجـــــامعي نـــــور البشـــــــــير – البيّـــــــــــض –
University Centre Nour El Bachir - El Bayadh Institute of Science and Technology Department

IP Routing

Course for first year Master Telecommunication students

Prepared by:

Dr. Mehdi ROUISSAT

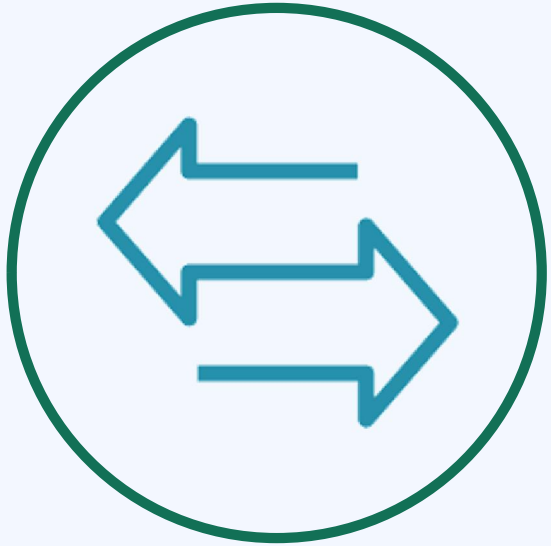
preface

This course is intended for students of first year Master Telecommunication. At the end of this course, the student will know the functions that allow to determine the best path in a mesh network to a destination identified by an IP network address. In this subject, we find static and dynamic routing. The course is organized into six chapters :

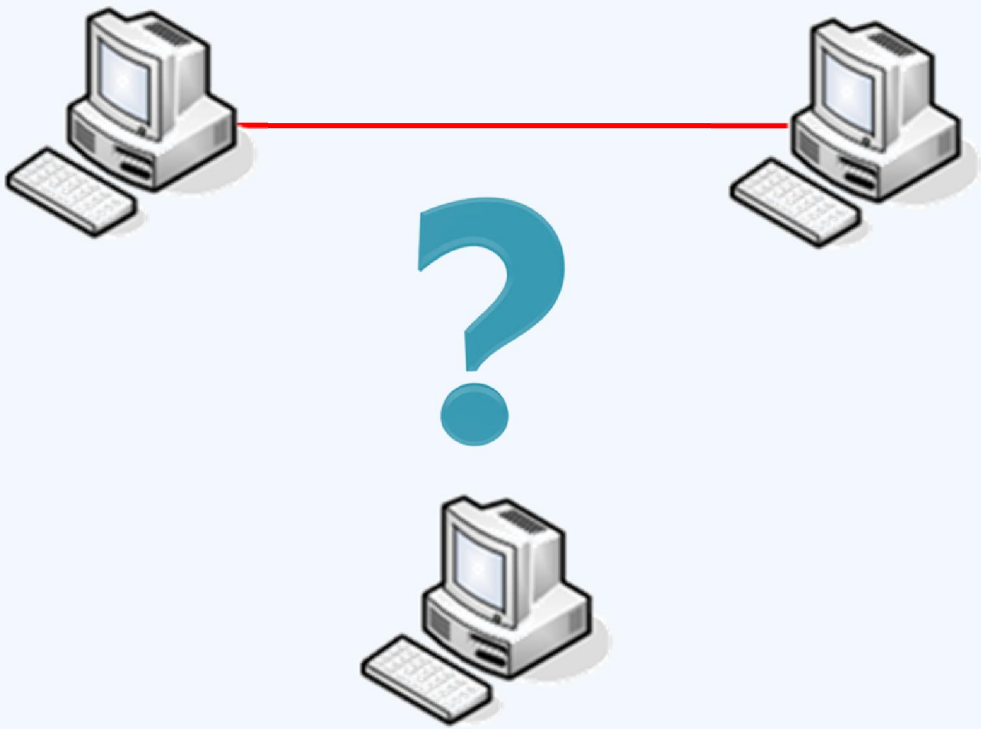
- Chapter 1. Switching in LANs
- Chapter 2. VLANS
- Chapter 3. Redundancies on switched links
- Chapter 4. Link Aggregation (Etherchannel)
- Chapter 5. Static Routing
- Chapter 6. Dynamic Routing

Chapter 01

Switching in LANs

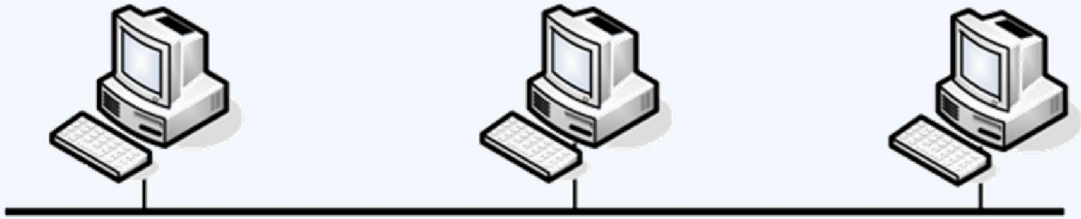


Local Area Network Interconnection

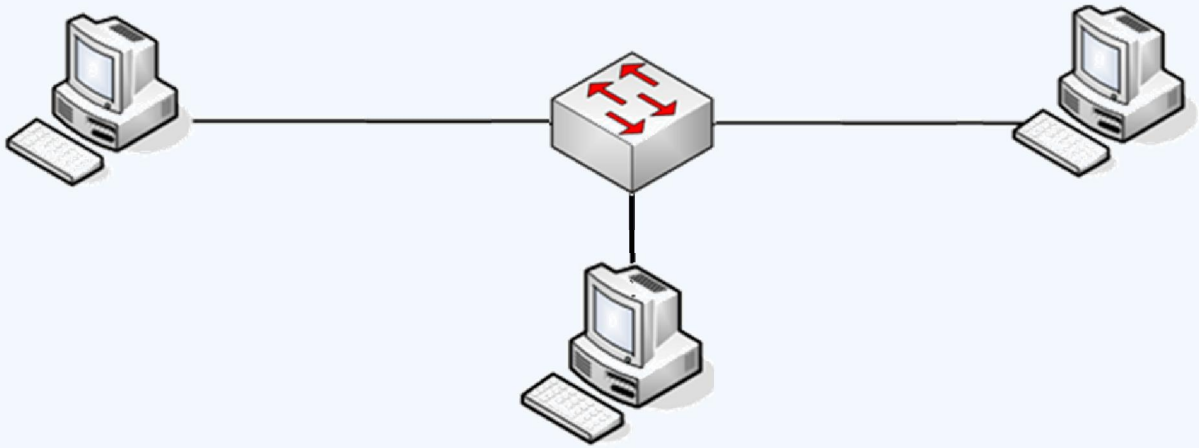


Local Area Network Interconnection (2)

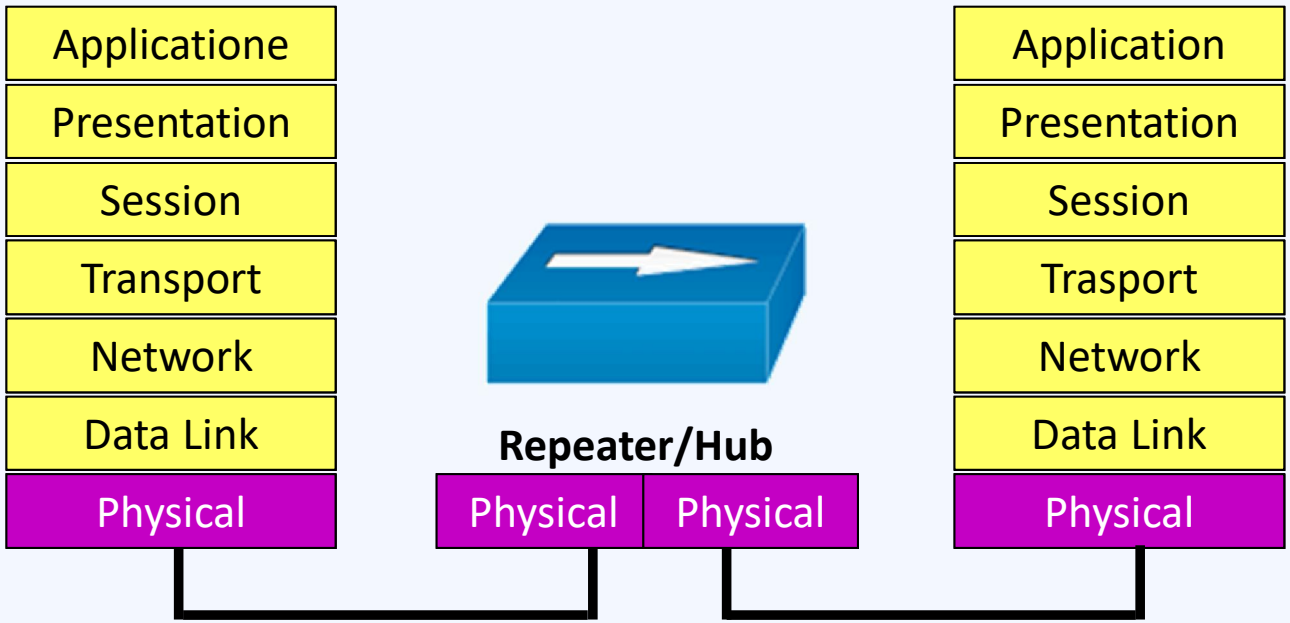
* Shared support:



* Interconnection equipment:



The Hub -1-

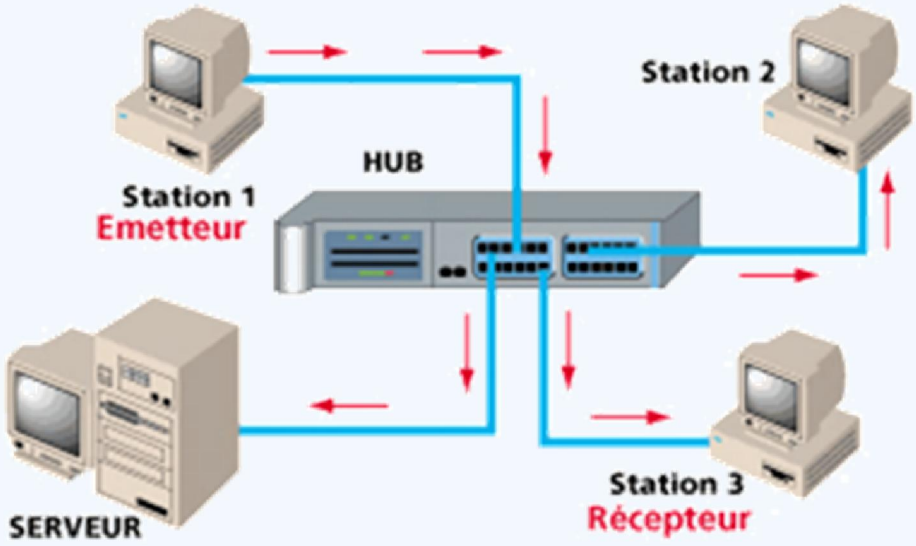


Situation of hubs in the OSI model

The Hub (The concentrator) -2-

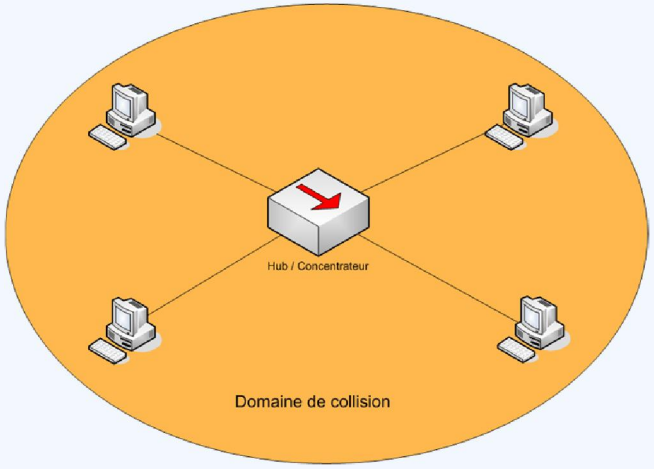
It has multiple interfaces, it receives an incoming signal from one interface and repeats it on all other interfaces.

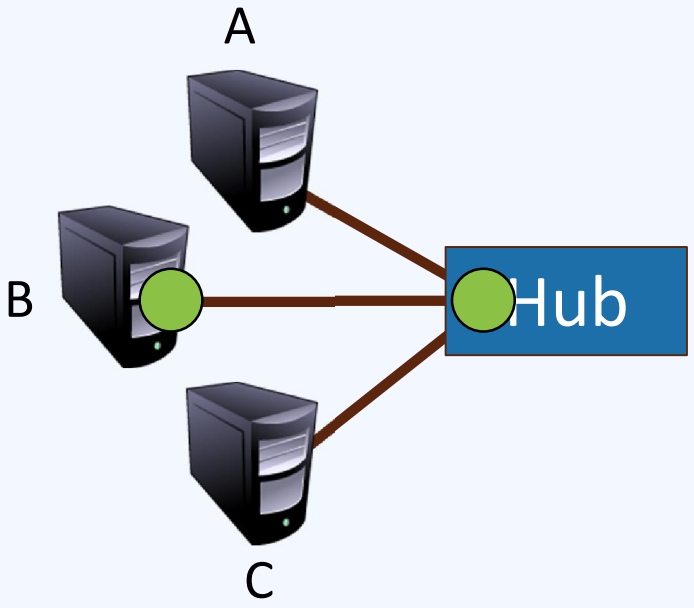
- Only works on signals,
- It has an extremely low cost.
- no need to "configure" the Hub, it's plug&play



The Hub (The concentrator) -3-

On a Hub, all the stations can potentially transmit at the same time and cause a collision, so they belong to the same collision domain.

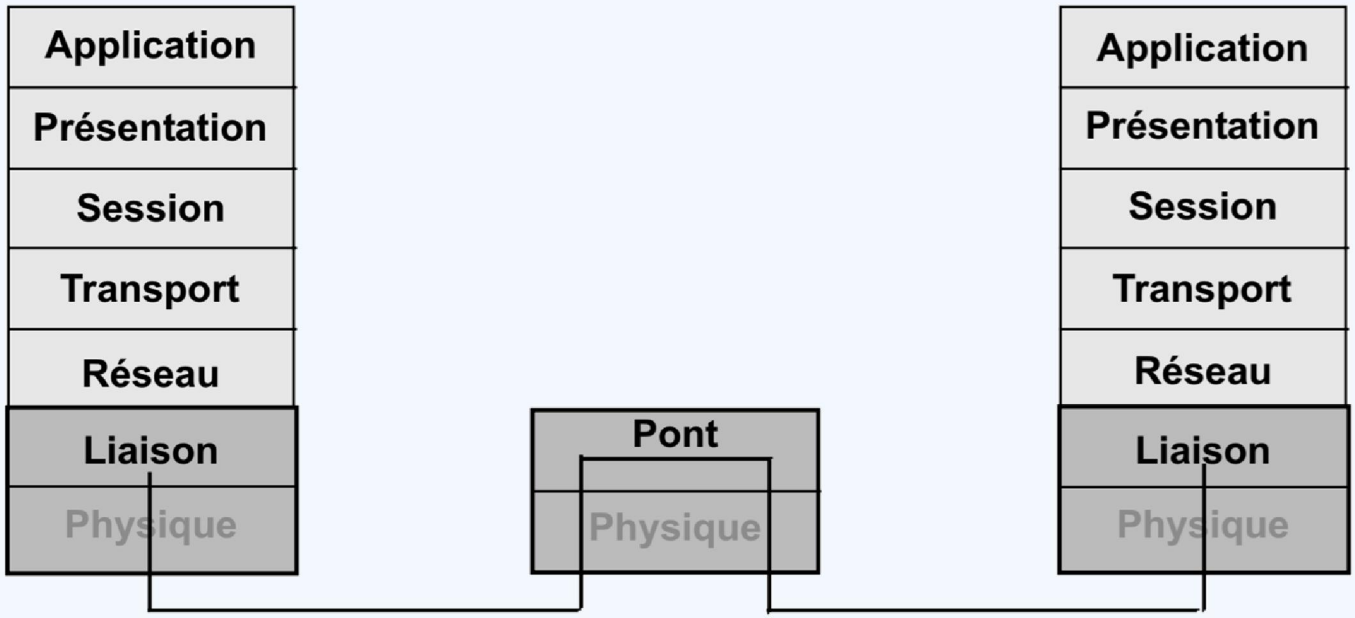




The Hub in action

Hubs operate in Half-Duplex. This mode allows electrical impulses to be sent and received on the same copper pair.
When you send data to the Hub (or vice versa), you can't receive data otherwise there is a collision!

The Bridge -1-



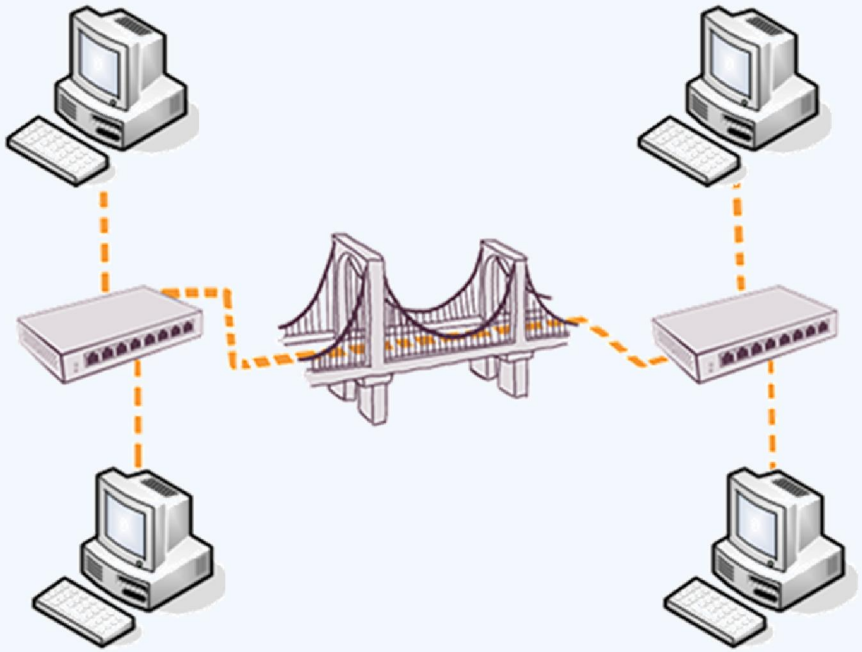
Situation of bridges in the OSI model

The Bridge -2-

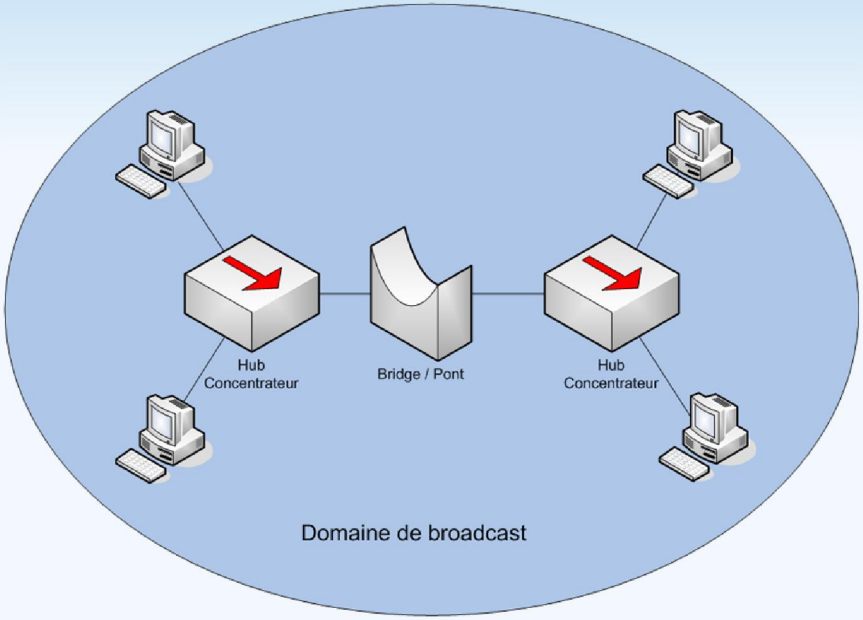
- It is used in **layer 2** of the OSI model (link).
- Its objective is to interconnect two distinct network segments,
- It is able to filter **frames** by allowing only those whose **address** corresponds to a machine located opposite the bridge to pass through.

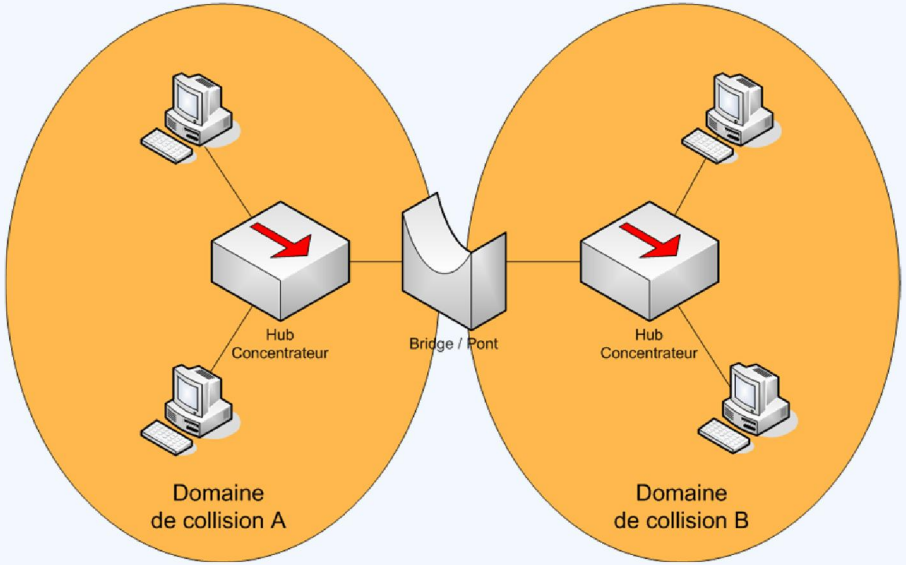
The Bridge -3-

Bridges are used to logically separate segments of the same network. They are independent of upper-layer protocols.



The Bridge - 4 -



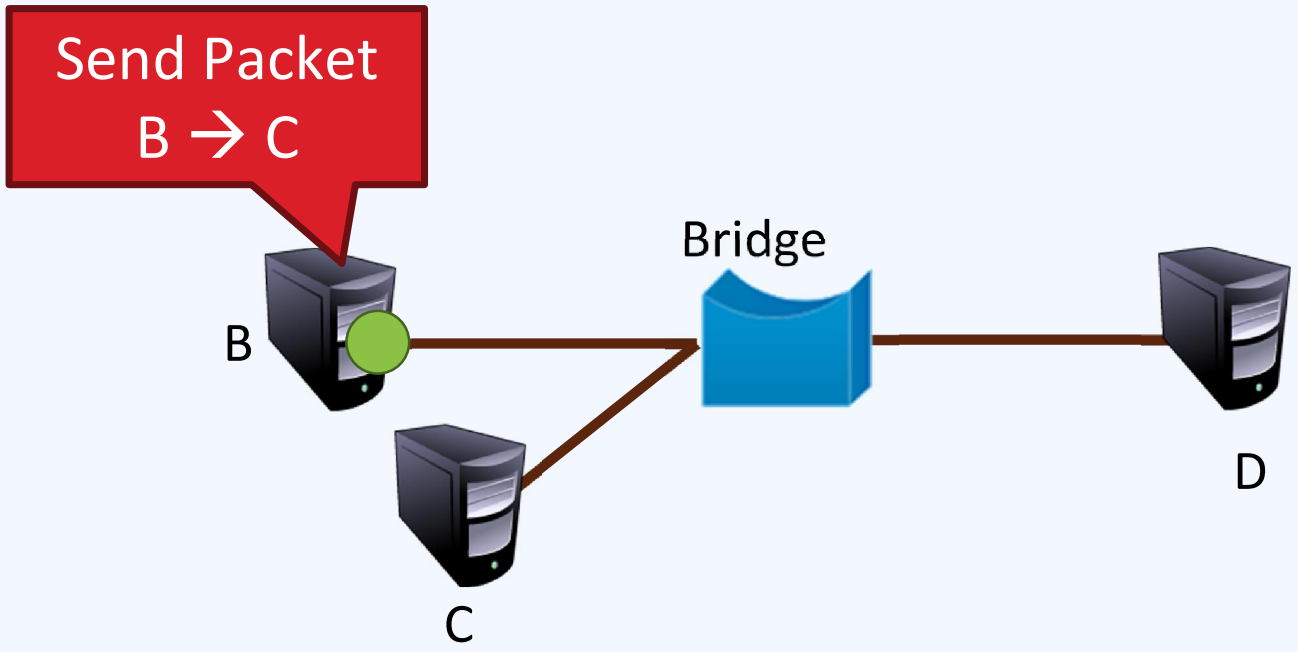


Diffusion

Collision

The Bridge - 5 -

- Transmit only frames to intended recipients
- No broadcast!
- The bridge is said to be transparent



Today, there are almost no Bridges anymore because they are replaced by Switches, which are more powerful with a higher density of ports

The Switch

A switch is a **multiport bridge**, allowing to connect several segments of the network and route **frames** to the right destination segment, based on **MAC addresses**.

- is a data link layer device.
- can perform error checking before transmitting data,
- Splits the hosts' collision domain, but the broadcast domain remains the same.

The Switch

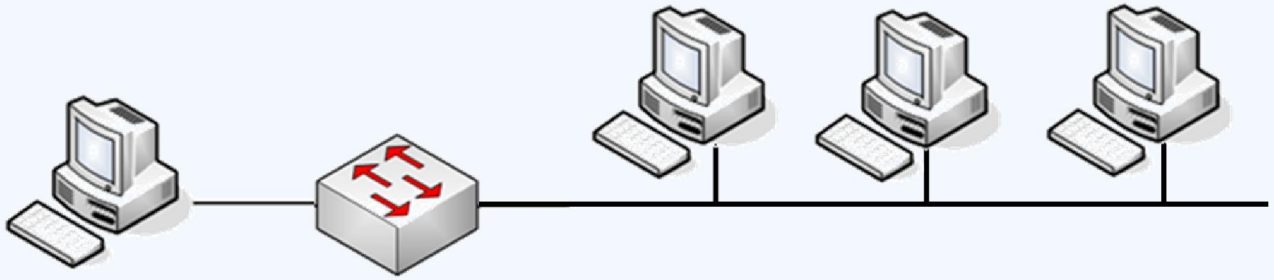
- Is physically similar to a hub,
- Is logically similar to a bridge,
- Works on frames,
- Includes MAC addresses,
- Transparent
- They can be used in busy networks to isolate the flow of data and improve performance,
- Switches are the Ethernet device of choice for Layer 2...

Types of switches

There are two types:

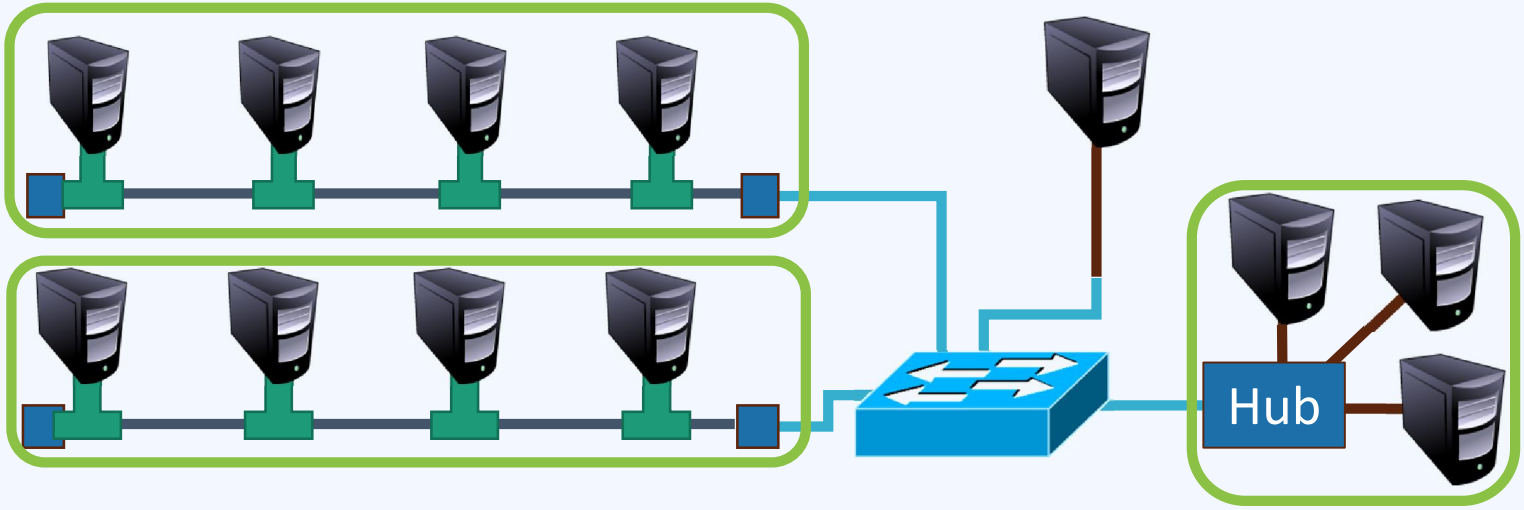
- Switches per port: Only one station is connected to each port.
- Switches per segment: A local subnet is connected to a port on the switch.

They allow a single station or a group of stations to be addressed to a port on the switch, respectively.



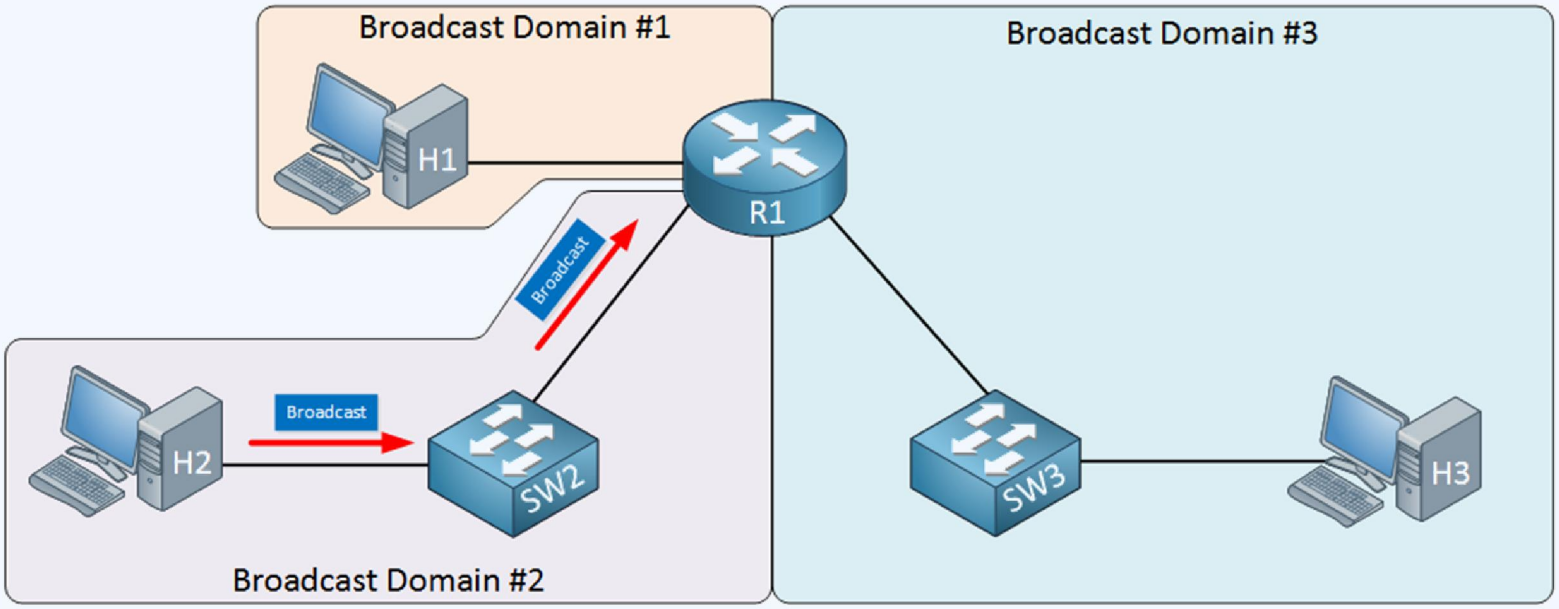
The Switch

- The Switch limits the size of collision domains
- the link between the switch and a computer is done in full duplex (CSMA/CD is not used)
- Increase the level of privacy
- Significantly improves scalability
- Need for buffer and switch tables

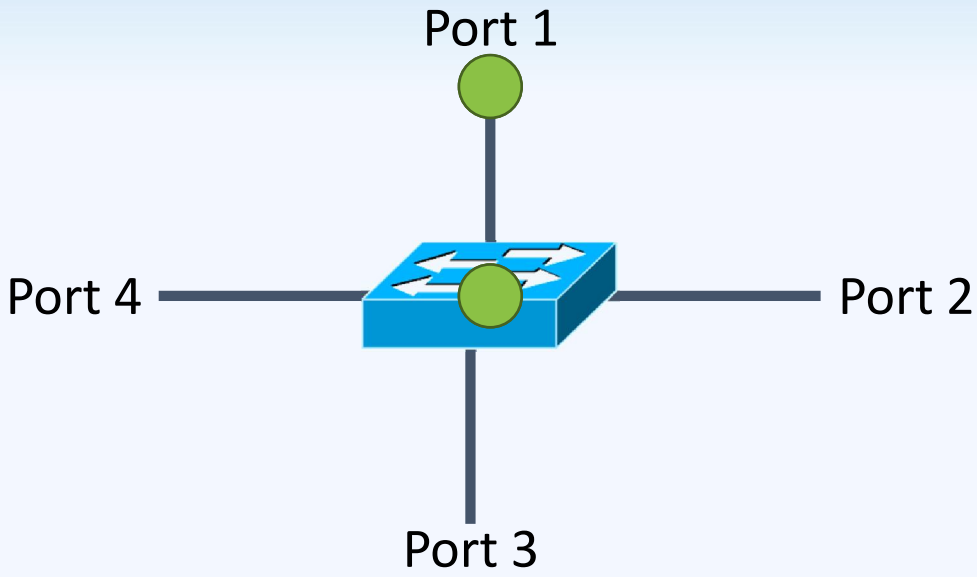


The switch

Important: the broadcast domain stops at a level 3 device, such as a router!



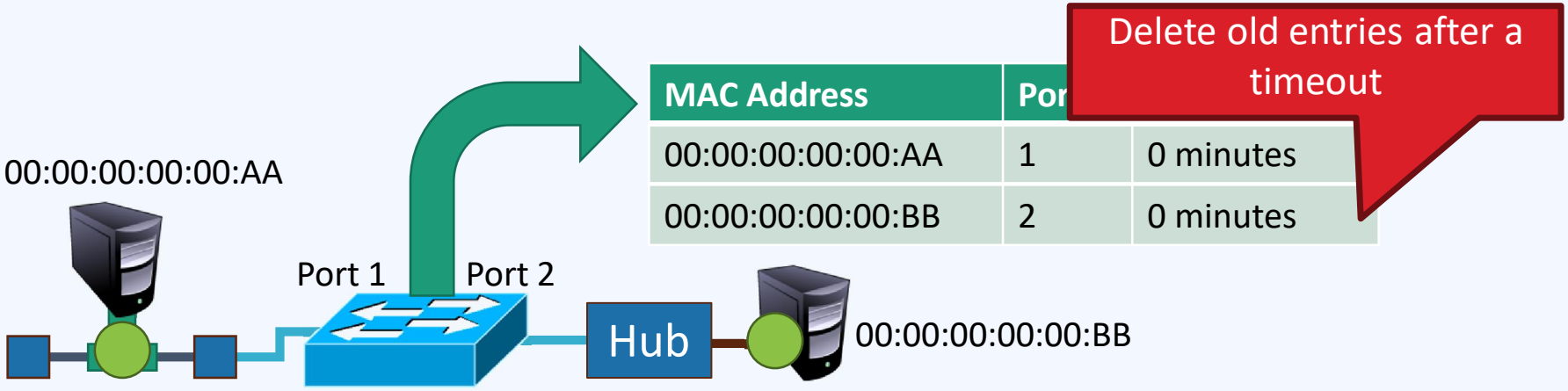
Frame Transfer in action



- Assuming that the frame arrives at port1
- If the destination MAC address is at the switch table, then it will be sent to the correct port
- If the destination MAC address does not exist on the switch table, then the frame will be broadcast to all ports except 1

MAC Address learning

- Manual configuration is possible, but...
  - ✓Long
  - ✓Error-prone
  - ✓Non-adaptable (hosts can be added or removed)
- Instead, learn addresses using a simple heuristic
  - ✓Look at the source of the frames that arrive at each port



Frame Transfer Tables

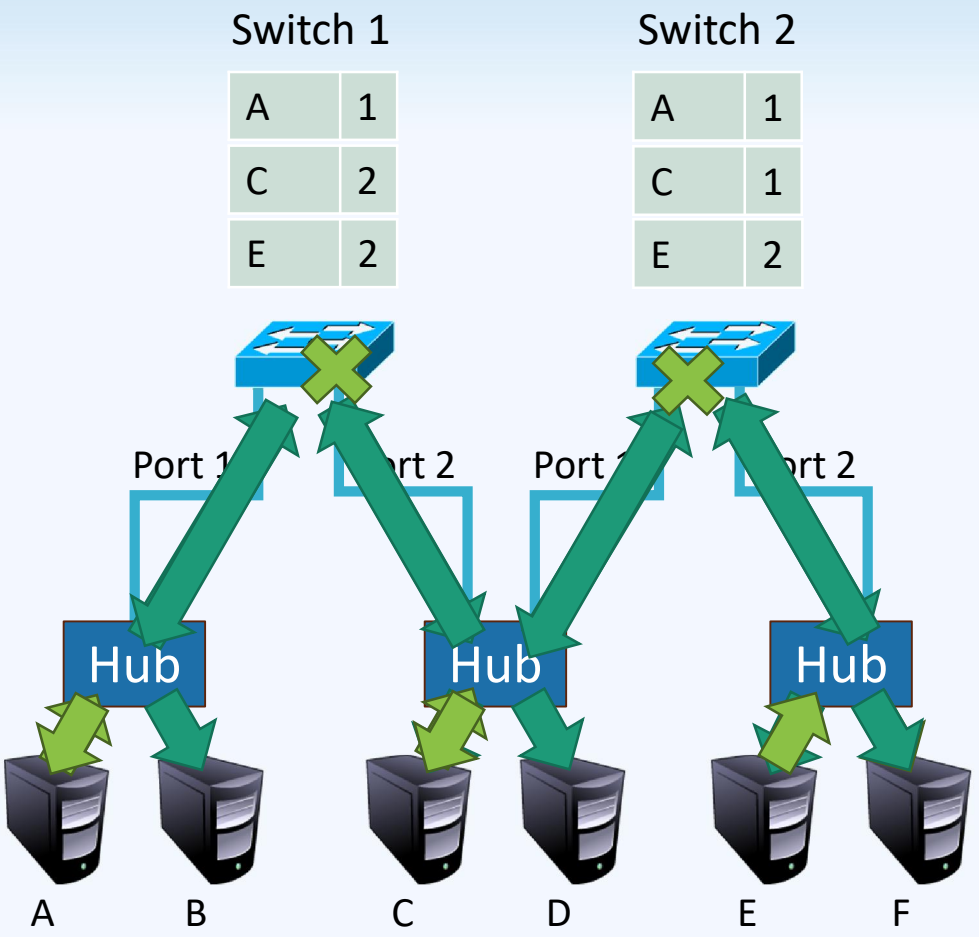
- Each switch maintains a Forwarding Data Base (FDB).

MAC Address	Port	Age
00:00:00:00:00:AA	1	1 minute
00:00:00:00:00:BB	2	7 minutes
00:00:00:00:00:CC	3	30 seconds
00:00:00:00:00:DD	1	1 seconds

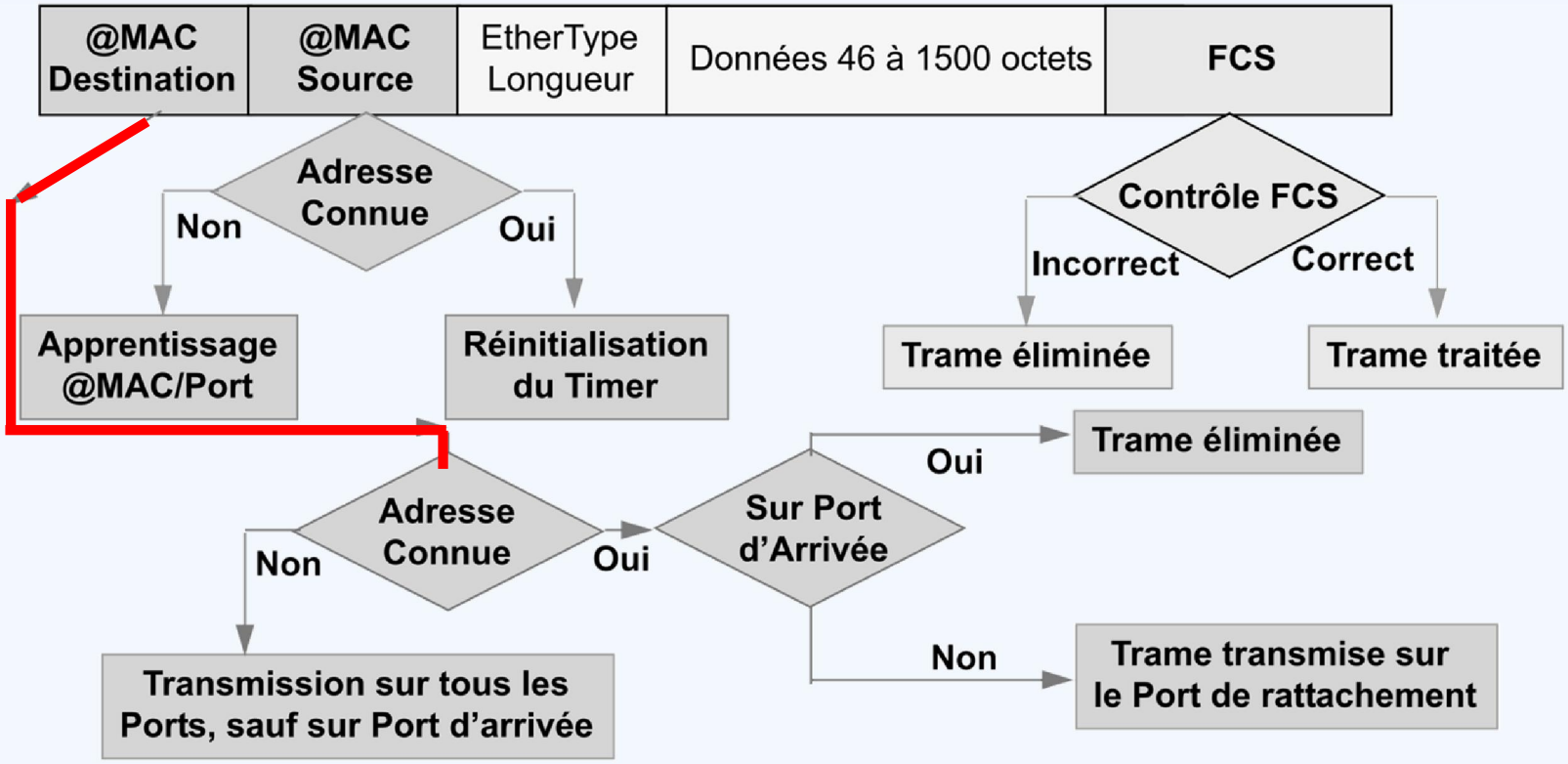
The value of the delay is set at 5 min, it can be set from 10 s to 11 days

Learning Example

- <Src=**A**, Dest=**F**>
- <Src=**C**, Dest=**A**>
- <Src=**E**, Dest=**C**>



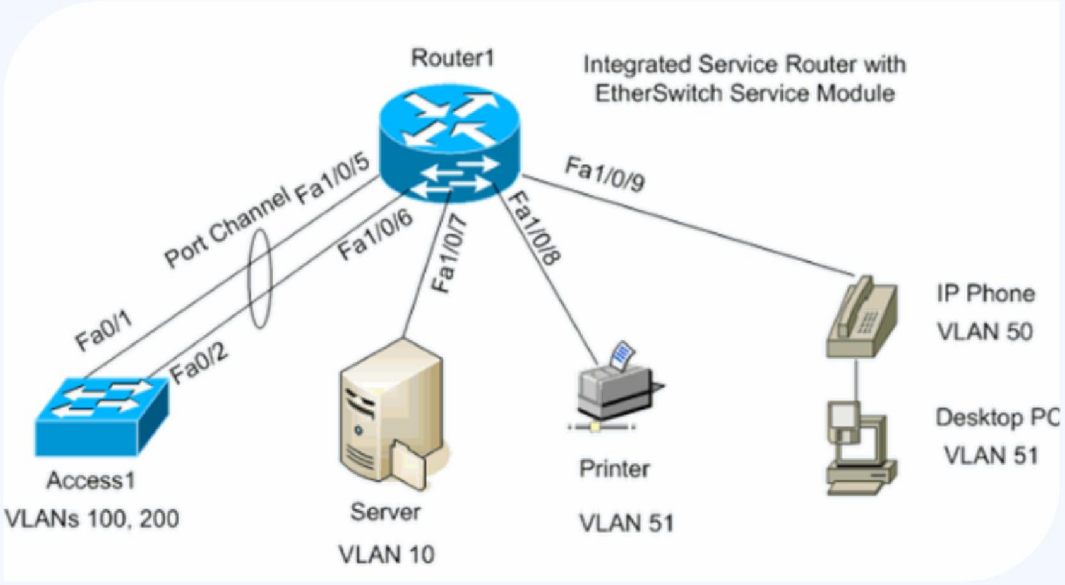
Switching Steps



Ports of the Switch

Like routers, switch ports follow either (motherboard/port) or (motherboard/slot/port) conventions.

- Fa0/1 is assigned to the first Fast Ethernet port.
- Fa0/2 is assigned to the second Fast Ethernet port.
- The Gigabit ports are labeled Gi0/1, Gi0/2 etc.



Switching Methods

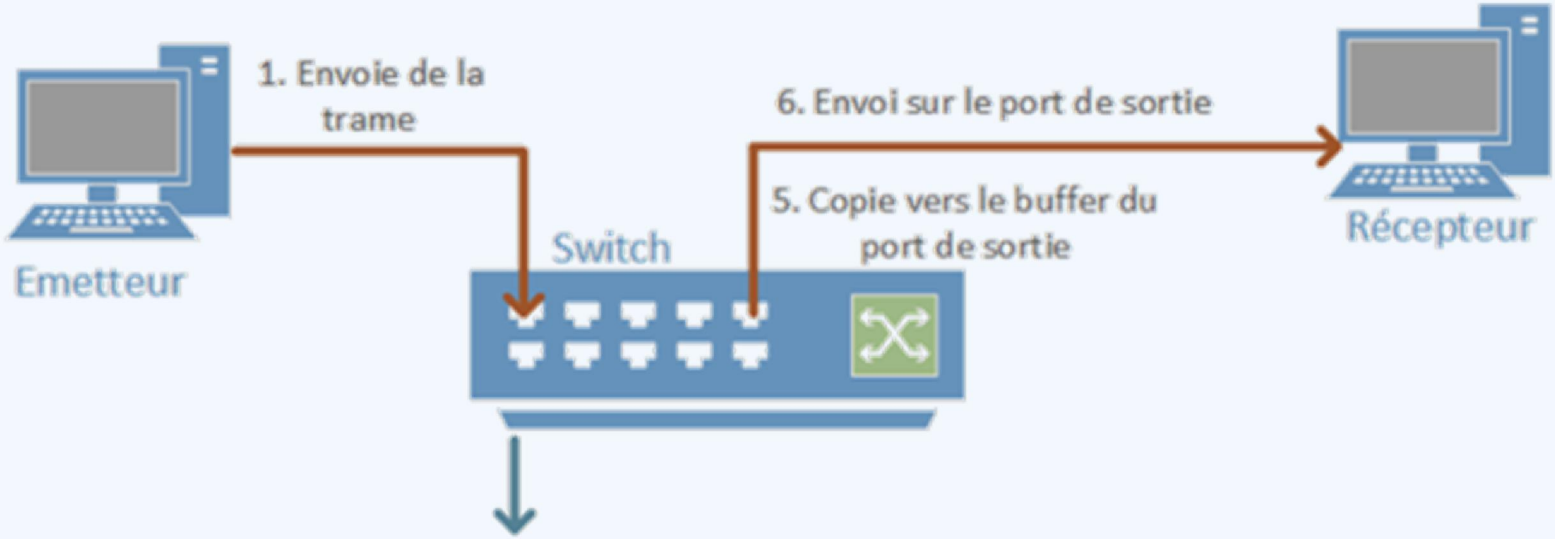
The switching is essentially carried out in 3 modes:

- Store-and-forward
- Fast forward
  - ✓ On-the-fly switching (Cut-through)
  - ✓ Fragment-free

Performance evaluation: latency (the time taken between the first incoming bit and the last outgoing bit) ...

Switching Store and Forward

The frame is sent to the recipient once it is received and the SCF was verified.



2- Buffering
3- Complete receipt, then verification of the CRC
4- if Ok, reading the switching table + transfer decision

Switching Store and Forward (2)

Automatic Buffering
The Switch puts the frame in a buffer, it waits for it to be fully received

Error Checking
The Switch performs a CRC (*Cycle Redundancy Check*) sometimes called FCS (*Frame Check Sequence*). It is calculated by the sending network adapter and included at the end of the frame. When received, the Switch recalculates it and compares it to the CRC sent.

Switching Store and Forward (3)

The frame is sent to the recipient once it has been totally received and the SCF was verified.

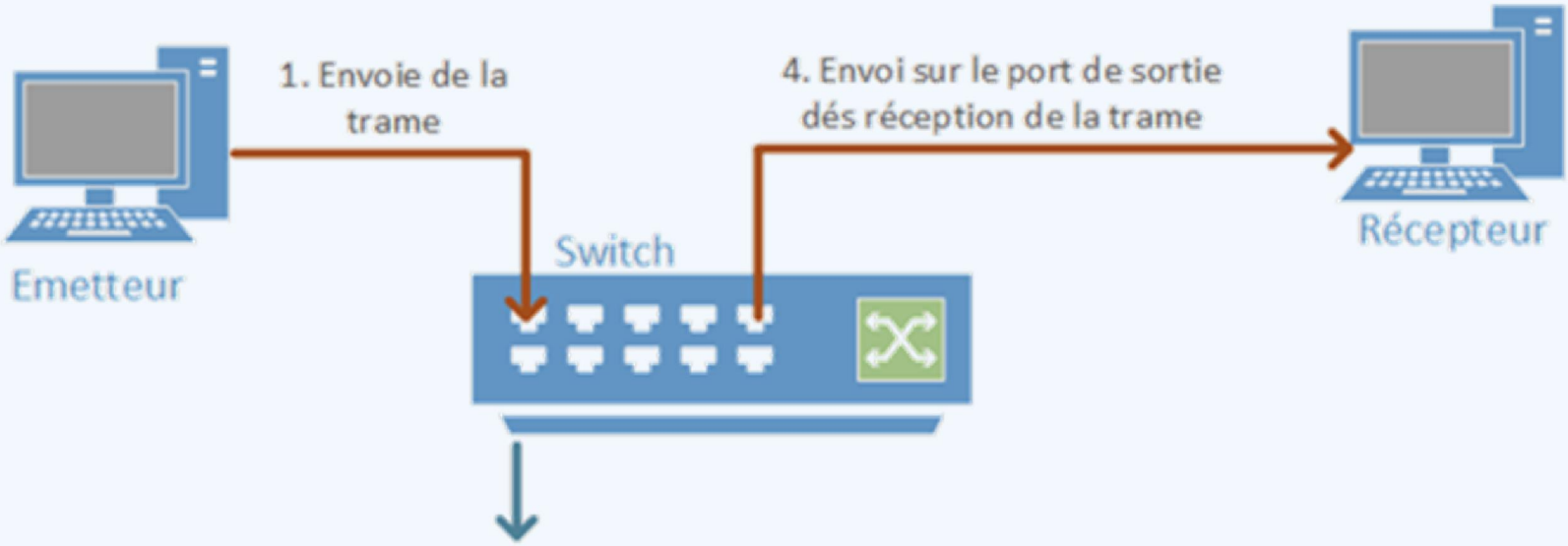
Benefits:
Error handling, using the CRC
Possibility of special treatments on the grid

Disadvantages:
Latency time depending on the length of the frame...

Switching Cut-through

Switching Cut-through

As soon as the destination address is parsed, the frame is sent to its destination. The retransmission is carried out while the rest of the frame is received.



2- Waiting for the header to be received (MAC destination)
3- Once the header received, read the switch table + transfer decision

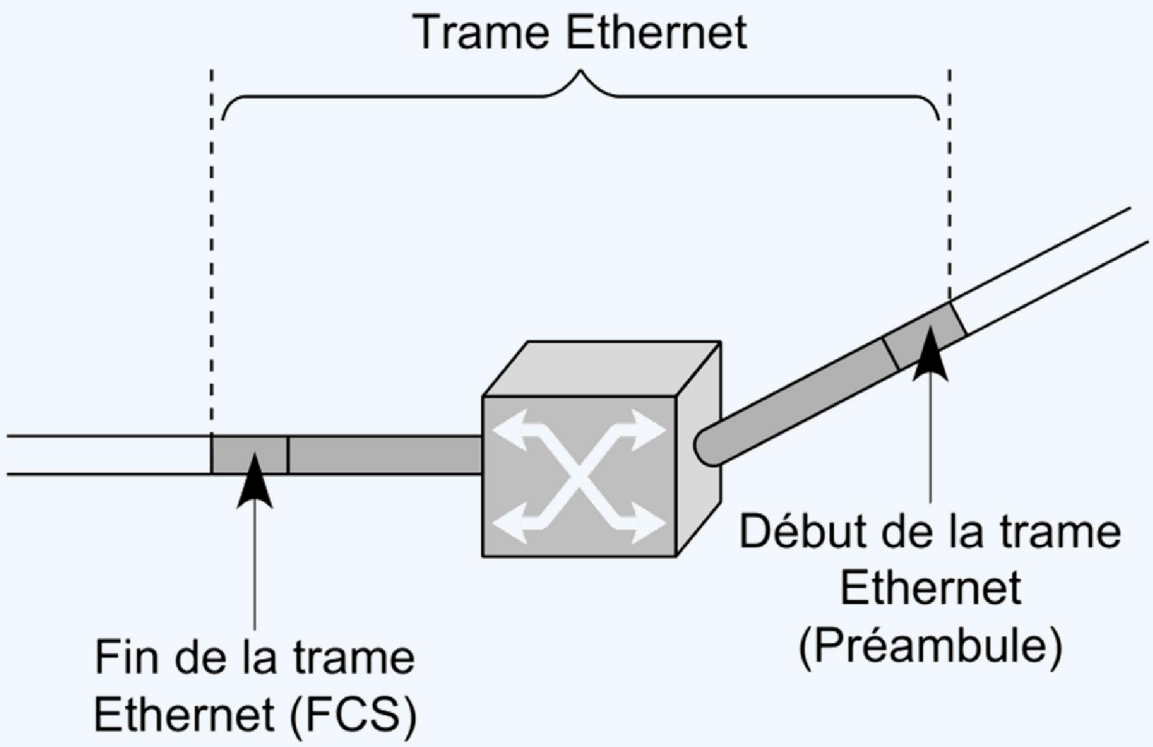
Switching Cut-through (2)

Benefits:

- Independent of the length of the frame
- Low latency

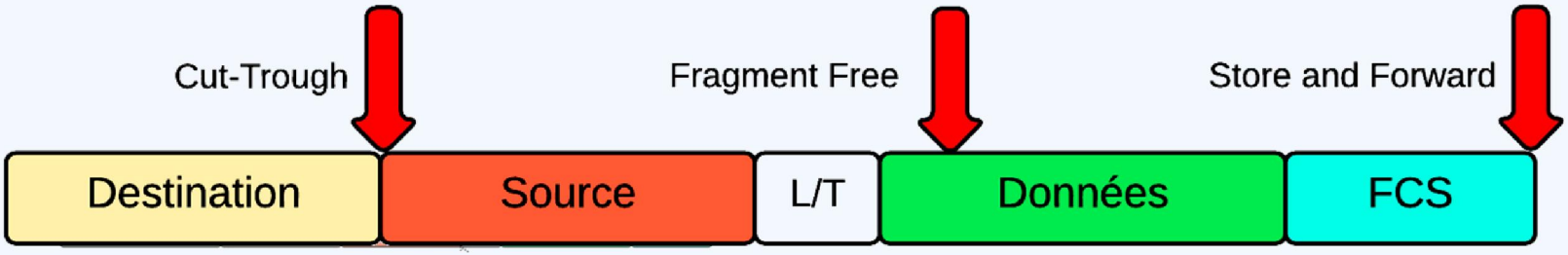
Disadvantages:

- Error retransmission...



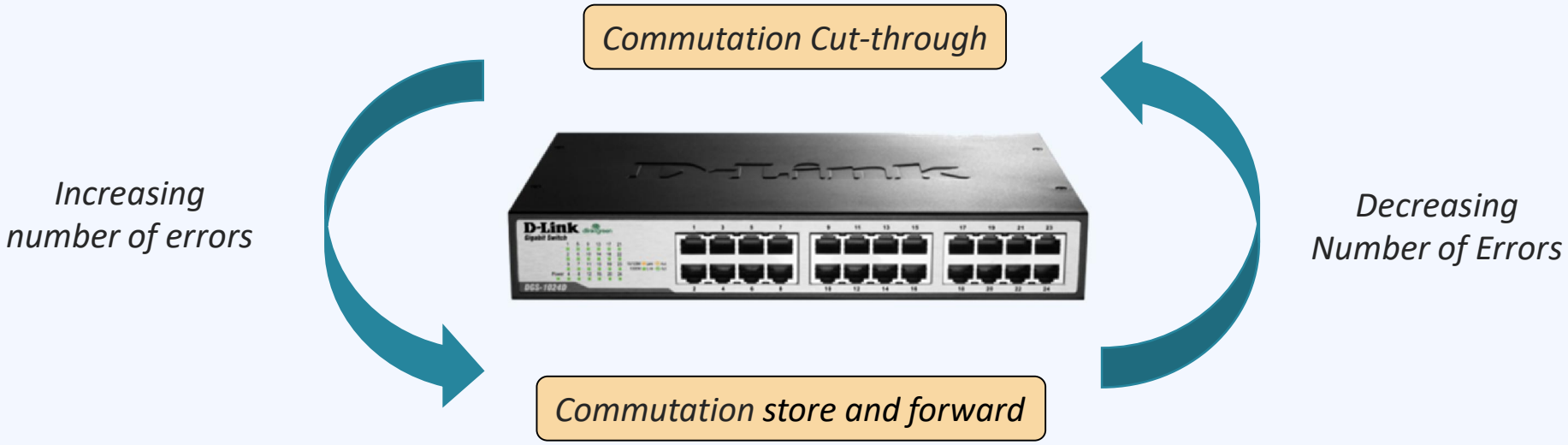
Switching Fragment-free

Read the first 64 bytes before retransmission. Typically, because collisions occur in those first 64 bytes, the switch can filter and eliminate collision frames.



Mode Cut-through adaptatif

- This method works on the fly at first, but checks the FCS in passing (error counter). Once an error rate threshold has reached, the on-the-fly method is abandoned in favor of Store and Forward.
- If the number of errors falls below the threshold, it goes back to Fast-Forward mode.



Types of Switchs

There are 3 types of Switch:

- Unmanaged Switches and
- Smart Switches
- Managed Switches.

To link a few computers together, an unmanaged Switch will be more than enough.

Unmanaged Switch

Administrable Switch (Fully Managed)

After 30 to 40 computers, "manual" management can quickly become a "headache". Management by managed Switch is therefore essential. It can be possible to:

- monitor the network,
- perform statistics,
- manage traffic priorities.
- Remotely programmable.
- Set permissions or restrictions for access to machines that are part of a local subnet or VLAN.

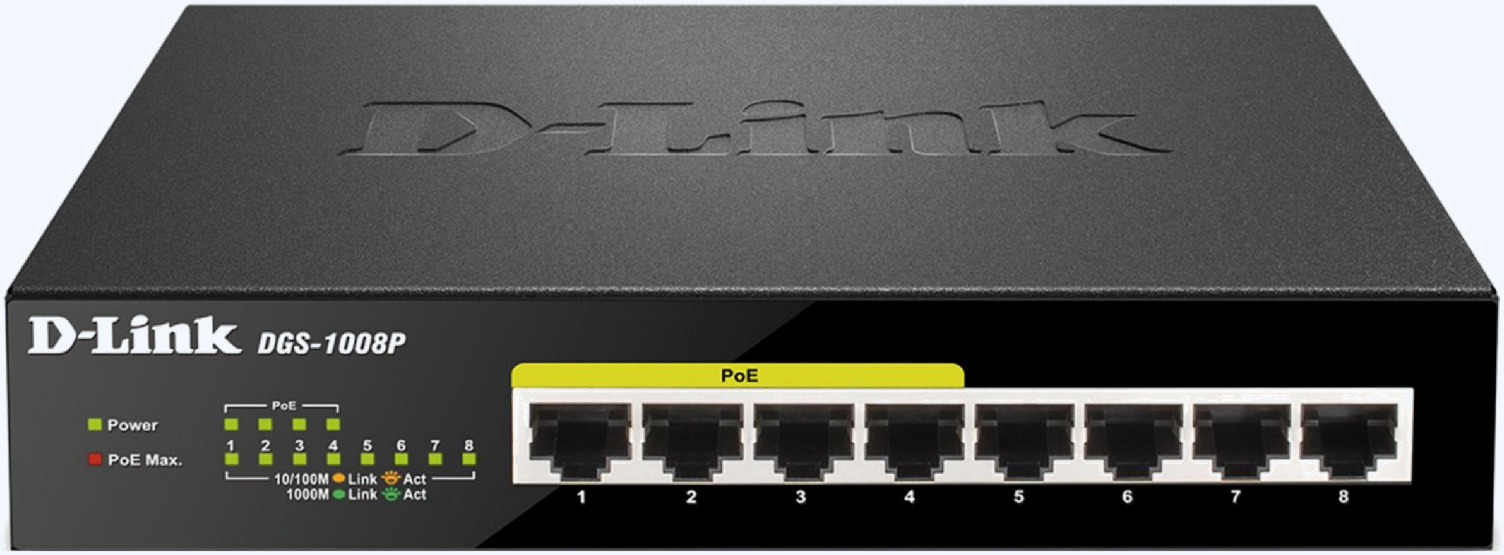
Smart Switch (*Smart Managed Switches*)

Smart Switch that can be partly managed via a web interface. These switches are a good compromise between an unmanaged switch and a managed switch.

Switch Features

The Ethernet PoE switch

A Power over Ethernet (PoE) switch is used to power connected devices, such as IP cameras or VoIP phones, via RJ45 Ethernet cables.



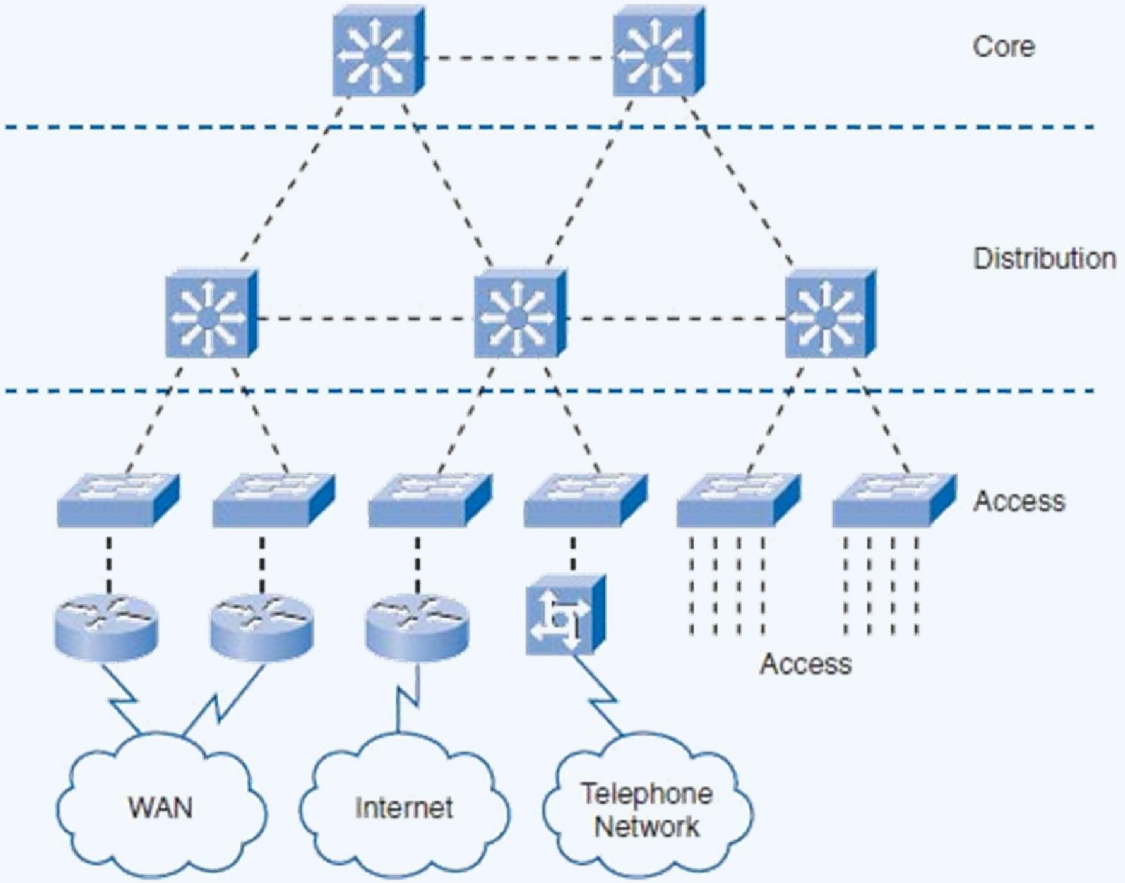
Switch Features

Modular switches

These switches allow expansion modules to be added to the switches as needed, providing flexibility as your network needs change.

Application example

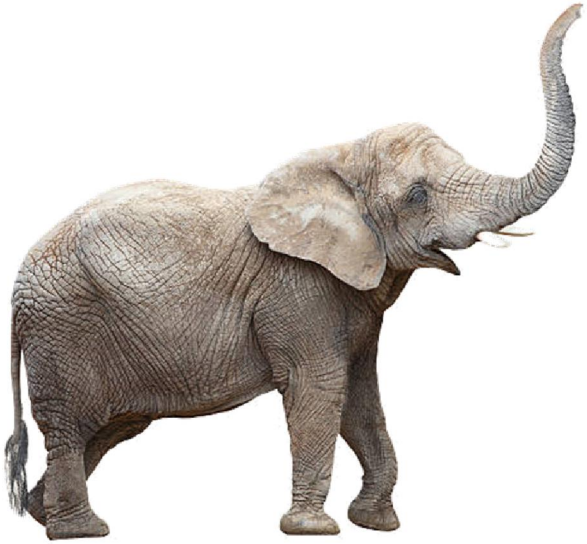
Core Switch Vs Distribution Switch Vs Access Switch

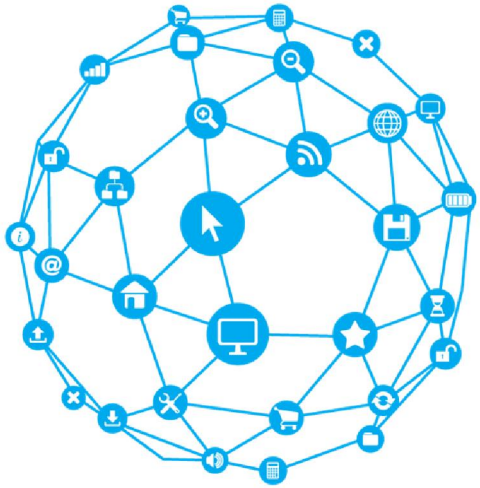


Core Switch

Chapter 02

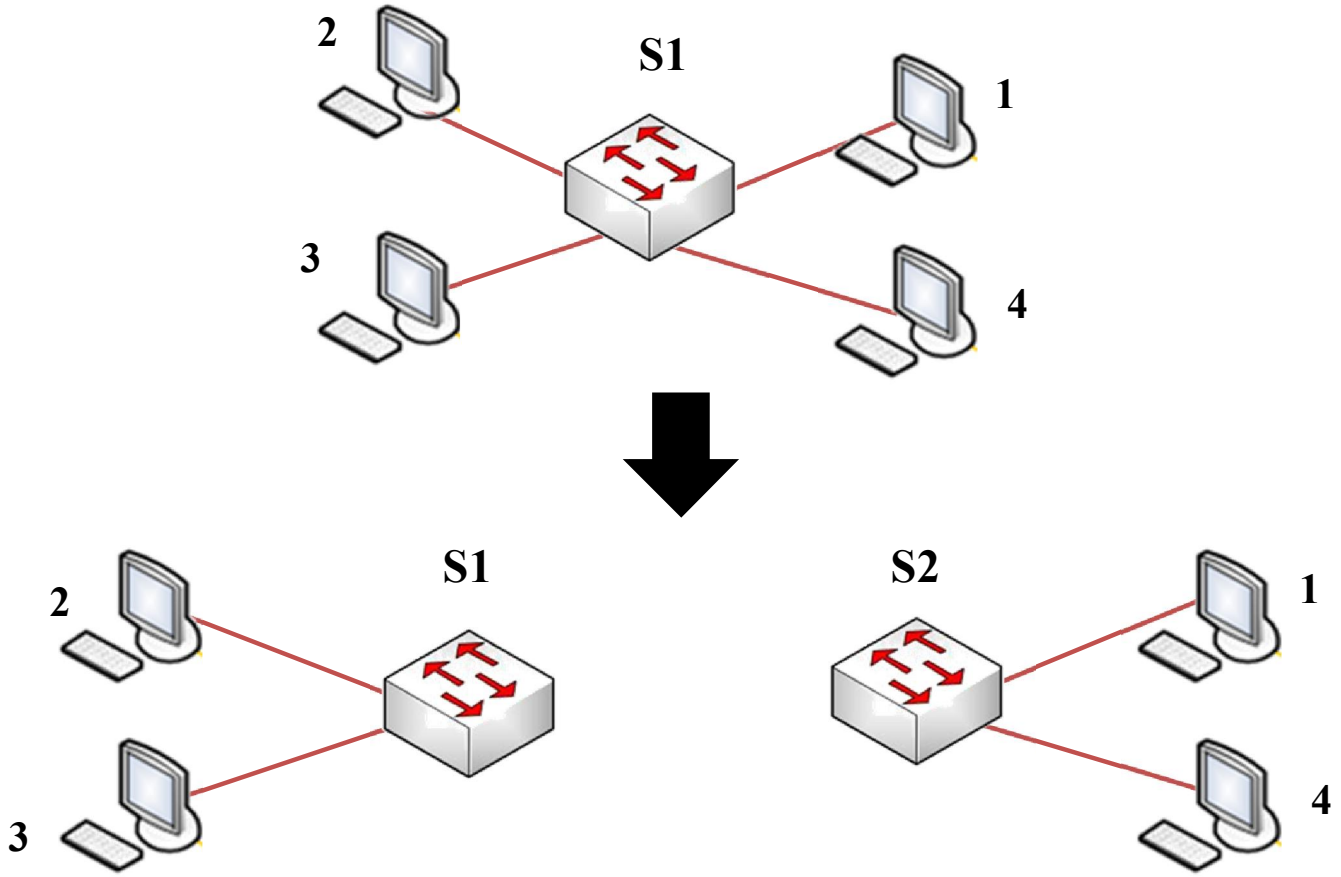
The VLANs





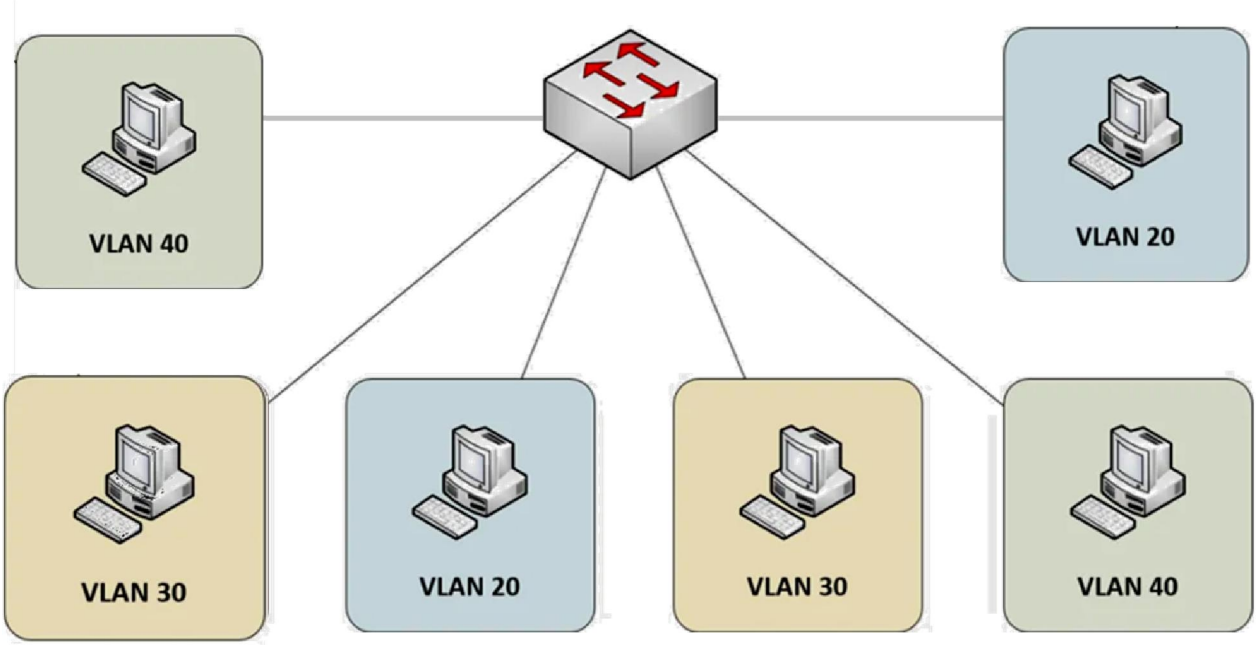
What exactly is the problem!!

If we want to use several subnets, then we need several Switches.



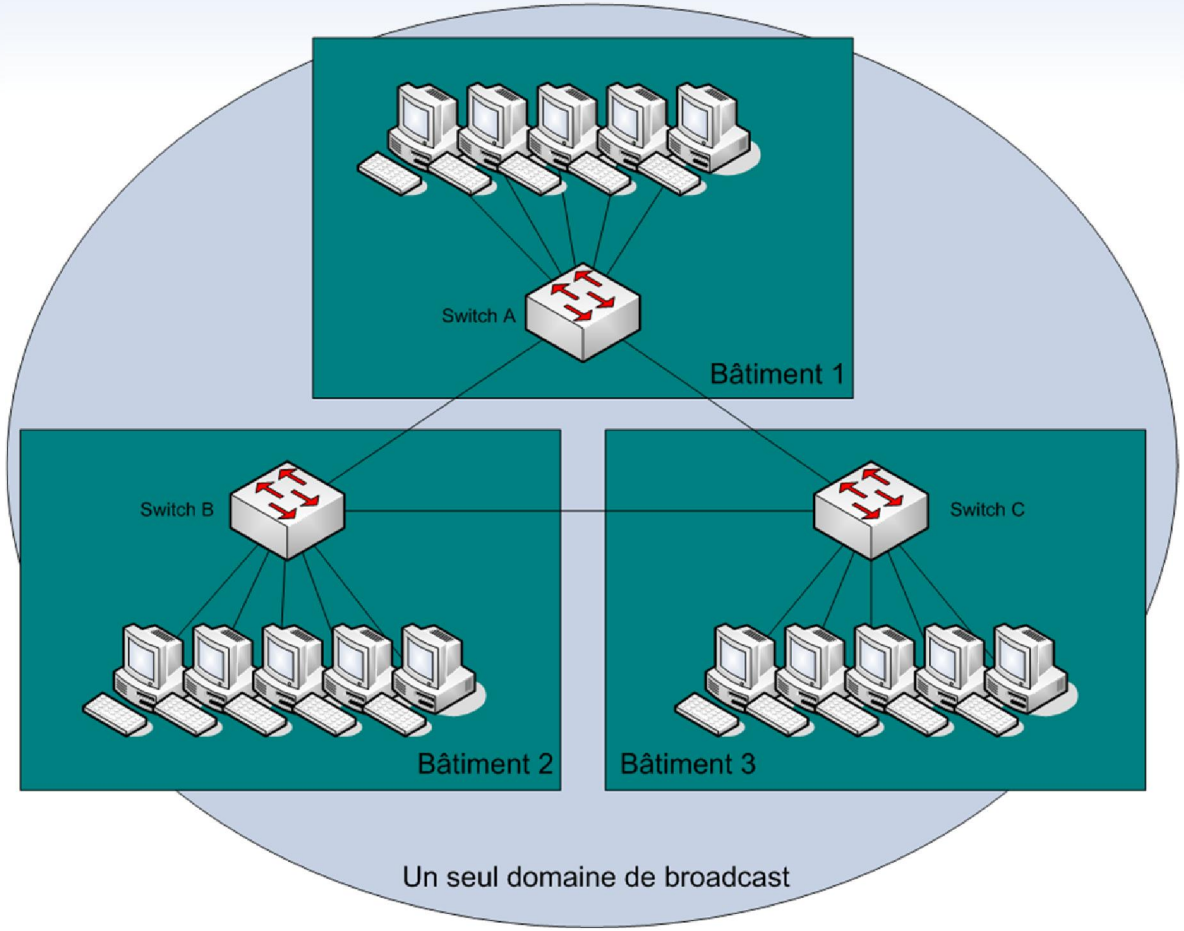
Virtual Local Area Network (VLAN)

Before, you had to physically separate the equipment, with VLANs you can get rid of this physical separation and make a logical separation within the same Switch,



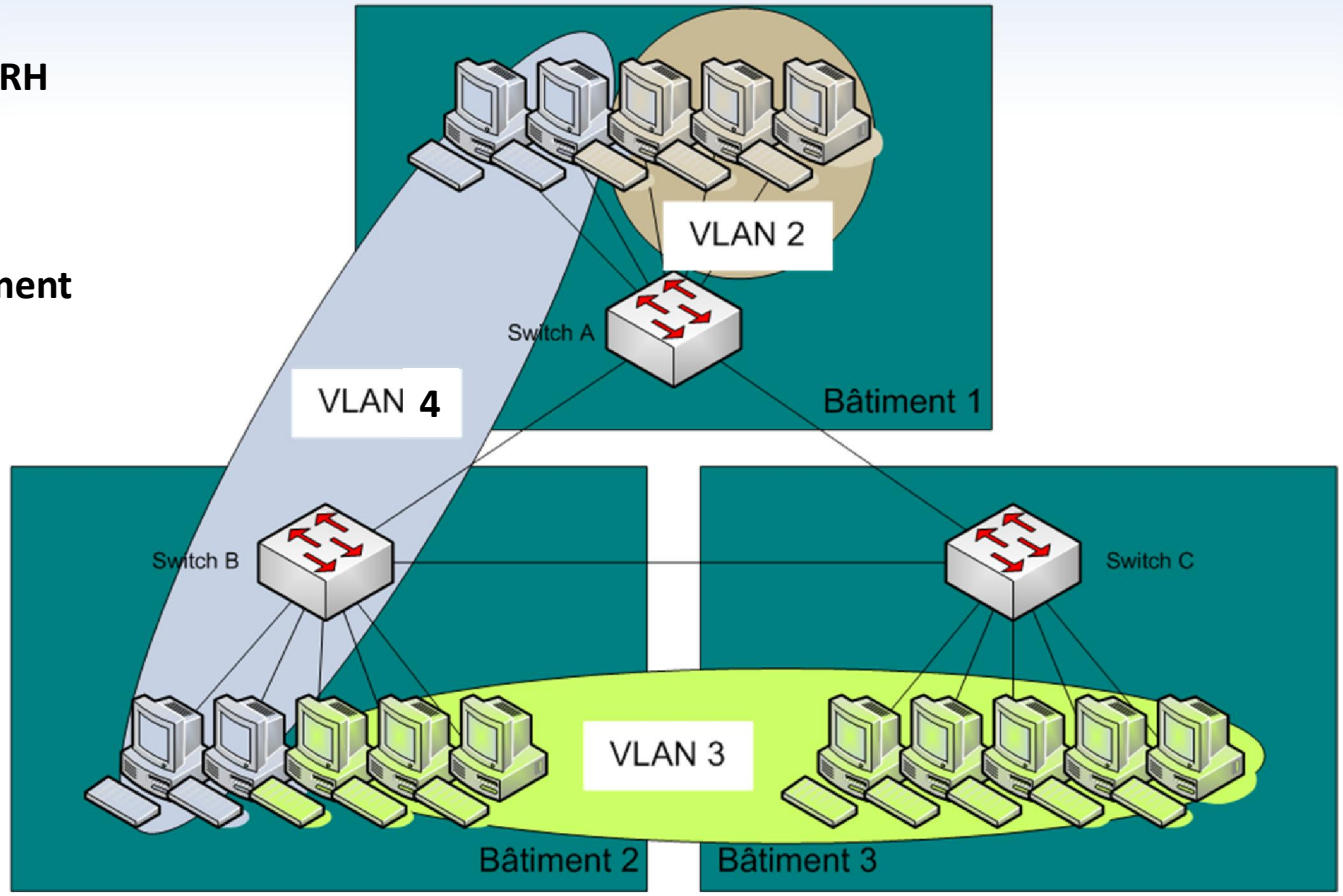
And more

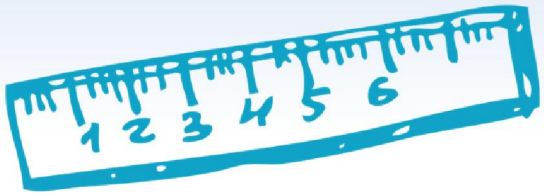
- English Department RH
- Biology Department
- Engineering Department



Shared VLAN

- English Department RH
- Biology Department
- Engineering Department





Decrease Broadcast Domain

1 Broadcast domain per VLAN

What for?

- Allows us to subdivide Switches into separate (virtual) Switches,
- Only members of a VLAN can receive traffic from that VLAN, secure domains,
- Inter-VLAN traffic must pass through a router,
- Broadcast frames are isolated. Limits the domains of distribution,
- Greater flexibility in job mobility and network segmentation (geographical independence),
- Easily be able to assign different permissions, depending on the rights and roles of each group of people....

With VLANs, the machines of the same subnet are no longer grouped together physically (connected to the same switch), but in a logical way.

VLAN levels

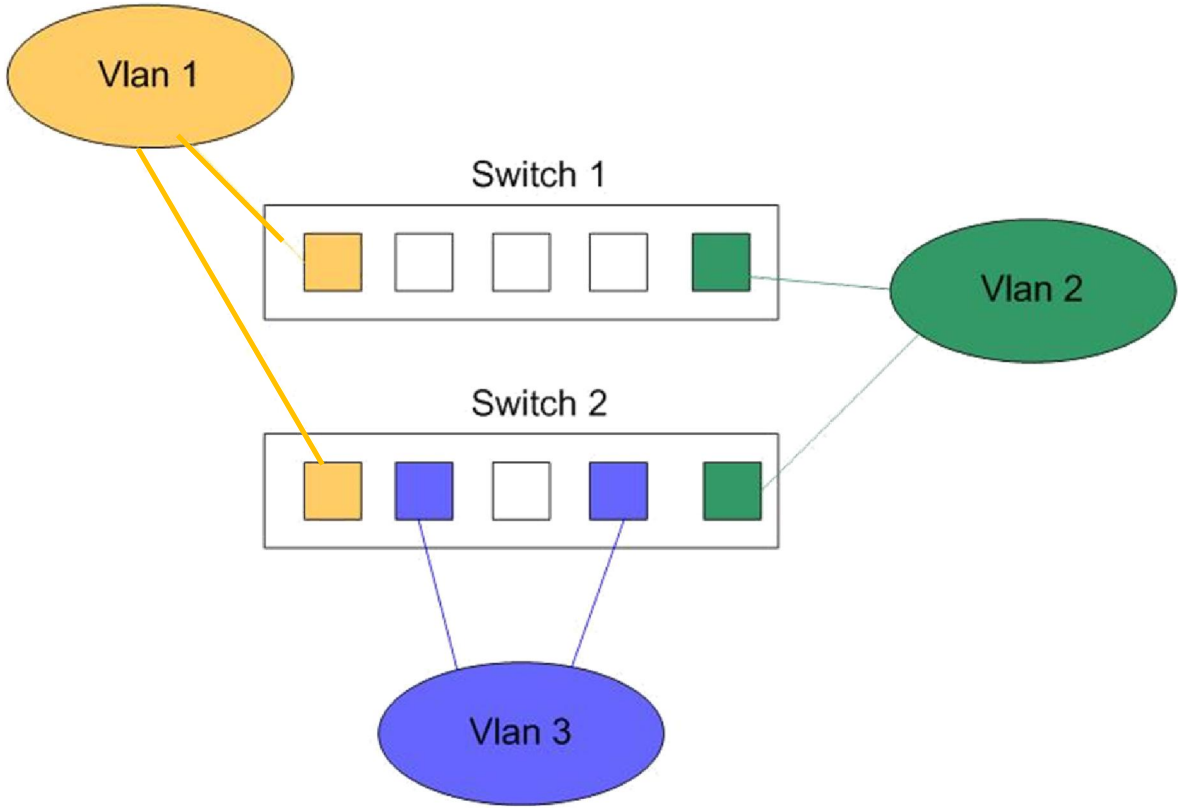
The different levels of VLANs

There are several types of VLANs, depending on how they work, we can associate them with a particular layer of the OSI model.

- Level 1 VLAN associated with the physical layer
- Link-layer 2 VLAN
- Level 3 VLAN associated with the network layer…

Level 1 VLANs

- It is the port that determines the VLAN to which the associated stations belong



Level 1 VLANs (2)

- There is no heavy processing for each frame in the switching process
- Configuration is static (static VLAN)
- It is the most secure.
- Stations that are connected to a port can only belong to one VLAN.

However, this type has some significant drawbacks

- Requires changes to VLANs when adding or removing users
- Does not support cascaded switches
- We don't have a lot of flexibility.
- 90% of the vLAN's used are level 1.....

Level 2 VLANs

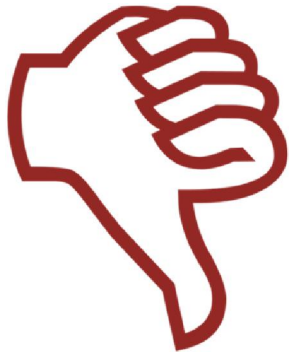
Also known as MAC Address-based VLANs:

- Assignment based on the **MAC address** of the machine and not its location on the physical network.
- The stations are associated by their MAC address. Two stations on a port can belong to 2 different VLANs.
- the ports automatically determine their VLAN to which they belong (dynamic VLAN).
- A station can belong to multiple VLANs. Layer 2 VLANs are independent of higher protocols.
- Switching, which is carried out at the MAC level, allows a low latency .....

Level 2 VLANs (2)

Disadvantages:

- The independence of the user's location makes it difficult to locate them on the network.
- Adds difficulty in applying filter rules.
- For each new user, it is necessary to introduce it.....



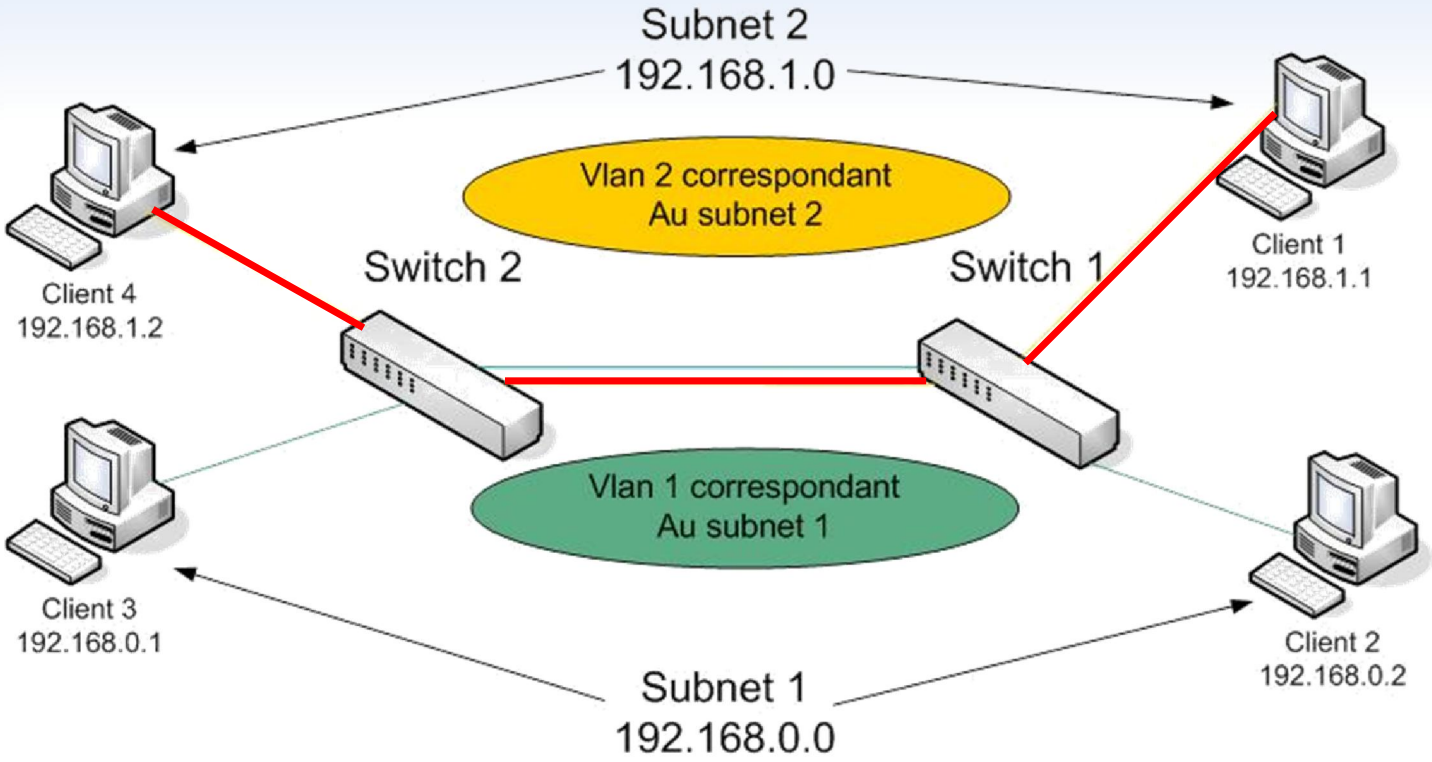
Level 3 VLANs

Also known as Network Address-based VLANs:

- Allows you to group machines according to the subnet to which they belong
- Users of a Layer 3 VLAN are affected dynamically to a VLAN, following an IP address,
- The switch must access the Layer 3 (network) address to set the VLAN of ownership (less performance)...

The level 3 address is used as a **label**, it is indeed switching and not routing. The header is not changed.

Level 3 VLANs (2)



Level 2 and 3 VLANs obviously allow for greater flexibility and mobility.
However, these two types are not widely used.

Level 3 VLANs (3)

Disadvantages:

- Requires more resources to isolate the many flows flowing through the network: the switch must decapsulate each packet to determine which vLAN it belongs to.
- A hacker can quite easily determine the IP address used by a user to impersonate them.
- Requires a Level 3 Switch (relatively expensive)..

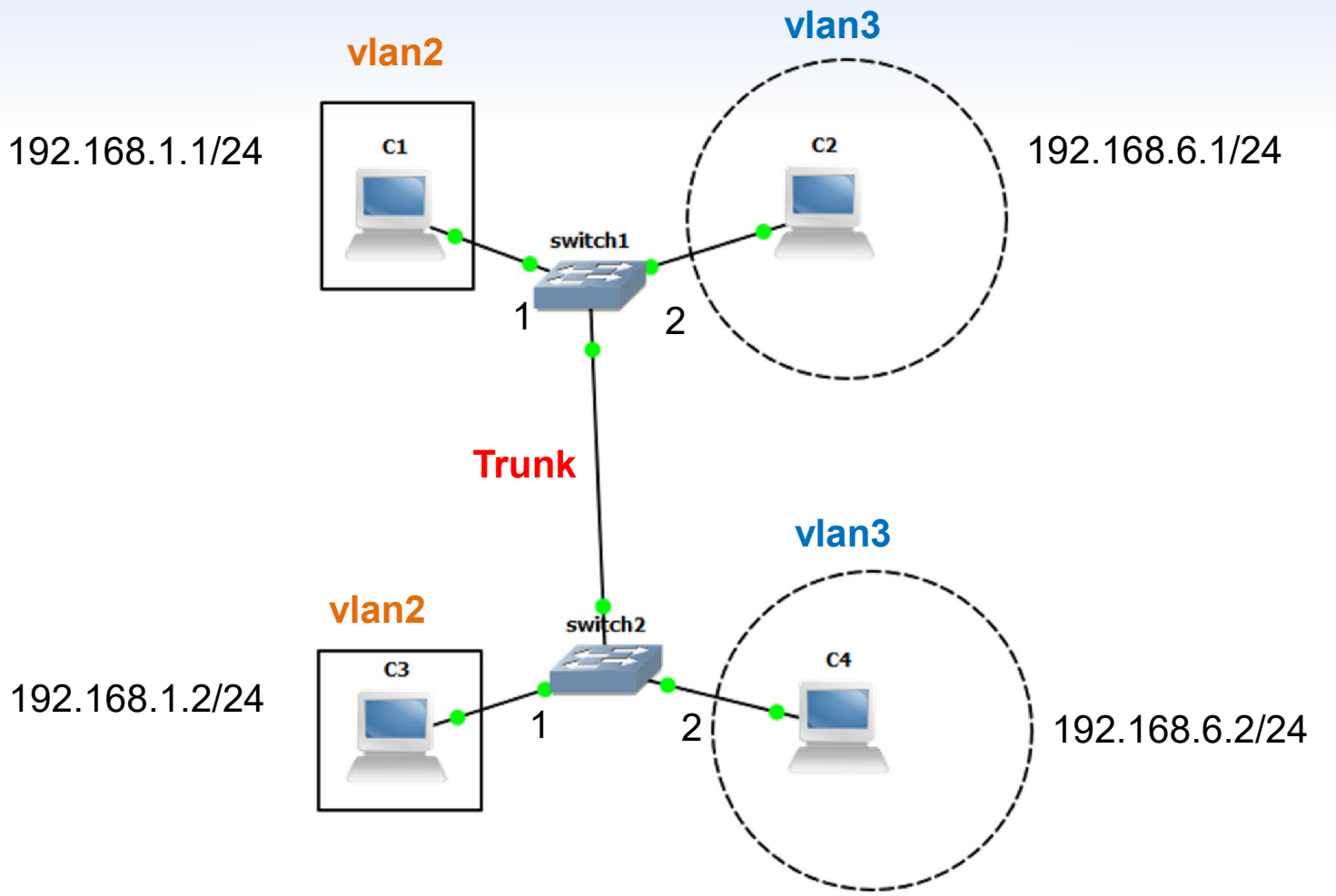
And others ...

It is also possible to create VLANs by:

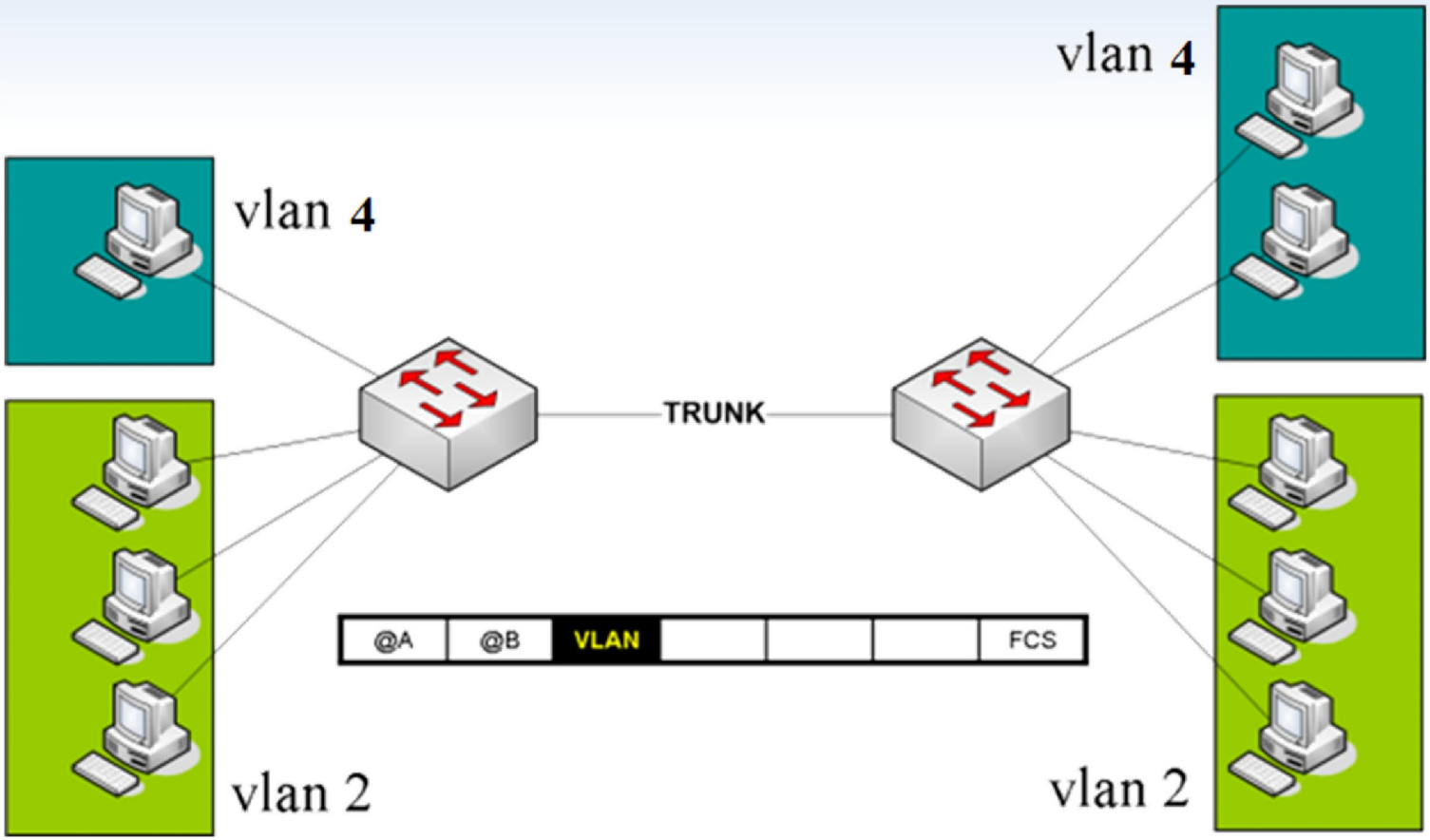
- protocol, communication can only be established between stations using the same protocol;
- per application (TCP port N◦), the constitution of VLANs is then dynamic, a user can successively belong to different VLANs depending on the application he uses;
- by password (dynamic creation of VLANs at user login).....

Concept of labelling

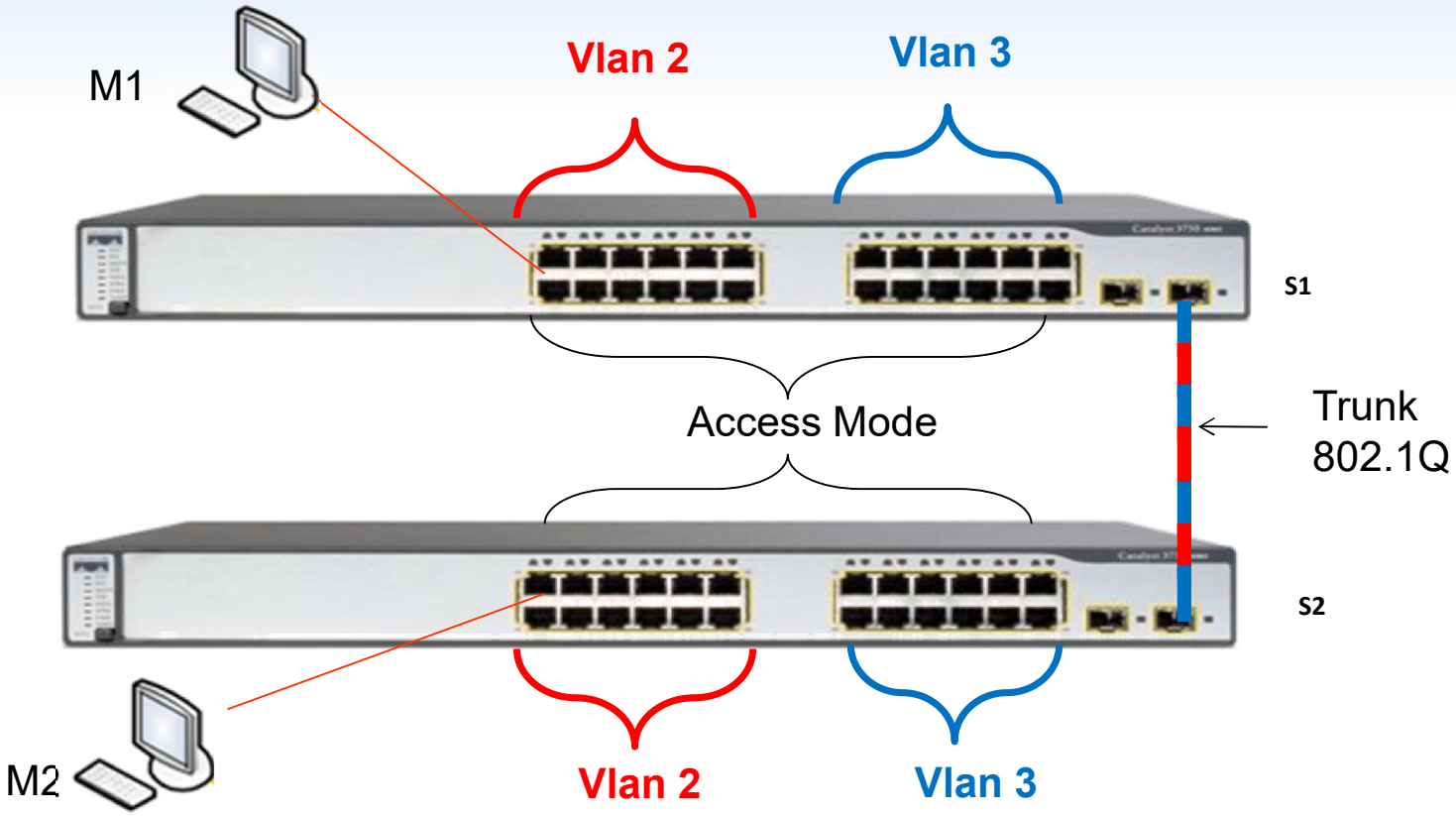
But how!!


vlan2
vlan3
192.168.1.1/24
C1
C2
192.168.6.1/24
switch1
1
2
Trunk
vlan3
vlan2
switch2
C3
C4
192.168.1.2/24
1
2
192.168.6.2/24

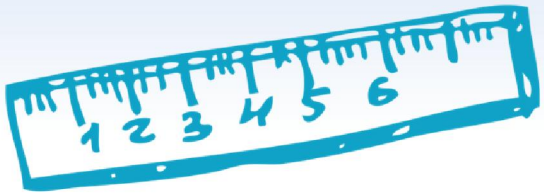
Aggregation link



The Trunk



Each frame has a **tag** that identifies the frame to which VLAN it belongs



The label
identifies the VLAN of the source station

The Trunk (definition)

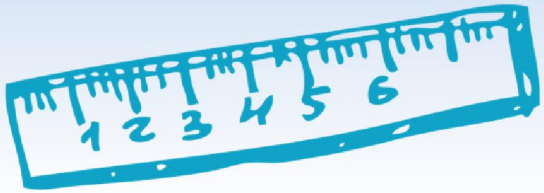
- A trunk is a physical connection over which traffic from multiple virtual networks is transmitted.
- It extends VLANs across an entire network
- A Trunk does not belong to a specific VLAN; it is rather a conduit for several VLANs.
- They are called a "Trunk" by Cisco Systems and "aggregation link" by others.
- The switch adds additional fields in the Ethernet frame to distinguish traffic from different VLANs...

VLAN Labeling Technologies

Each frame with a VLAN label has fields indicating that it belongs to a VLAN. There are two main formats of VLAN labels:

- Cisco ISL inter-switch link format. It is less and less encountered in favor of:
- the standard 802.1Q format.

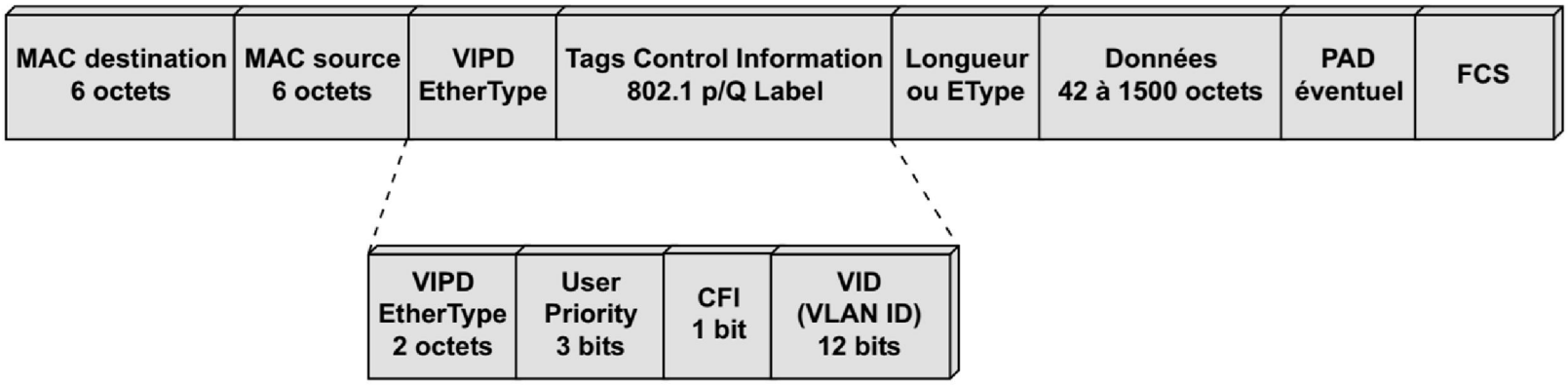
The ISL inter-switch link format is a Cisco proprietary format for VLAN labels. This VLAN label adds 26 bytes of information to the front of the frame and a CRC of 4 bytes at the end of the frame..



By default, the ports on a switch are part of VLAN 1

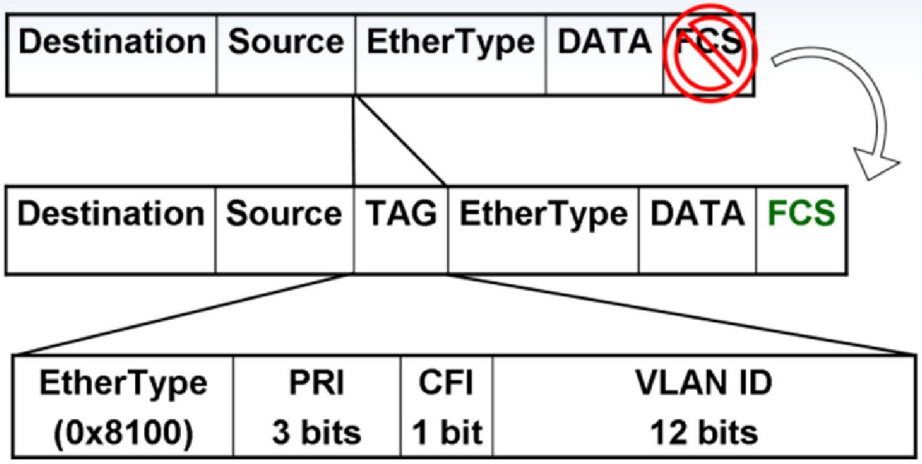
Format standard 802.1Q

The 802.1Q standard introduces four **additional bytes** into the MAC frame. They are used to identify VLANs (VLAN tagging) and manage **8 levels** of priority (QoS).



MAC frame of 802.3 networks + Tag

standard Format 802.1Q



- ✓ TPID Tag Protocol Identifer: 0x8100 for 802.1Q
- ✓ Priority: Priority levels defined by IEEE 802.1Q
- ✓ CFI: Ethernet or token-ring (value 0 in Ethernet)
- ✓ VID: VLAN identifier,

How many VLAN identifiers are possible?

VLAN range

Plage de VLAN	Utilisation
0, 4095	Réservé pour le système, non utilisable
1	VLAN par défaut
2–1001	VLAN Ethernet
1002–1005	VLANs créés par défaut pour les technologies FDDI et Token Ring
1006–4094	VLAN Ethernet – plage étendue (attention aux switchs anciens)

Types of ports

Types of ports

The two types of ports that can be used in a VLAN environment are access ports and link ports.

Access ports (untagged)

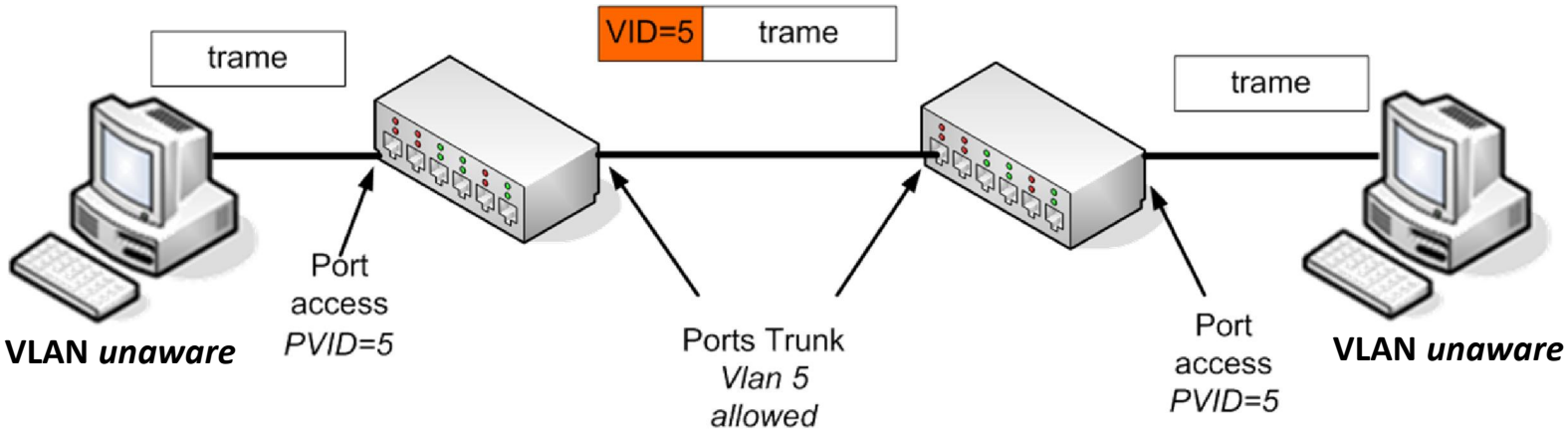
When an access port receives a frame, it does not have a VLAN label.....

Transmitting and receiving devices are unaware that a VLAN label has ever been used.

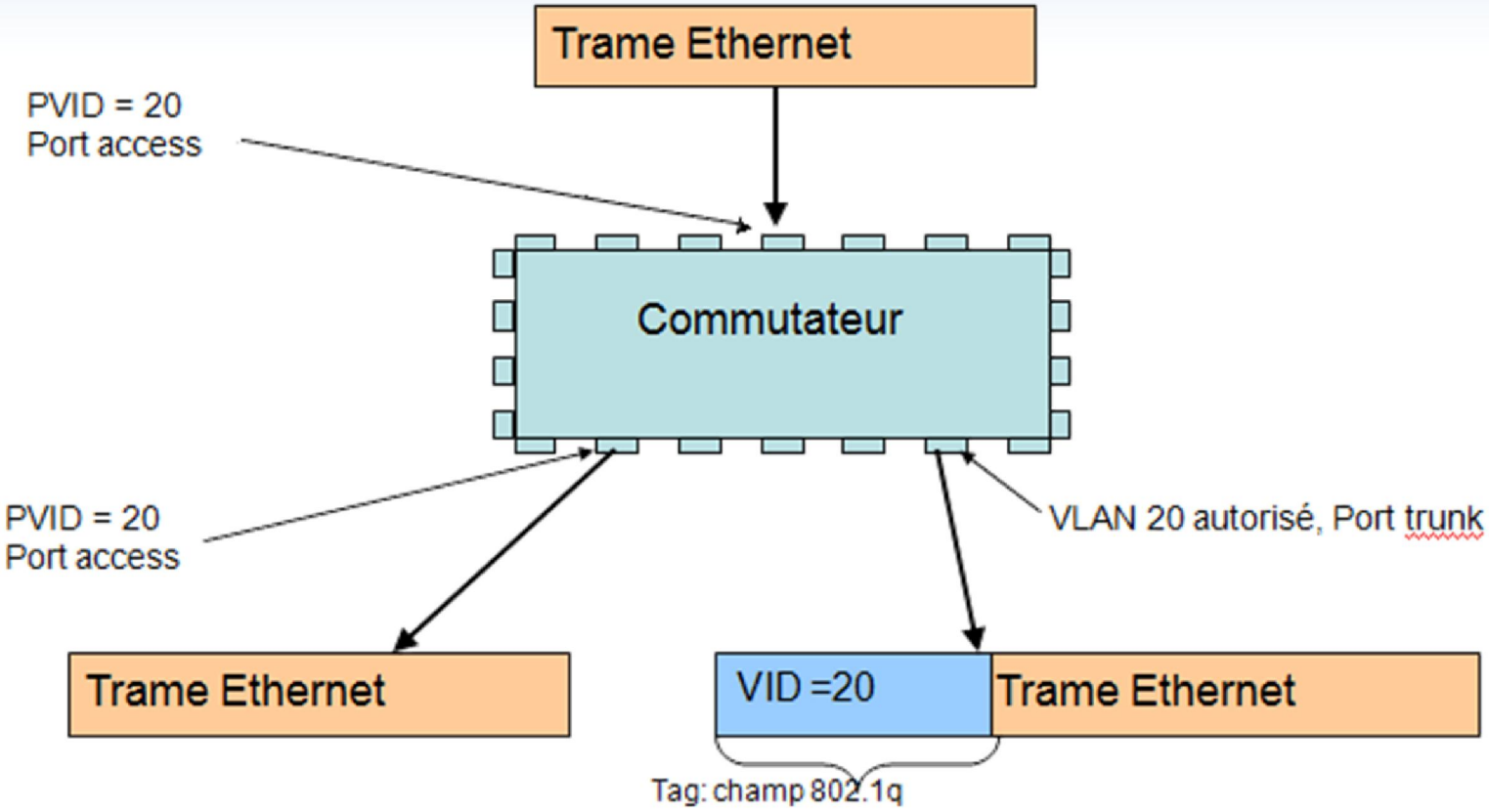
Port Types (2)

Link Ports Port Trunk or Tagged (Labeled)

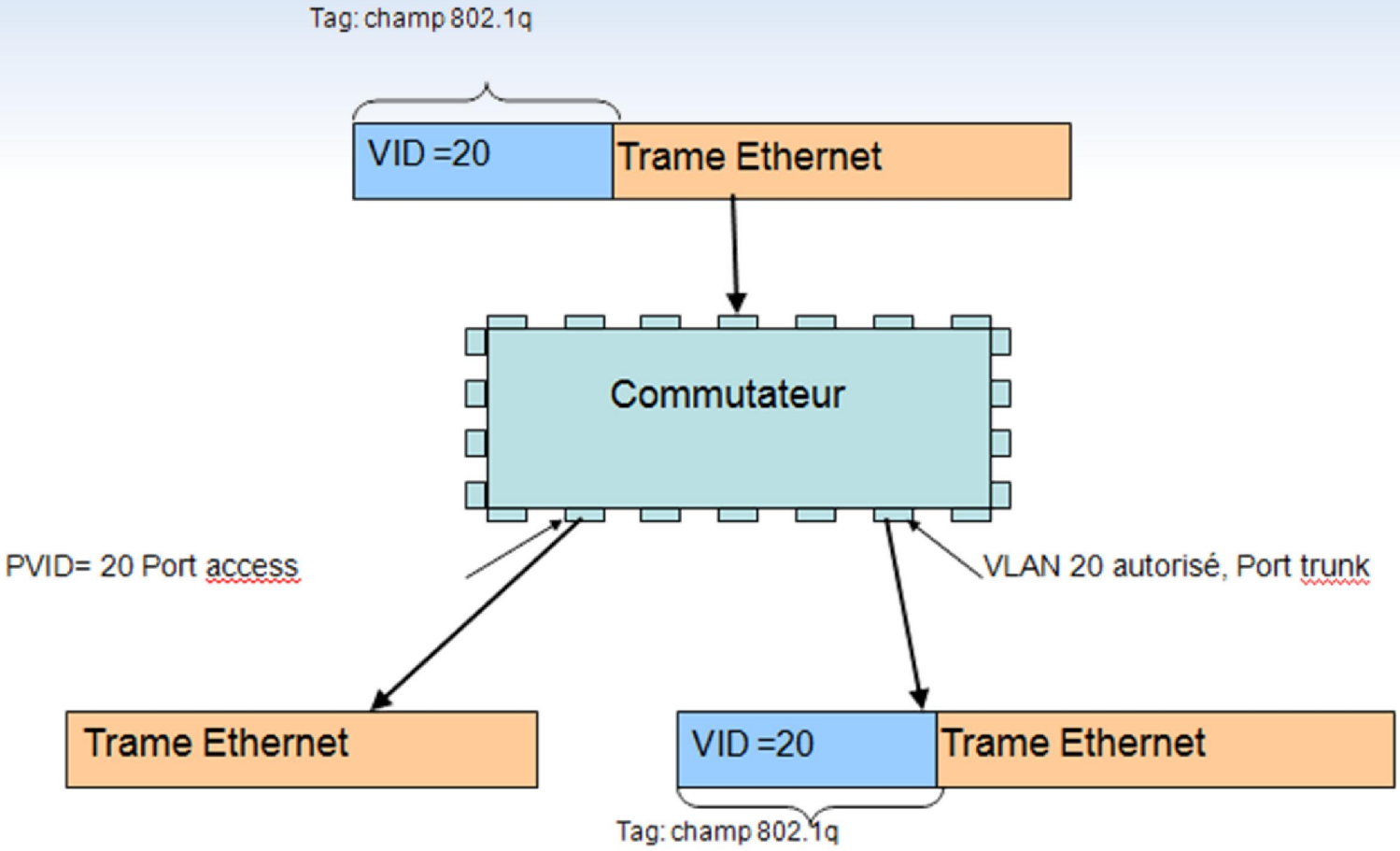
The difference between the access and link ports is that the link ports do not remove the VLAN label from the frame when they send it.



Port Types (3)


Trame Ethernet
PVID = 20
Port access
Commutateur
PVID = 20
Port access
VLAN 20 autorisé, Port trunk
Trame Ethernet
VID =20
Trame Ethernet
Tag: champ 802.1q

Port Types (4)



If a tagged frame arrives on an access port, it will be systematically rejected

inter-vlan routing

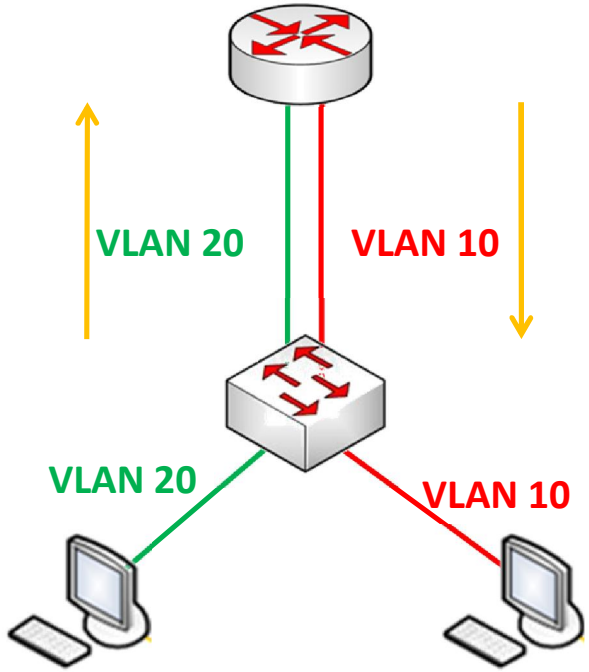
A router is the boundary of a VLAN like that of a LAN. As a result, for VLANs to communicate with each other, a routing function is required. This is called "inter-VLAN routing".

Three options are available when implementing inter-VLAN routing:

- Traditional or legacy inter-VLAN routing
- Router on a stick
- Multi-layer switch

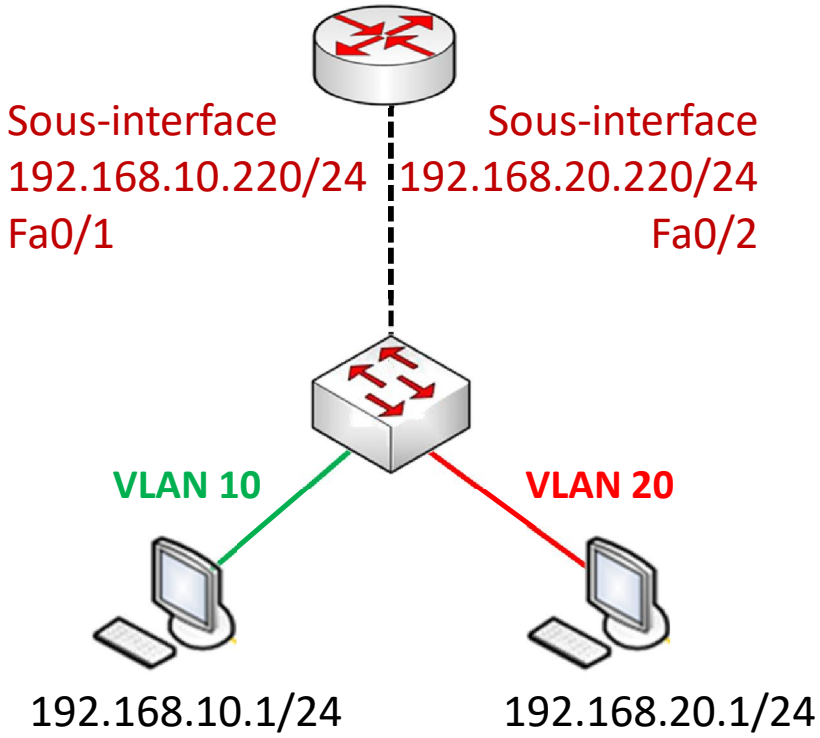
Traditional inter-VLAN routing

Traditional inter-VLAN routing requires multiple physical interfaces on the router and switch.



Router on a stick

For inter-VLAN routing `Router on a stick`, a single physical interface configured trunk on the router is sufficient, in the form of a multitude of virtual gateways (subinterfaces) with different IP addresses.

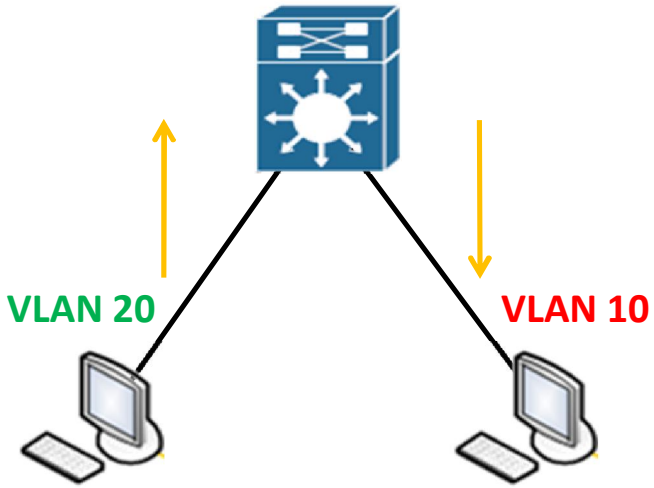


Router on a stick

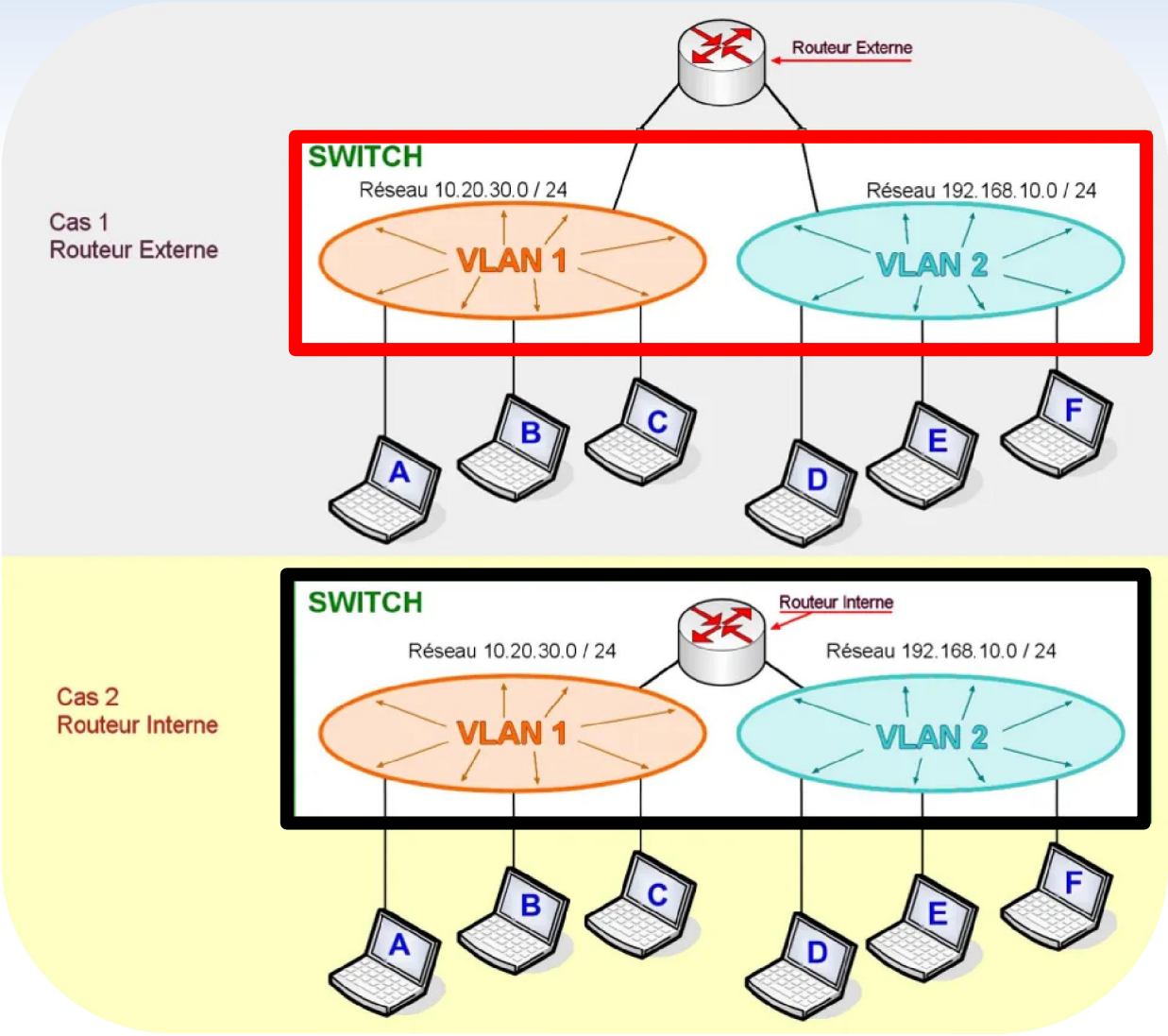
- Today, routers allow to configure a router interface as multiple trunks using subinterfaces.
- The physical interface is logically subdivided into several logical interfaces.
- Each subinterface of the router is assigned a VLAN.
- The router redirects routed traffic, marked VLAN for the destination VLAN, to the same physical interface that it used to receive the traffic..

Multi-layer switch

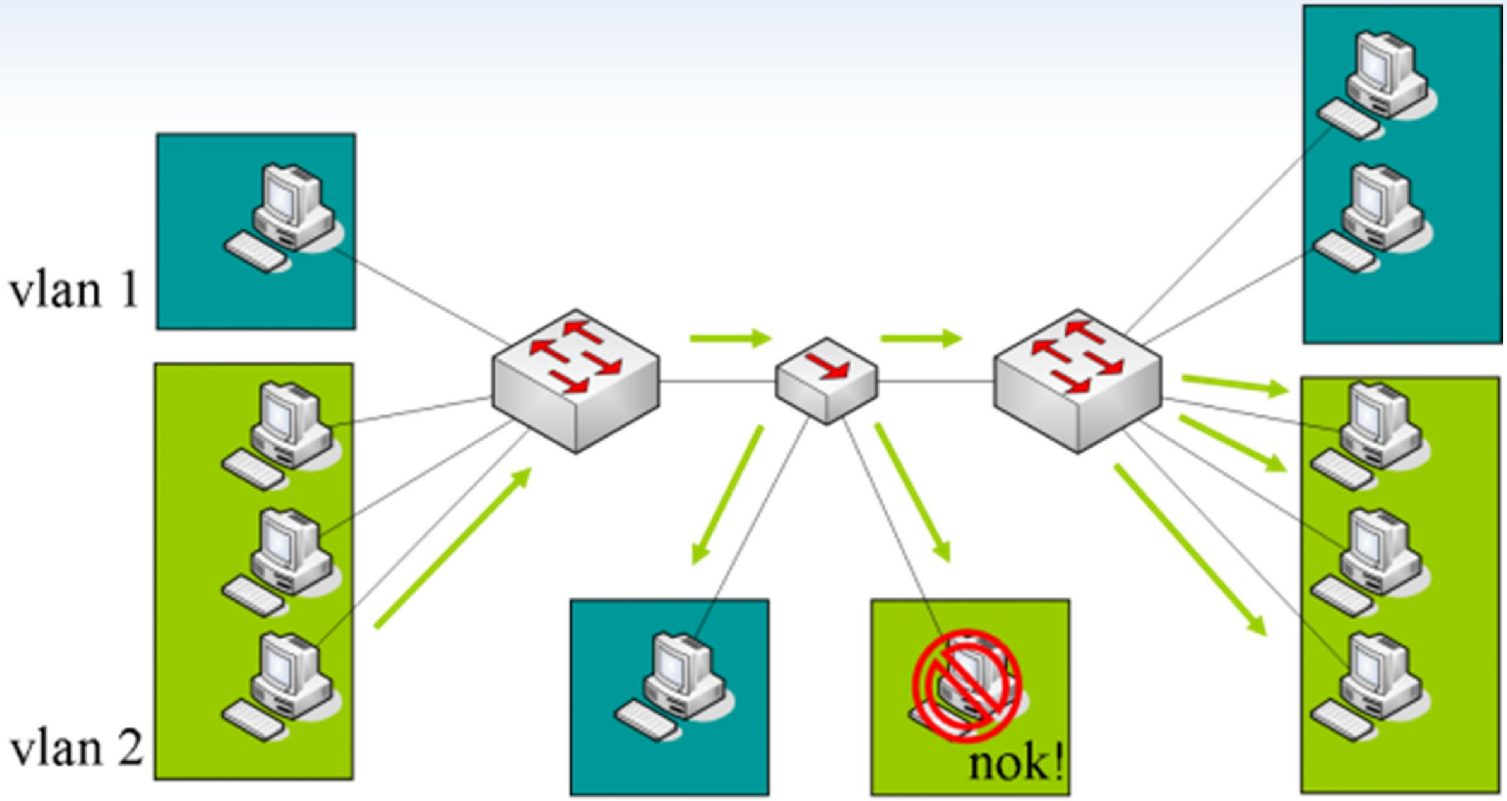
The most scalable solution in today's enterprise networks is to use a multi-layer switch to replace both the router (routing traffic between VLANs) and the switch (switching traffic within the same VLAN)



Inter-vlan Routing (Multilayer Switch)



Natif VLAN



In the **native VLAN**, the switches would let the initial frame pass through without adding a TAG.

Vlan Trunking Protocol

VTP (Vlan Trunking Protocol)

- It is used for the propagation of **creation/deletion/modification** of VLANs on all Switches of a network from a single Switch (server).
- It is a Cisco proprietary protocol. Due to its simplicity and power,
- The IEEE has released a similar protocol to allow this feature between switches from different manufacturers: GVRP (GARP VLAN Registration Protocol). The standard is IEEE 802.1ak.....

The role of VTP is to propagate VLAN configurations.

VTP Architecture

The Switch has 3 VTP modes: client, transparent or server

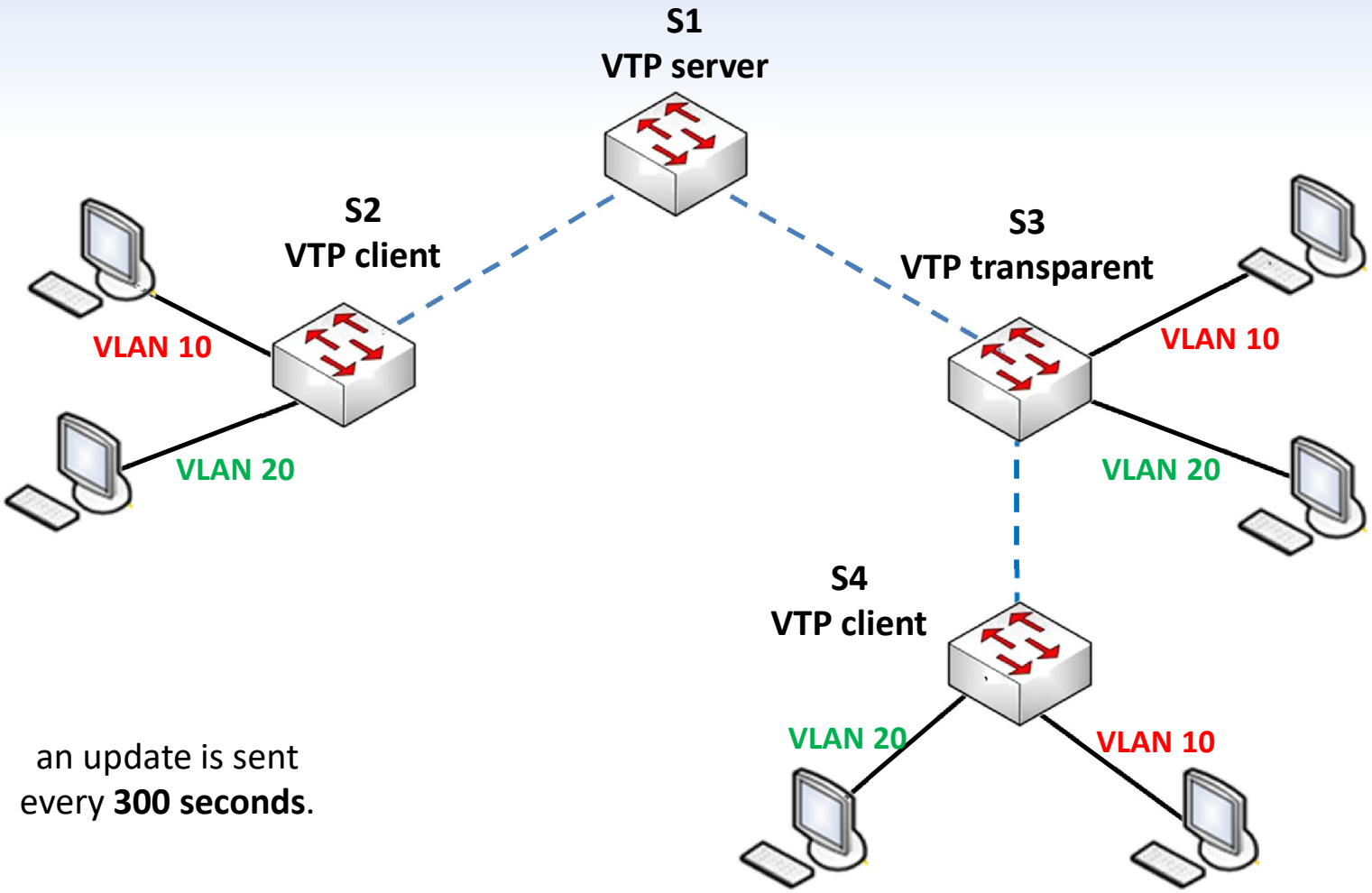
- VTP Server: Switch that creates VTP announcements
- VTP Client: Switch that receives, synchronizes, and propagates VTP announcements
- VTP Transparent: Switch that doesn't process VTP ads

Swtichs must have the same VTP domain name, in order for them to exchange updates. In addition, it is possible to set up a password for the domain.....

Modes VTP

- **Switch in VTP Server mode**
  It allows the administrator to make any changes to the VLANs and automatically propagate his changes to all switches on the network.
- **Switch in VTP Client mode**
  It does not allow the administrator to make changes to VLANs. Only VTP updates can change the configuration.
- **Switch in VTP Transparent mode**
  They allow the administrator to make any changes on VLANs locally only and therefore do not propagate his changes to all switches on the network.....

VTP Architecture



an update is sent
every **300 seconds**.

Types of VTP messages

There are three types of VTP messages, which are:

1. **Client advertisement request**: Generated by a client to request information from a VLAN server, this message is sent when:
   - The switch has been rebooted
   - An incremented revision number is received
   - The domain name has been changed.....

Servers respond with 'Subset advertisement'

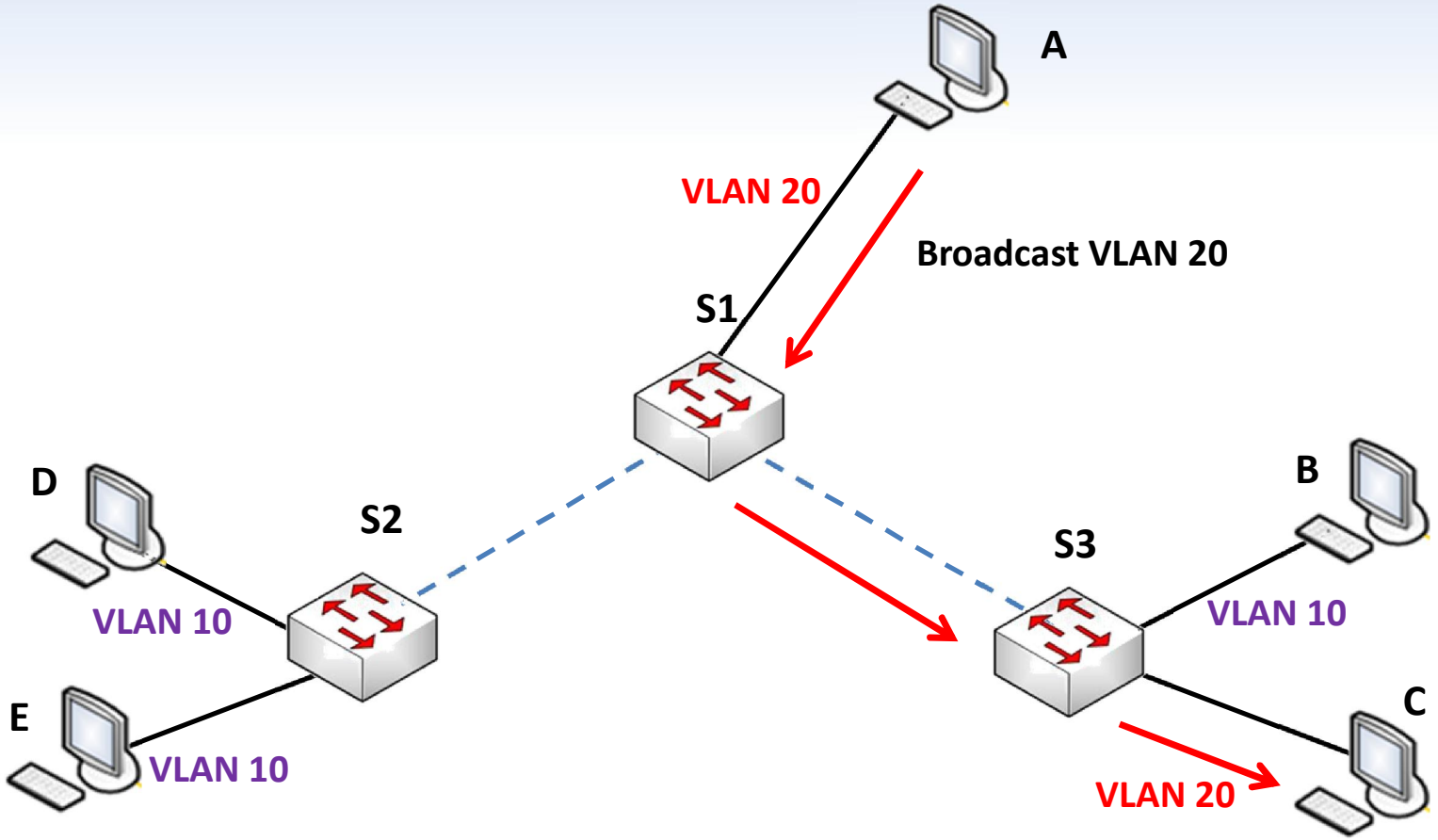
2. **Summary advertisement**: sent every 300 seconds, by default, or when a configuration change occurs. It contains the VTP domain name, as well as the revision number.

3. **Subset advertisement**: are sent when a configuration change occurs on the server switch. They are VLAN-specific and contain details about each VLAN...

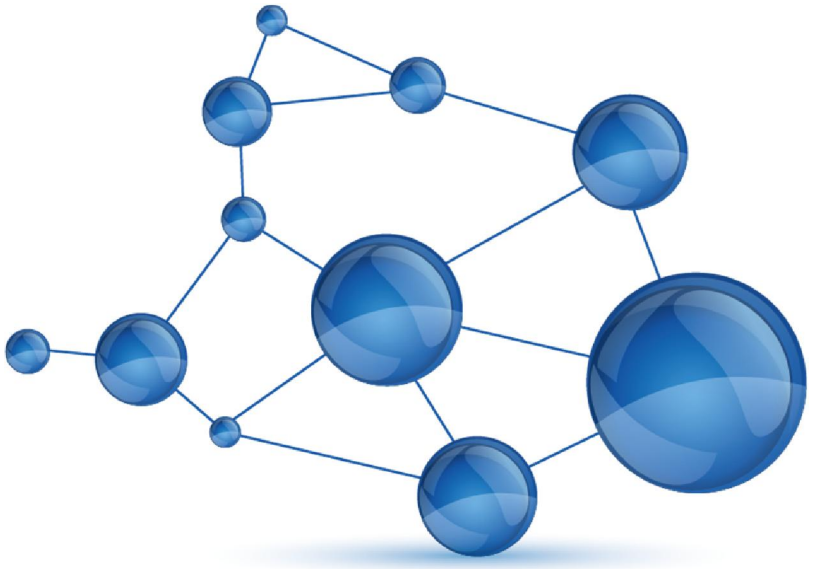
VTP Pruning

It prevents Broadcasts from propagating to switches that do not have Access ports in the VLAN concerned by a Broadcast. This is done by allowing switches to chat with each other, in order to tell which VLANs they are using

Example of VTP Pruning



Dynamic Trunking Procotol



Principle

Configuring the Trunks on all the inter-switch links in the company can be very tiring. Cisco has created a protocol that will automatically mount a Trunk between 2 switches, it's the DTP protocol. **Dynamic Trunking Procotol,**

The principle is very simple, when a port rises, DTP announcements are sent;

- If the port is connected to a neighboring switch, the switch will receive the advertisement
- If the port is connected to a pc, the latter will not respond to the ad because it does not understand the protocol. On the switch port, the Trunk is not activated and therefore remains in Access mode.....

Operation

A physical port on a switch can have multiple states (or modes) regarding DTP. By default, the ports on a switch are in Dynamic Auto mode.

Mode	Fonction
Dynamic Desirable	Annonce sa volonté de monter en trunk (négociation)
Dynamic Auto	Attends une sollicitation du voisin. Il n'envoie pas de requêtes mais répond aux requêtes d'en face
Trunk (on)	Le switch se met en mode trunk automatiquement et en informe le switch voisin
Nonegotiate	Le switch se met en mode trunk automatiquement sans en informer le switch voisin
Off	Désactivation du Trunk
Access	Désactivation du Trunk et prévient le voisin

Operation

	Dynamic Auto	Dynamic Desirable	Trunk	Access
Dynamic Auto	Access	Trunk	Trunk	Access
Dynamic Desirable	Trunk	Trunk	Trunk	Access
Trunk	Trunk	Trunk	Trunk	?
Access	Access	Access	?	Access

Dynamic Desirable	Annonce sa volonté de monter en trunk (négociation)
Dynamic Auto	Attends une sollicitation du voisin. Il n'envoie pas de requêtes mais répond aux requêtes d'en face
Trunk (on)	Le switch se met en mode trunk automatiquement et en informe le switch voisin
Nonegotiate	Le switch se met en mode trunk automatiquement sans en informer le switch voisin
Off	Désactivation du Trunk
Access	Désactivation du Trunk et prévient le voisin

To remember

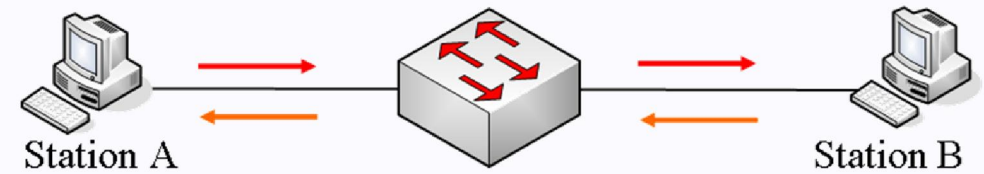
- One VLAN = one subnet
- Multiple VLANs, i.e. multiple subnets, on a single switch
- The same VLAN on multiple switches
- Routing between VLANs using a router (or Layer 3 switch)
- Cost reduction
- Greater flexibility.....

Chapter 03

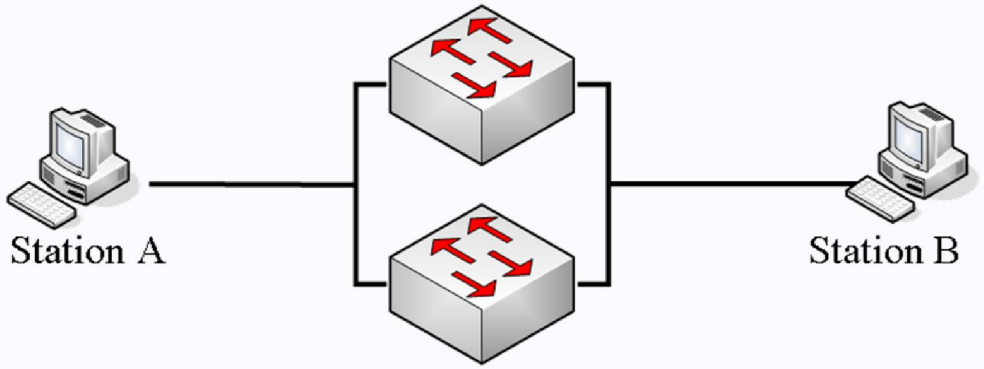
Redundancies on links

Notion of redundancy

To ensure the reliability of computer links, it is useful to duplicate the interconnection equipment so that in the event of a failure of one of them, the other equipment takes over; This is called redundancy.



not redundant Architecture



Redundant Architecture

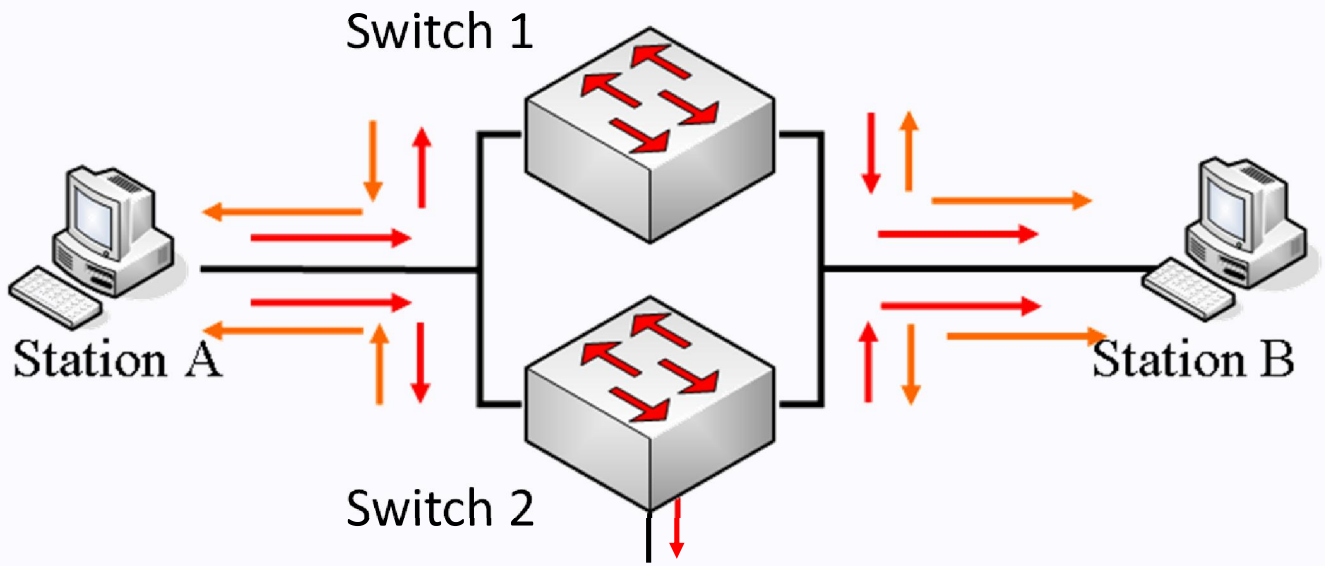
Redundancy issues

If the switches forward broadcast and multicast traffic through all ports except the original port, and if the Ethernet frames do not have a TTL, then several problems can occur:

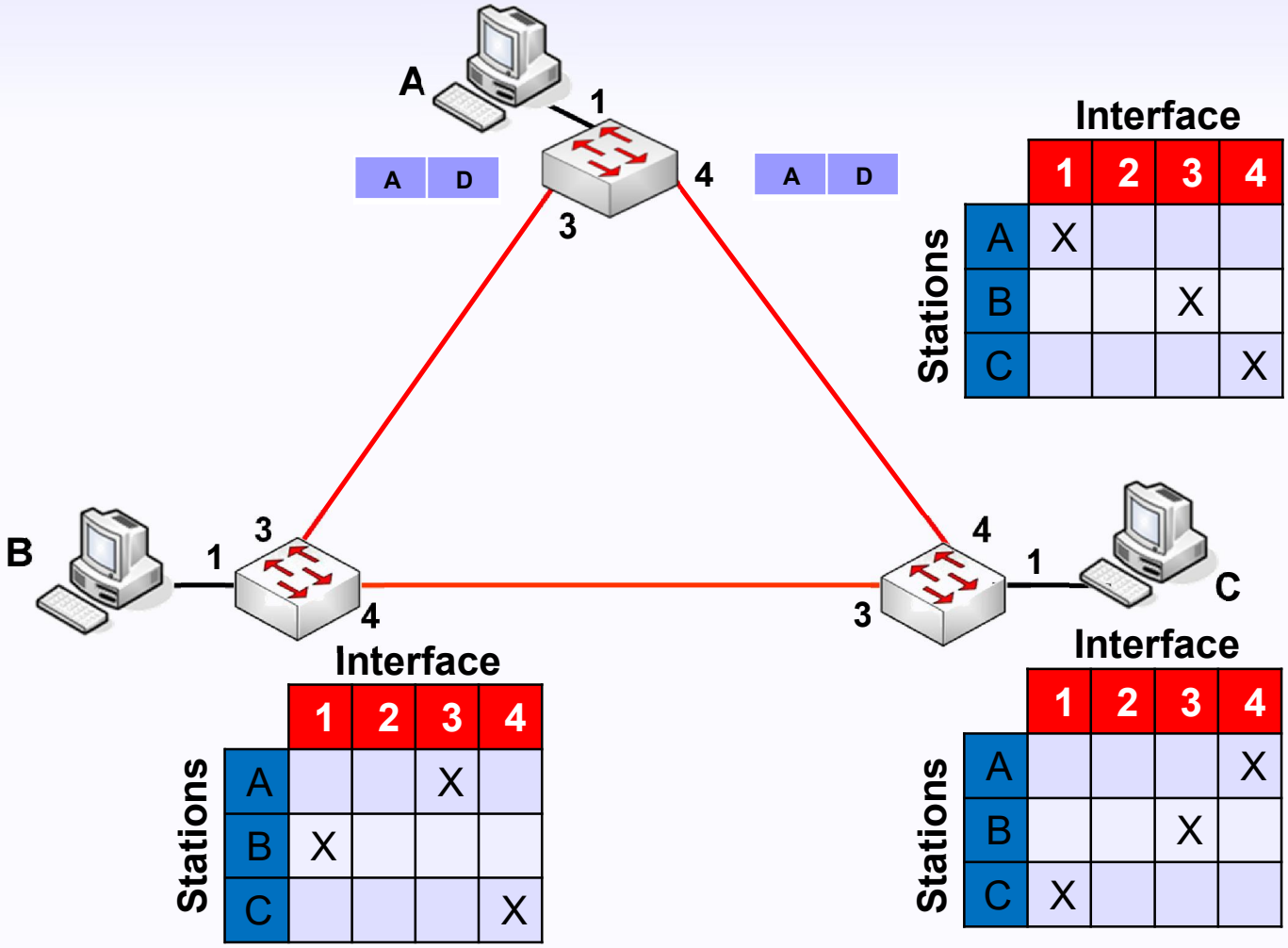
- Broadcast Storm,
- duplicate frames,
- an instability of the switching tables....

Problem 1: Broadcast storm

If station A sends a Broadcast frame (FFFF. FFFF. FFFF), the Switch 1 extracts the destination MAC address and duplicates it on all its ports. same for the Switch 2. These frames run non-stop (Broadcast Storm).



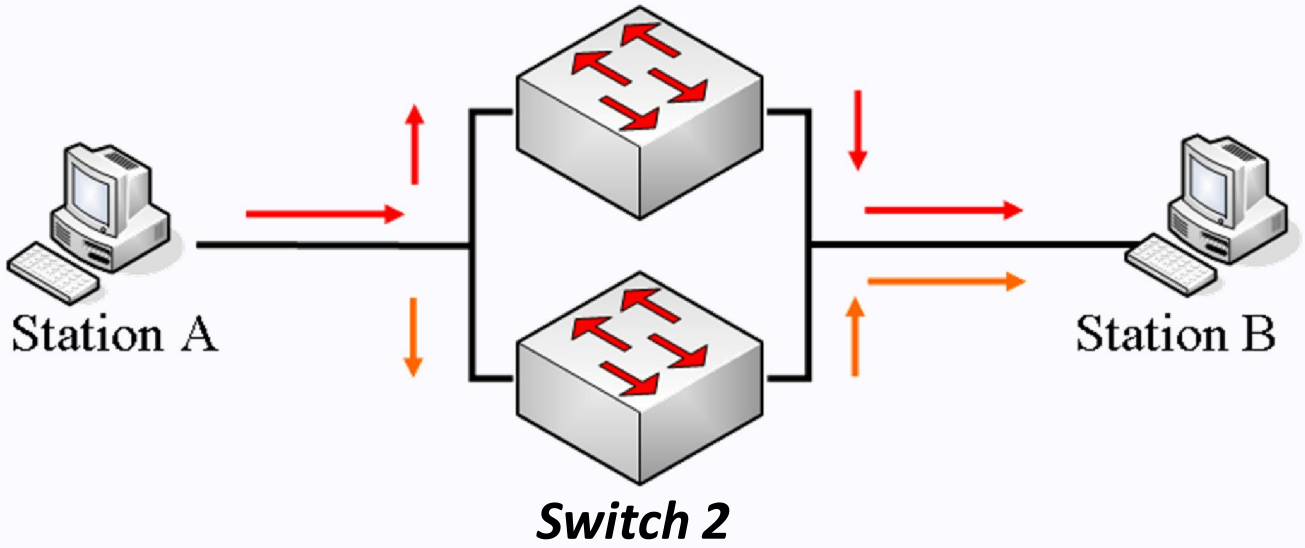
Example


A
1
A D
4
3
A D
Interface
1 2 3 4
Stations
A X
B X
C X
B
1
3
4
3
4
1
C
Interface
1 2 3 4
Stations
A X
B X
C X
Interface
1 2 3 4
Stations
A X
B X
C X

Problem 2: Frame Duplication

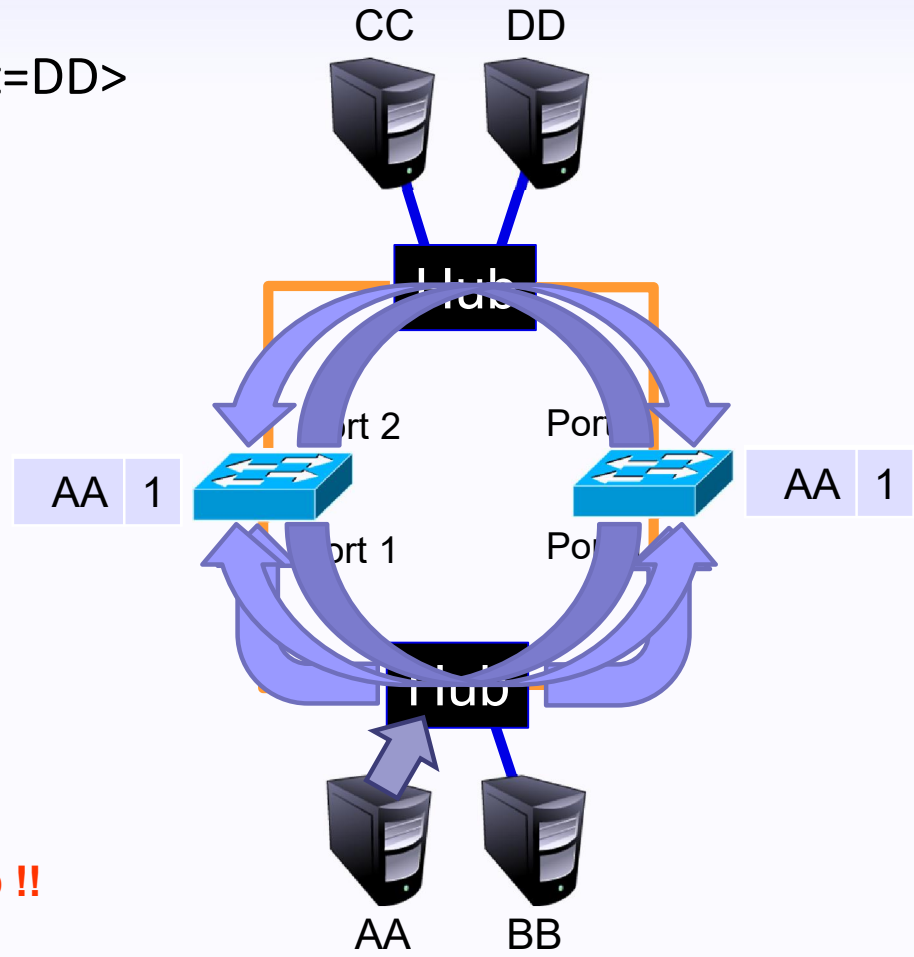
Captivating:
- ✓ Source MAC Address: A
- ✓ MAC Address Destination: B

The Switch 1 receives the frame and switches it to the the port on the right. Station B does receive the frame of Station A. But the Switch 2 also receives the frame and switches it to the port on the right. Station B therefore receives the frame of Station A for the second time!!



Problem 3: Switch Table Instability

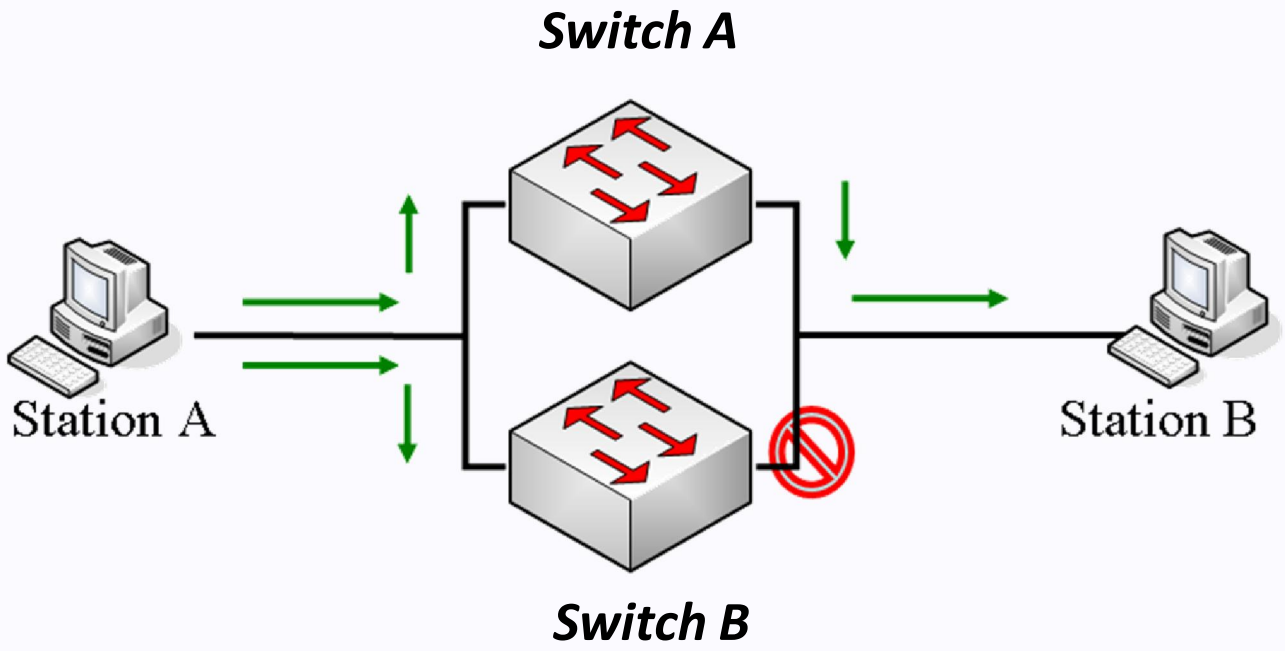
<Src=AA, Dest=DD>



Infinite Loop !!

Solution !

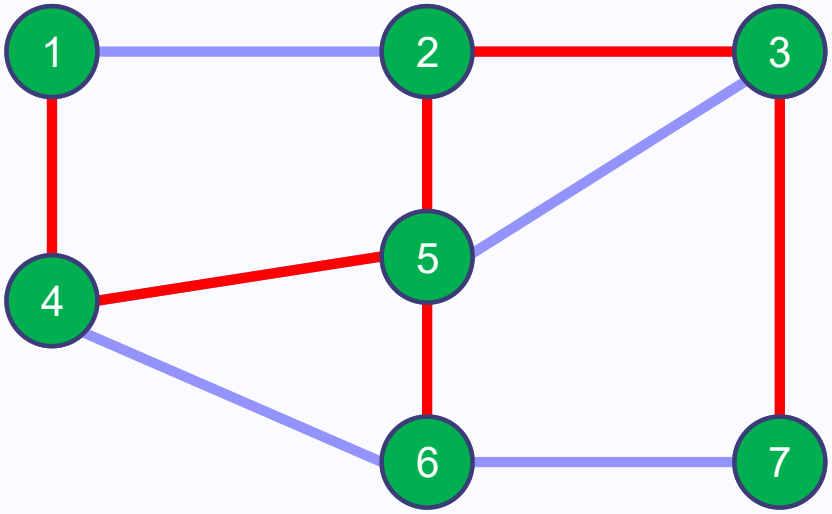
To avoid the 3 aforementioned problems, the STP protocol was created. It makes it possible to identify loops and block them "logically"
All traffic will go through Switch A, with the bottom path blocked at the port of Switch B.
If Switch A fails, the STP will detect it and will unblock the bottom port.

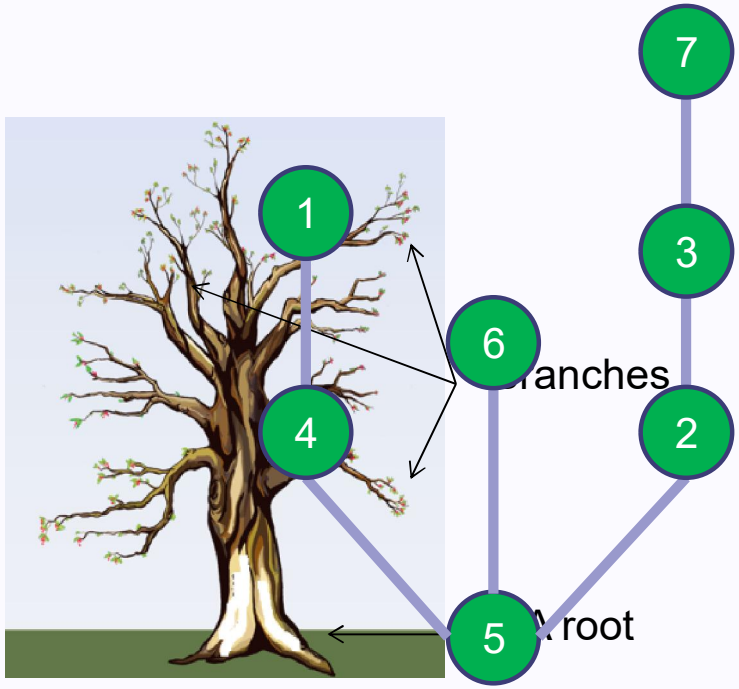


Tree Configuration

Ensures that:

- Scope on all nodes
- Creates no loops
- This structure is a tree





A redundant physical topology while creating a single logical path.

Spanning Tree Protocol

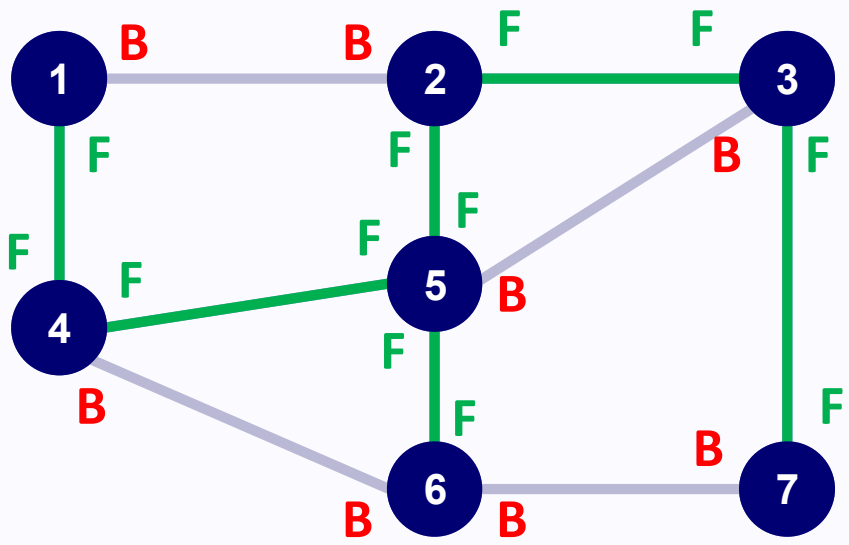
Originally developed by DEC and standardized by the IEEE (IEEE 802.1D), the Spanning Tree Protocol (STP) is a network topology learning protocol whose purpose is:

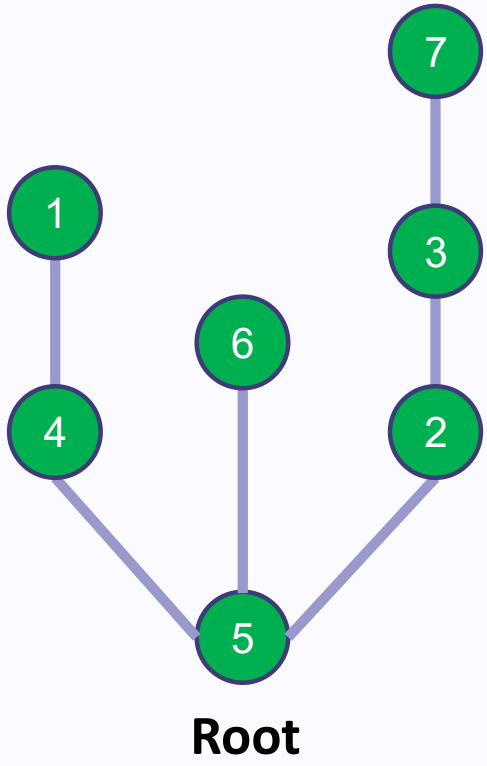
- eliminate loops by disabling the ports that generate these loops;
- Continuously monitor the availability of active ports.
- and, in the event of an active bridge failure, to fail over traffic to the dormant port.

IEEE 802.1 : Management of local networks, VLANs, authentication, etc.

How !!!

F: forwarding
B: blocking





The implementation of the STP algorithm

The configuration and implementation of the STP algorithm goes through 4 steps, which are:

Step 1: Determine the **Root Switch**

Step 2: Determine the **RootPort** on Other Switches

Step 3: Determine the **DesignatedPort** on each segment

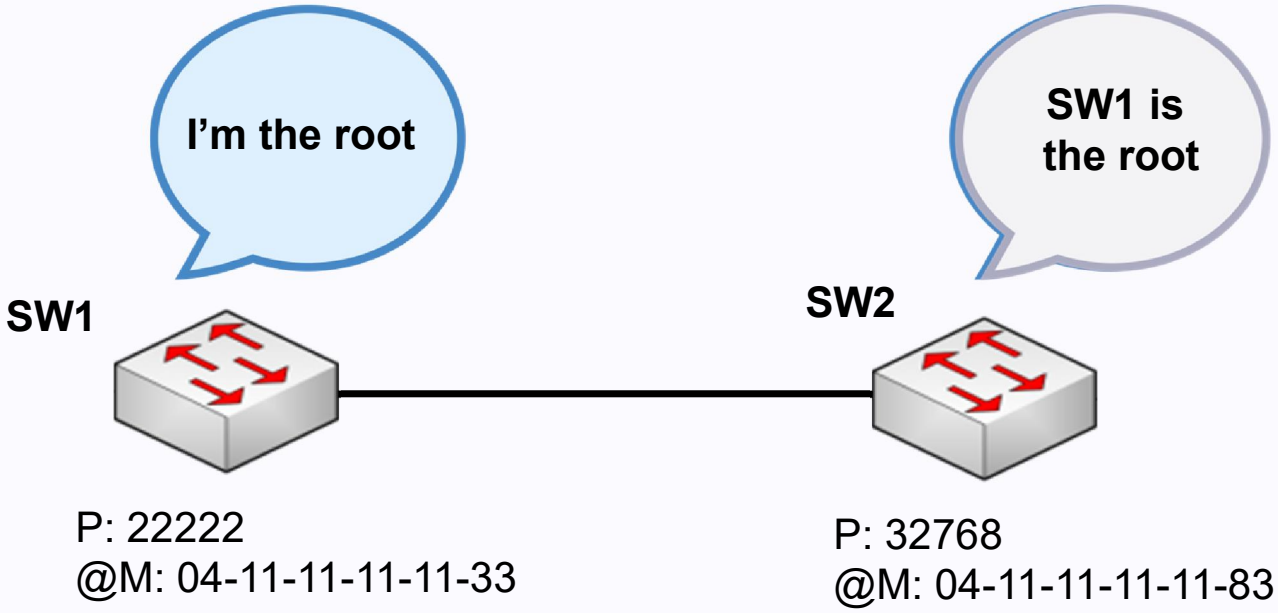
Step 4: Block Other remaining Ports

Step 1: Determine the Root

- ✓ Election: each Switch pretends that it is root and announces its **identity** (ID, Switch IDentifier) (BID, Bridge ID) to all other neighboring Switches by a configuration message called **BPDU** (Bridge Protocol Data Unit),

Identity : priority-MACaddress

- ✓ If a Switch X receives a BPDU from a Switch Y that advertises a **lower** (**better**) identity than the one it has, then X announces Y as root in its next BPDUs,
- ✓ The Switch with the lowest identity becomes root, with equal priority, it is the one with the smallest MAC address....


I’m the root
SW1 is the root
SW1
SW2
P: 22222
@M: 04-11-11-11-11-33
P: 32768
@M: 04-11-11-11-11-83

Step 1: Determine the Root (2)

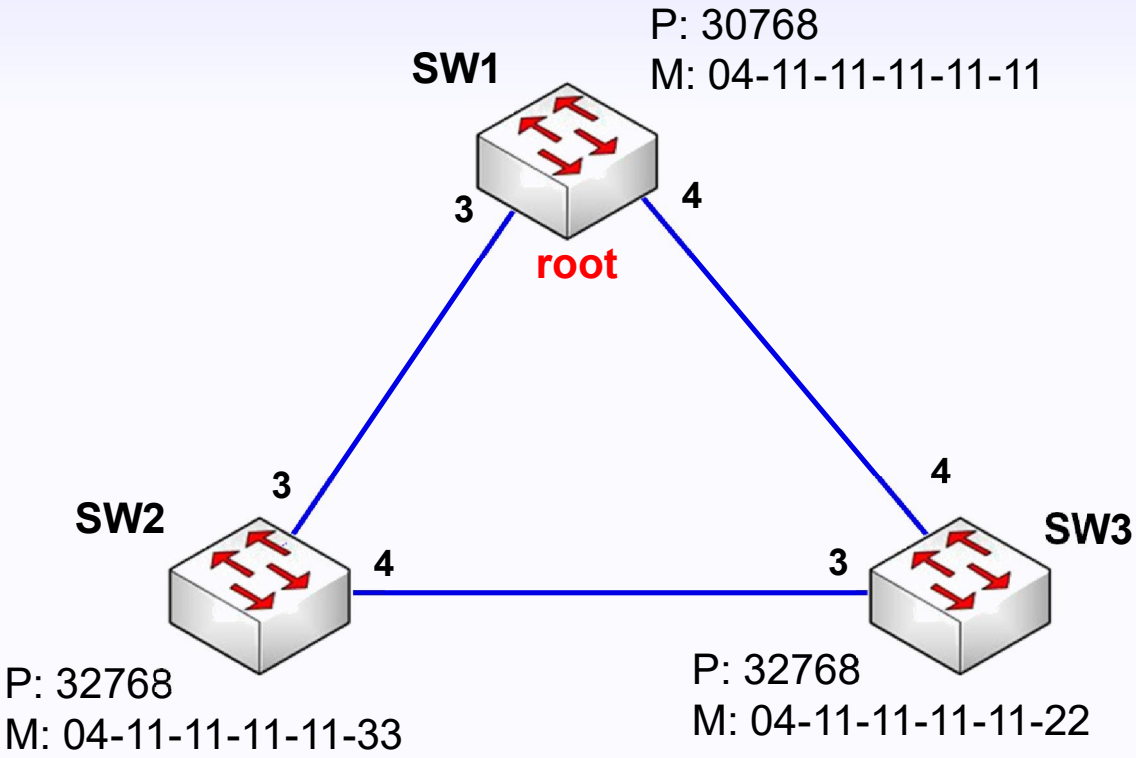
The priority field is set to 32768 by default. Its value can be changed by the administrator. The lower the number, the higher the priority.

Example:

Id1 < Id2 si: (Priority1 < priority 2)
Ou : (Priority1 = Priority2) and (MAC1<MAC2)

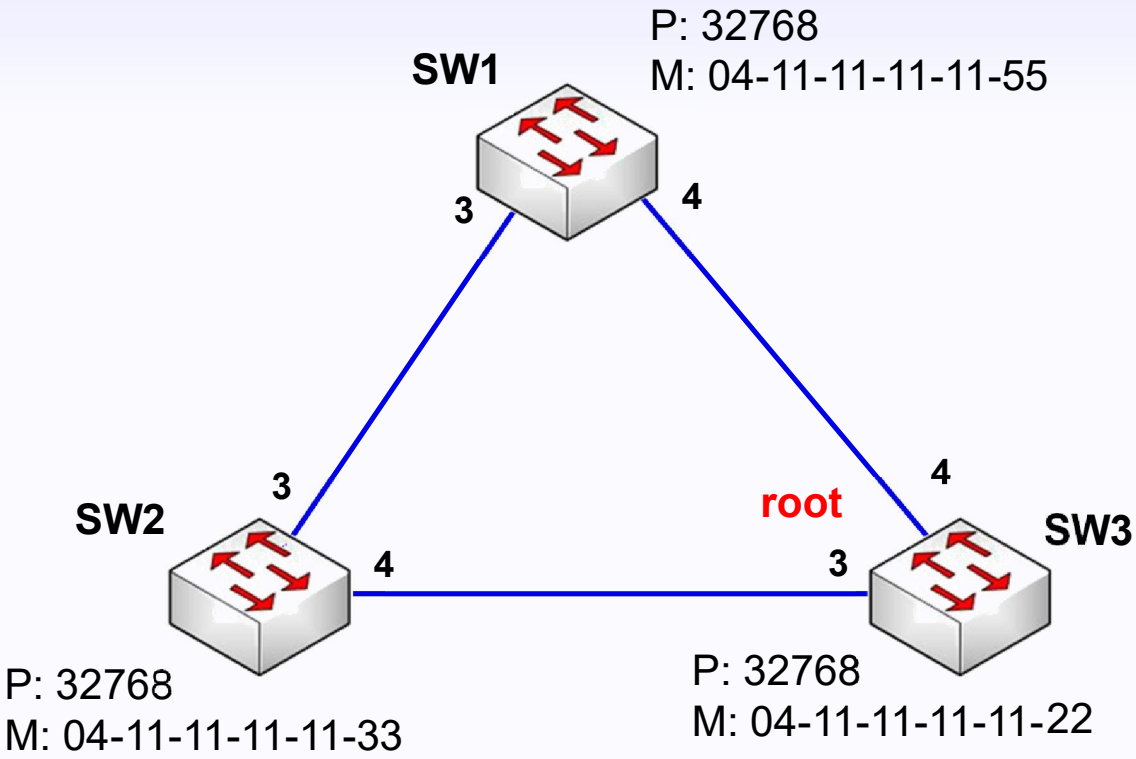
Periodically (recommended value 20 s) a **diffusion frame** is emitted. If a backup bridge remains more than this time interval without receiving anything, it deduces that the bridge, of which it is the backup, is failing. It then issues a configuration frame.

Example 1:



Identity : priority-MACaddress

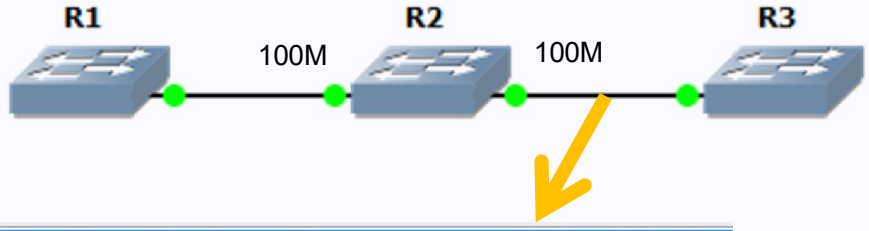
Example 2:



Identity : priority-MACaddress

Concept of cost

BPDU cc:00:0b:28:00:00 cc:01:0b:28:00:00 cc:02:0b:28:00:00

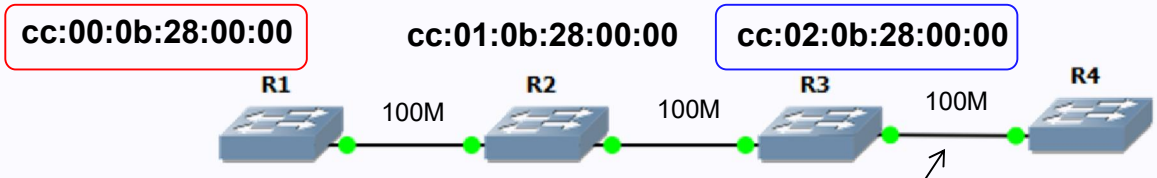



```
Frame 9: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
IEEE 802.3 Ethernet
Logical-Link Control
Spanning Tree Protocol
    Protocol Identifier: Spanning Tree Protocol (0x0000)
    Protocol Version Identifier: Spanning Tree (0)
    BPDU Type: Configuration (0x00)
  + BPDU flags: 0x00
  + Root Identifier: 32768 / 0 / cc:00:0b:28:00:00
    Root Path Cost: 19
  + Bridge Identifier: 32768 / 0 / cc:01:0b:28:00:00
    Port identifier: 0x8038
    Message Age: 1
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
```

Cost	Speed	range allowed
100	10M	50 - 600
19	100M	10 - 60
4	1G	3 -10
2	10G	1 -5

Step 2: Determine the RootPort

- ✓ Each switch (other than root) has a single RootPort (RP).
- ✓ RP has the lowest cumulative cost compared to the root.

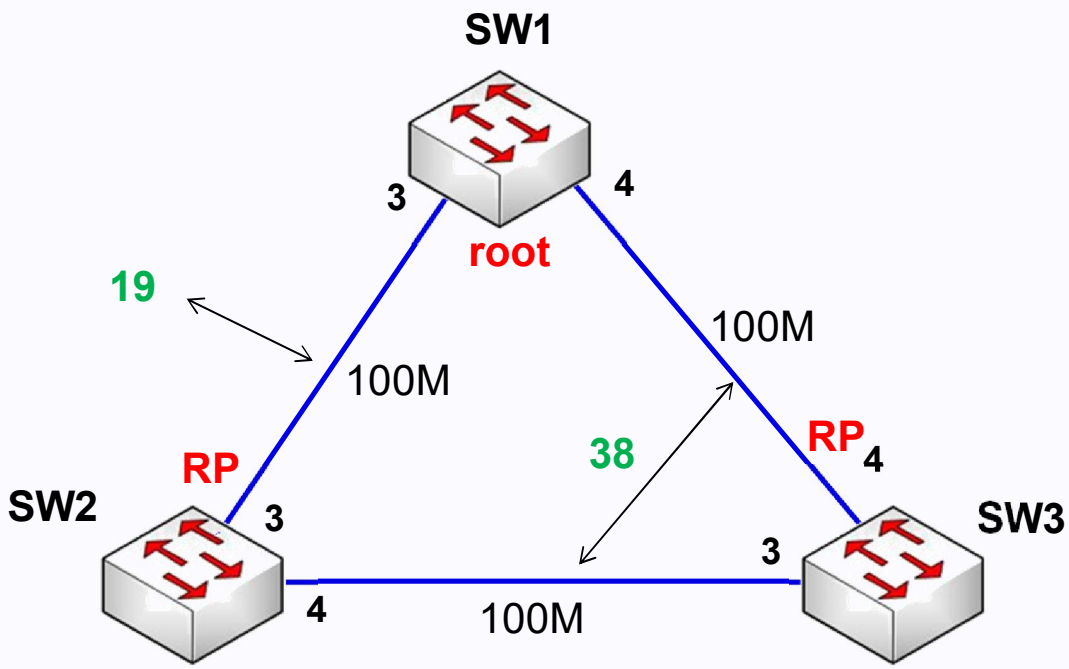



```
Frame 60: 60 bytes on wire (480 bits), 60 bytes captured
IEEE 802.3 Ethernet
Logical-Link Control
Spanning Tree Protocol
    Protocol Identifier: Spanning Tree Protocol (0x0000)
    Protocol Version Identifier: Spanning Tree (0)
    BPDU Type: Configuration (0x00)
  BPDU flags: 0x00
  Root Identifier: 32768 / 0 / cc:00:0b:28:00:00
    Root Path Cost: 38
  Bridge Identifier: 32768 / 0 / cc:02:0b:28:00:00
    Port identifier: 0x8038
    Message Age: 2
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
```

Speed	Cost
10M	100
100M	19
1G	4
10G	2

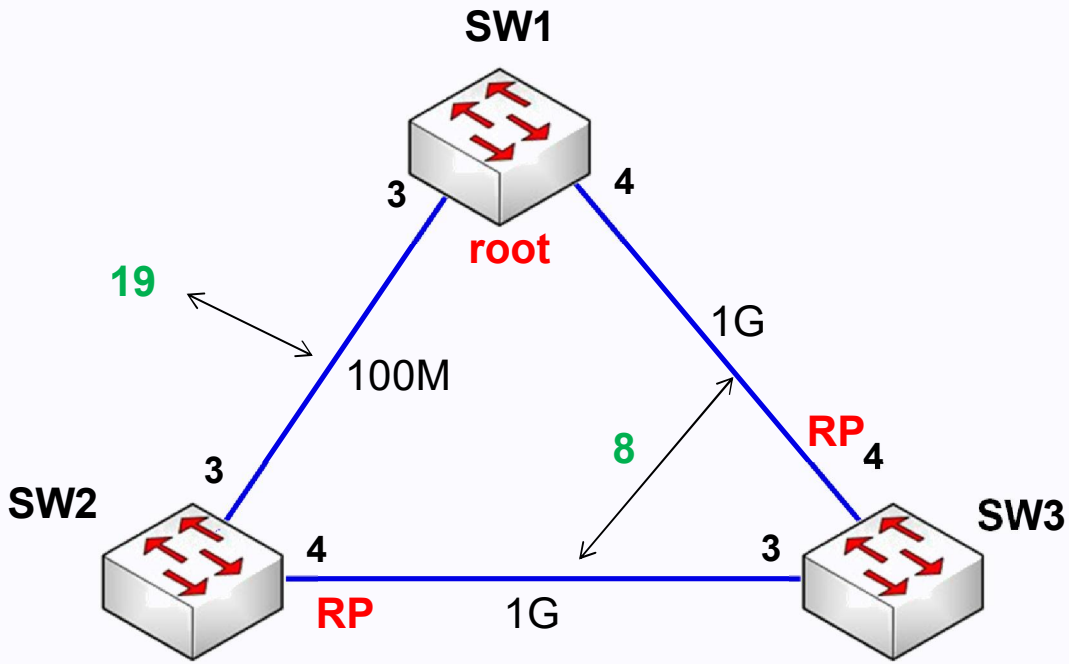
Example 1:

Speed	Cost
10M	100
100M	19
1G	4
10G	2



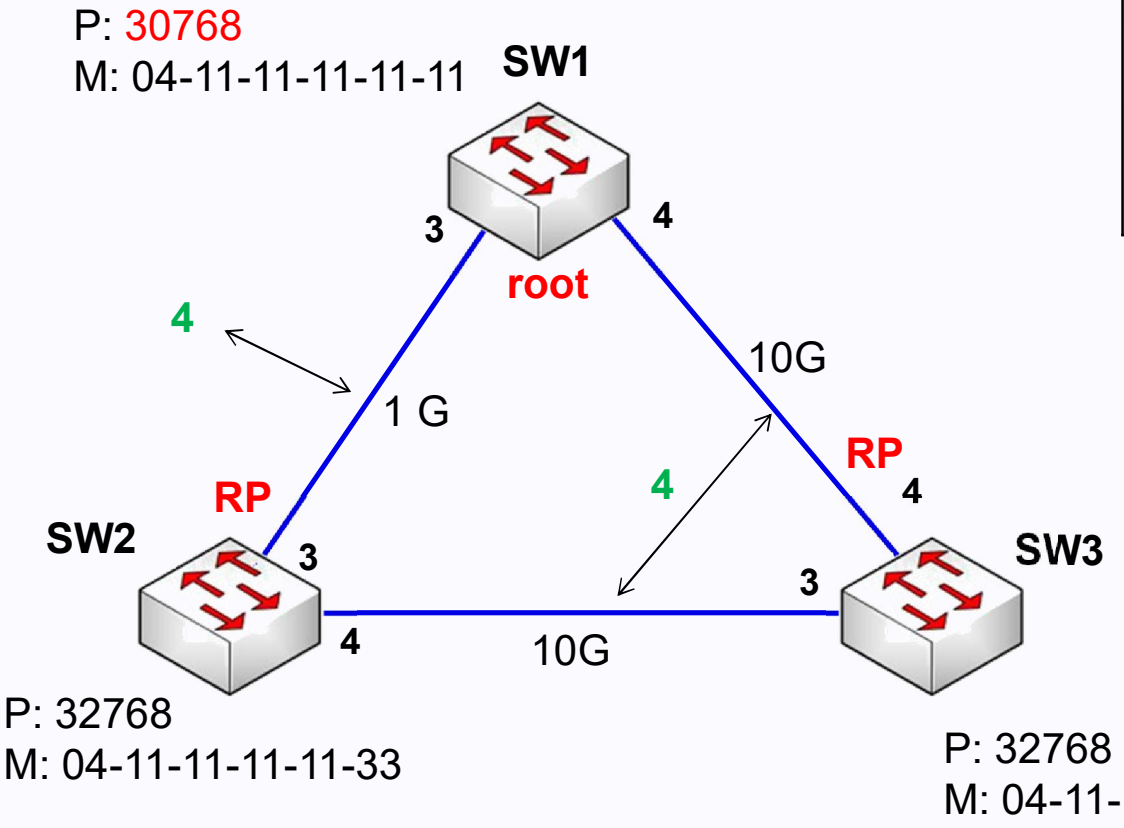
Example 2:

Speed	Cost
10M	100
100M	19
1G	4
10G	2

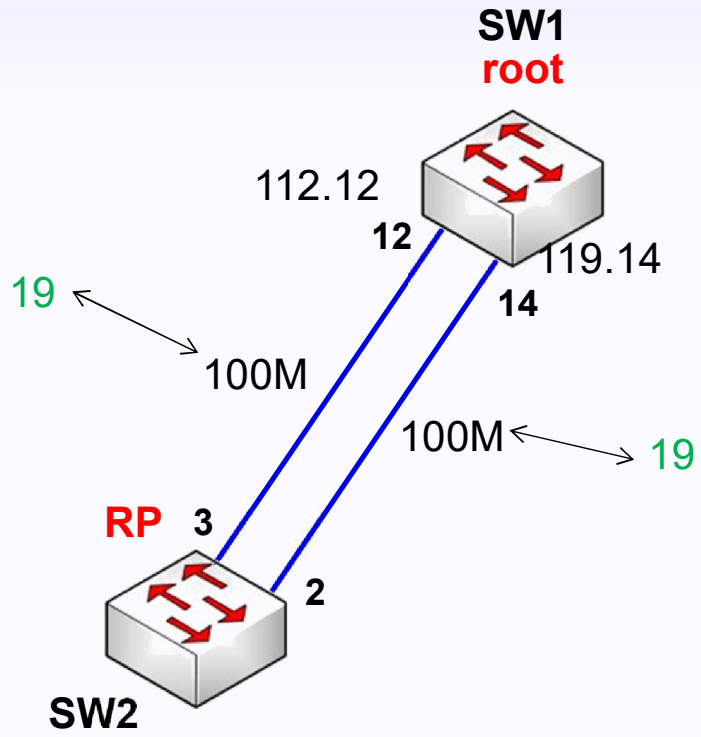


Example 3:

Speed	Cost
10M	100
100M	19
1G	4
10G	2

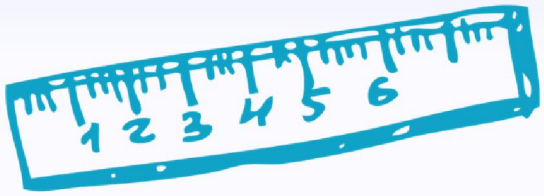


Example 4:



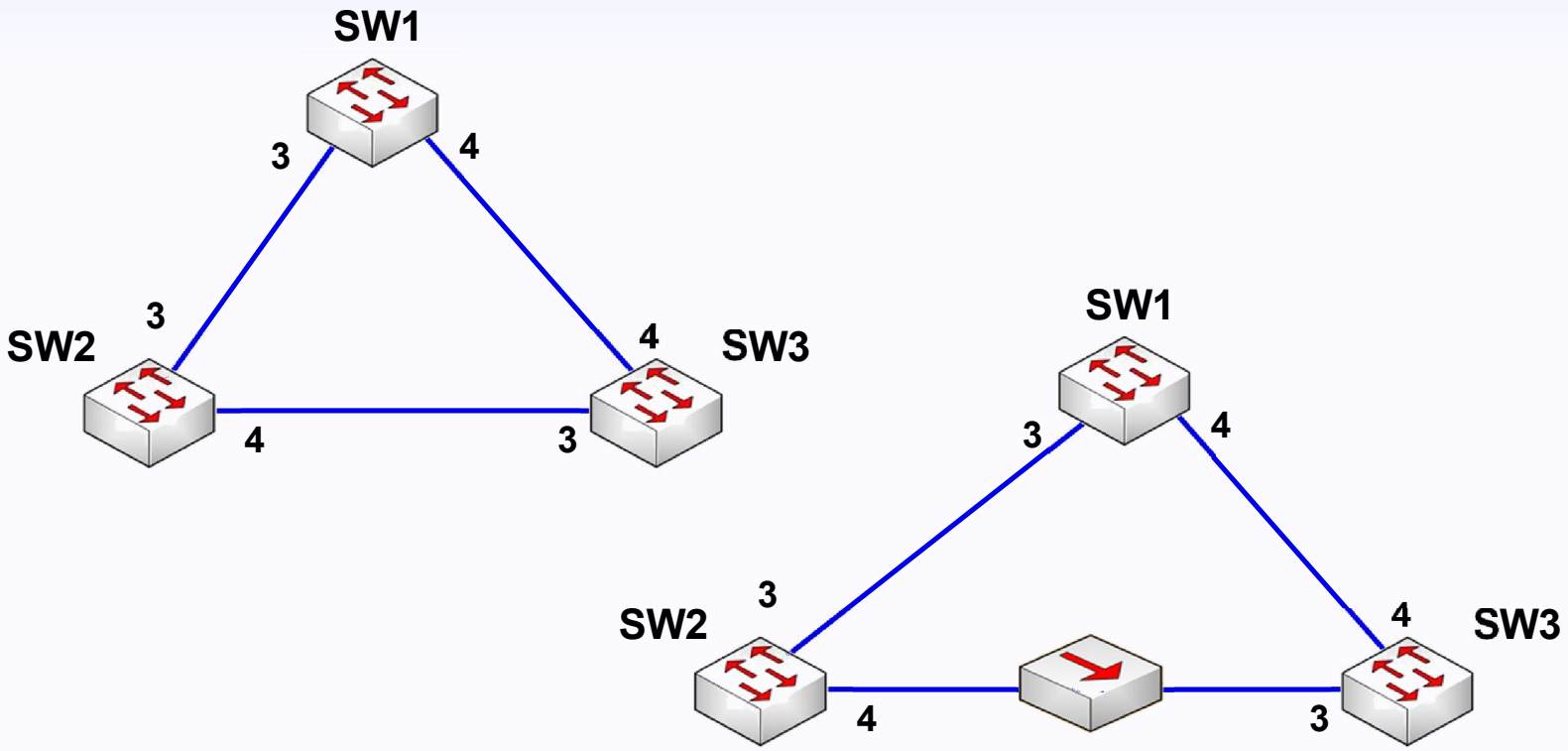
Speed	Cost
10M	100
100M	19
1G	4
10G	2

Port-ID : 112.12 (112 is the priority, 12 is the port number on the Switch).



All **Rootports** are in Forwarding mode

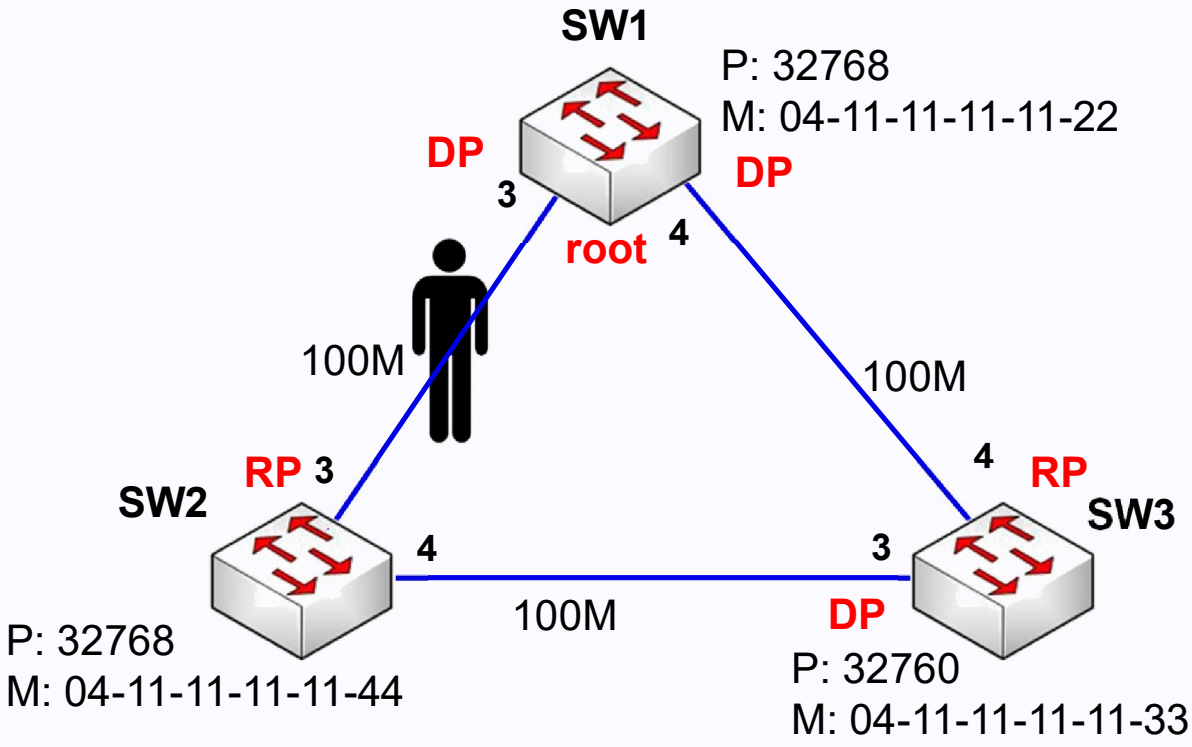
Concept of segment

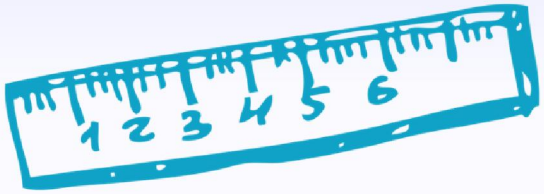


Physical Segment, Collision Domain

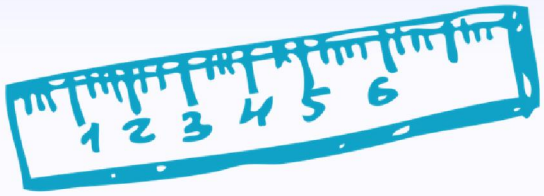
Step 3: Determine the DesignatedPort

- ✓ In each segment there is only one Designated port.
- ✓ The Designated port is the one that offers the least cost to the root.



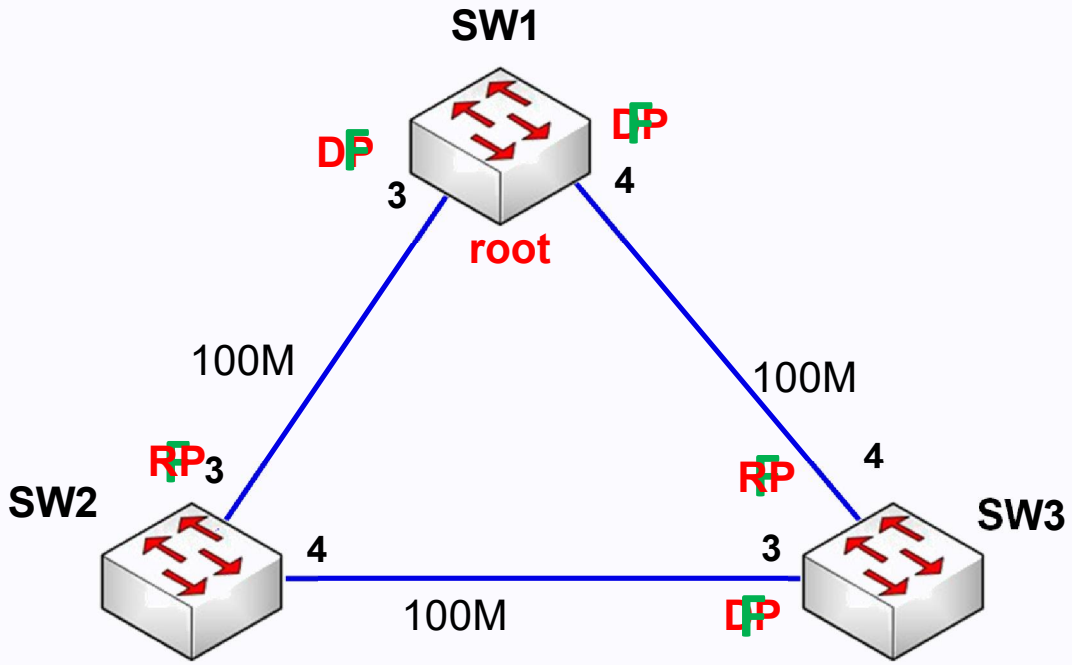


All **root ports** are **DesignatedPort**

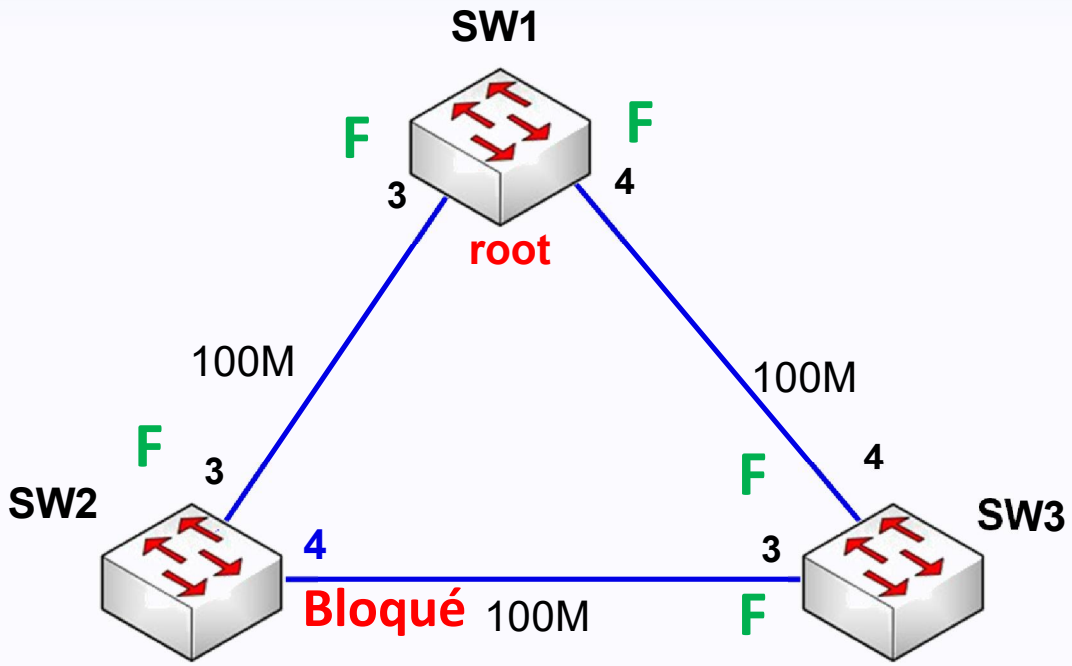


All **Designatedports** are in mode: **Forwarding**

Application

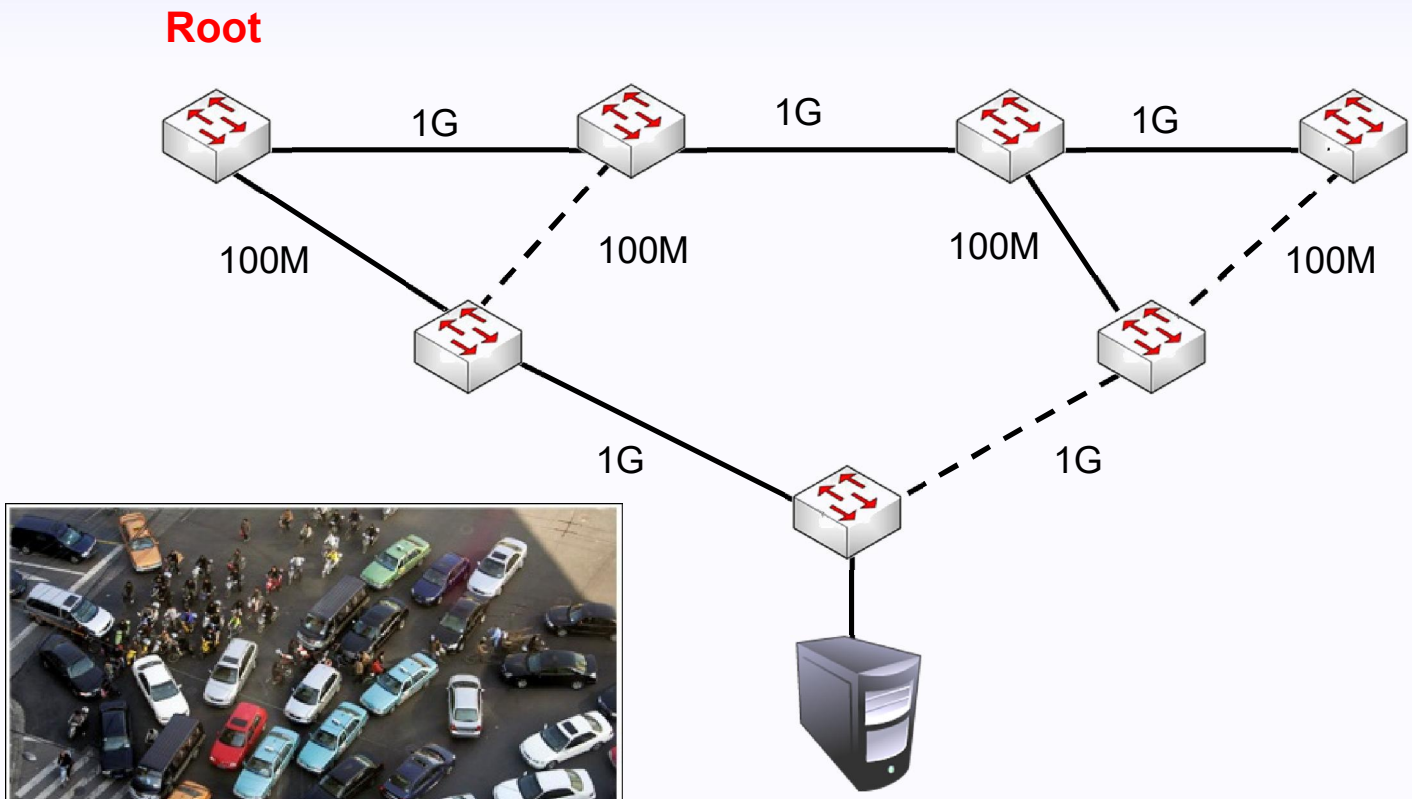


Step 4: Block Other remaining Ports

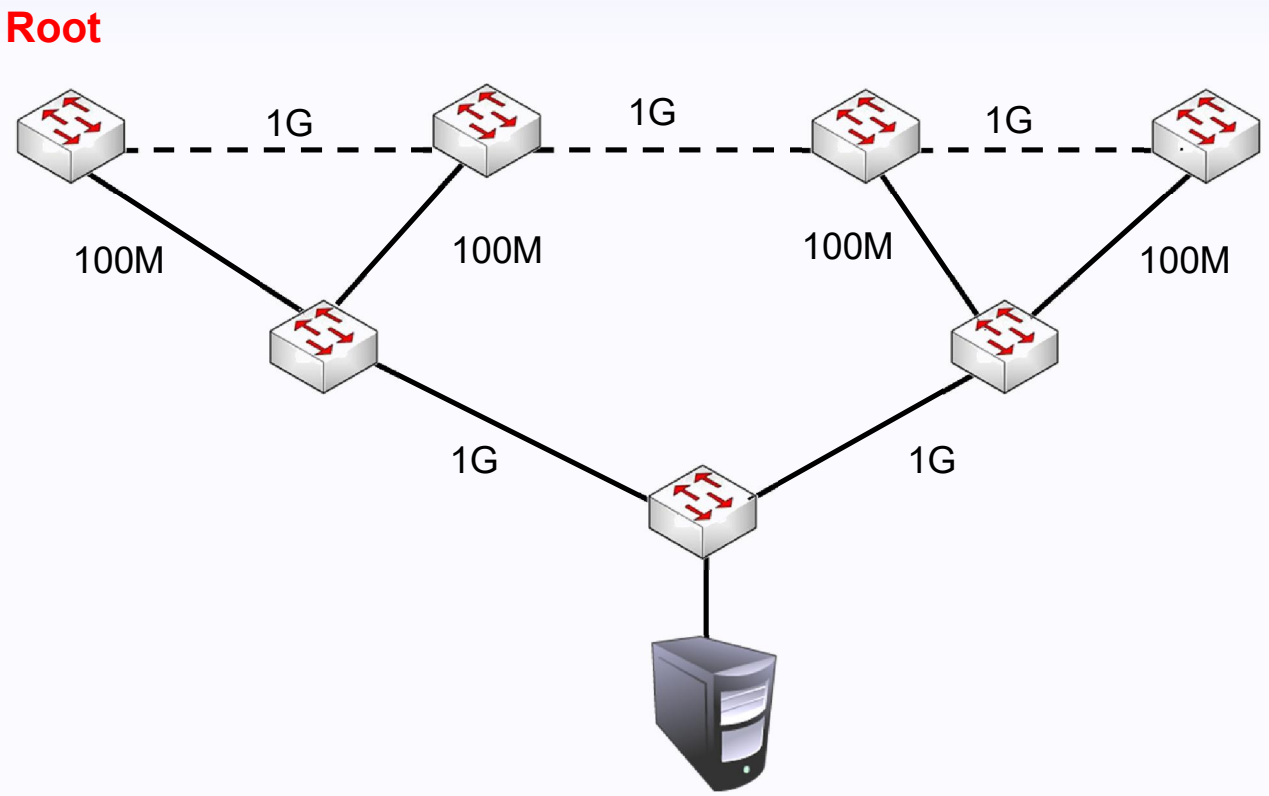


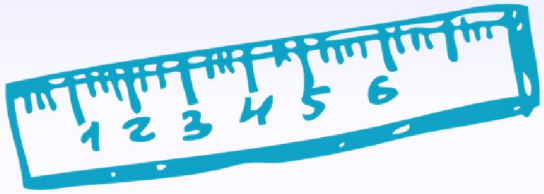
The **Root ports** and **Designated ports** forward traffic ("Forwarding" state) and the other ports cut the link ("Blocking" state).

Possible case



Impose the root switch





For the **Designatedports** we compare the ports of a segment.
For the **RootPort**, we compare the ports of the same Switch

Summary

- ✓ 1 Root switch per network with all ports Designated (Forwarding)
- ✓ 1 Root (Forwarding) port per Non-Root switch
- ✓ 1 Designated (Forwarding) port per collision domain (link)
- ✓ all other ports are Non-Designated (Blocking)

Spanning-Tree States

States	Time	Transfert data	MAC Learning	Sending BPDUs	Listening to the BPDUs
Listening	Forwarding Delay = **15** s	*non*	*non*	Yes	Yes
Learning	Forwarding Delay = **15** s	*non*	Yes	Yes	Yes
Forwarding	-	Yes	Yes	Yes	Yes
Blocking	Max Age = **20** sec.	*non*	*non*	*non*	Yes

- "Max-Age", from 6 to 200 seconds, 20 seconds by default.
- Forward-time, from 4 to 200 seconds, 15 seconds by default.

Types of BPDU

There are 3 types of BPDUs:

- Configuration BPDUs, used for Spanning Tree calculation: contain the ID of the switch that sends the message, the port ID, the cost of the link.
- Topology Change Notification BPDU (TCN), used when the topology changes. Then, the Root will send a configuration BPDU to everyone.
- Change Acknowledgement BPDU: is done in response to the Change BPDU

Limits and alternatives

Disadvantage of 802.1d STP:

- Slow Convergence Speed.
- The STP (802.1d) typically takes 50 seconds to converge.

Solution:

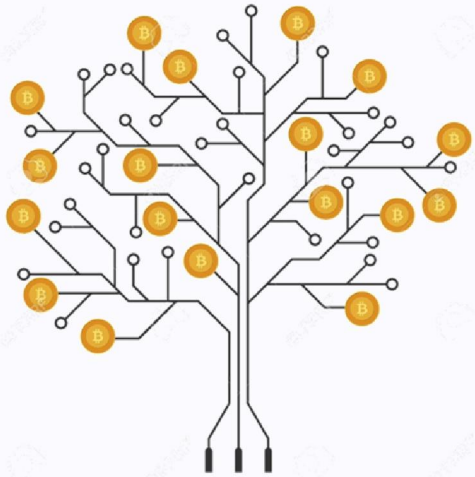
IEEE 802.1w: Rapid Spanning Tree Protocol, RSTP.

The **R**apid **S**panning **T**ree

802.1W

Rapid Spanning Tree Protocol, RSTP

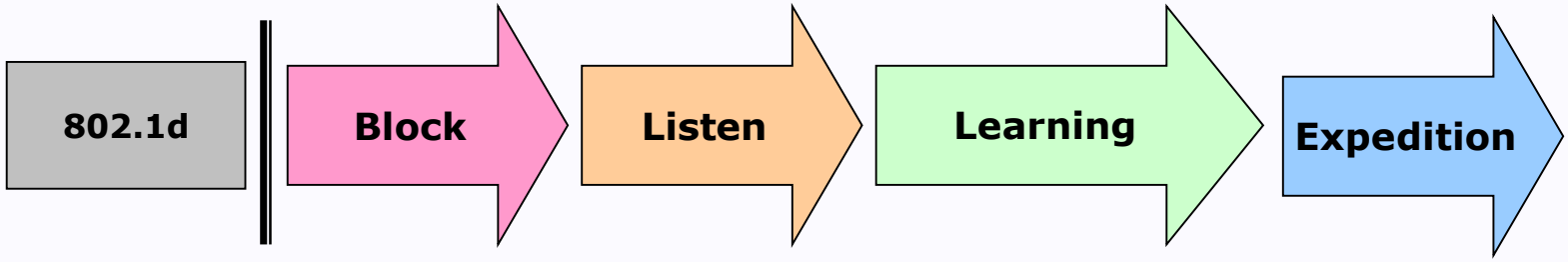
- Standardized 802.1W by the IEEE.
- Provides significant convergence speed improvements for the mesh network by immediately swapping root and designated ports in the transmission state.



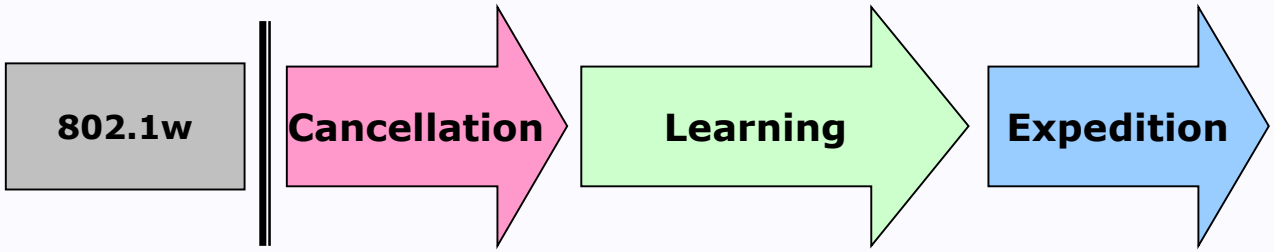
How RSTP works ?

Port States:

- In 802.1d, 4 different port states are defined:
- blocking, listening, learning and shipping.



- In 802.1W, 3 different port states are defined:
- cancellation, learning and shipping.



Port Roles

In Rapid Spanning Tree, ports can have these roles:

- Root Port: The port that provides the best path to the Root Bridge
- Designated Port: you need one and only one port designated per link
- Alternate Port: port blocked by Spanning Tree, but which can very quickly switch to Forwarding in case of failure
- Edge Port: is not connected to a switch (equivalent to Portfast)

Differences between RSTP and STP

Differences from STP:

There are now only three states for RSTP ports:

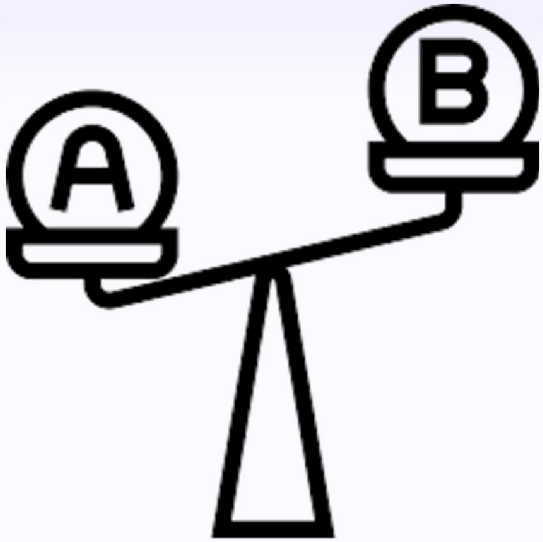
- Discarding (instead of Disabled, Blocking, and Listening)
- Learning and Forwarding (keeping the same function)

The Port Root and Port Designated roles remain. The best alternative ports are called the backup link of the latter: Alternate port and Backup port. They take on the role of Root port and Designated port in case of failure.

How RSTP Works (2)

- Convergence time:
  - STP, 802.1d: 50 sec.
  - RSTP, 802.1w: 2-3 sec.

- Maximum number of loops:
  - STP, 802.1d: 7 Loops
  - RSTP, 802.1w: 18 Loops



802.1w is "backwards compatible" with 802.1d. However, the benefit of rapid convergence will be lost....

Summary on the RSTP:

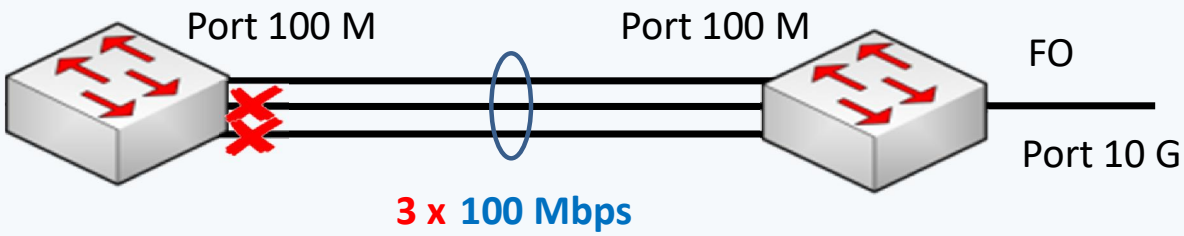
- There can only be one STP "parameterization" in a network (i.e. a tree).
- If VLANs are parameterized, all VLANs will share the RSTP. This means that all VLANs will have the same logical topology, hence poor flexibility.
- Solution: Multiple Spanning Tree Protocol, MSTP (IEEE 802.1s)

Chapter 04

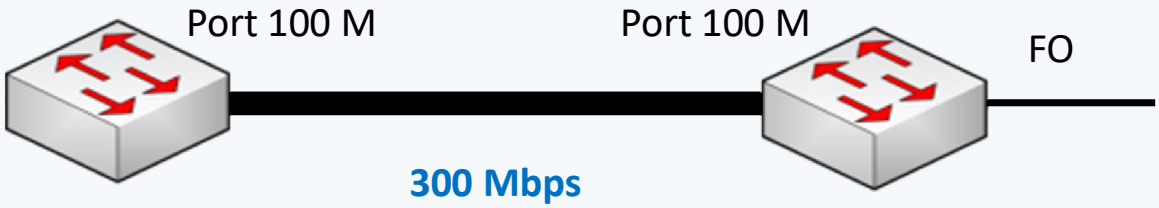
Link aggregation (Etherchannel)

What is it for?

Physically



Logically



I WANT A LOOP IN THE NETWORK

Definition

EtherChannel is a **link aggregation** technology that allows **multiple** identical Ethernet physical links to be assembled into **a single** logical link. It is also called bonding, LAG, etherchannel, or portchannel.

The goal is:

- increase speed (bandwidth) and
- fault tolerance between switches, routers, and servers (redundancy)....

Ethernet Generations

Ethernet bandwidths have historically increased tenfold each generation: 10 Mbps, 100 Mbps, 1000 Mbps, 10,000 Mbps. If one started to fall on the bandwidth caps, the only option was to upgrade to the next generation which could be prohibitively expensive.

Ethernet Technology	Segment Bandwidth
Ethernet	10 Mbps
Fast Ethernet	100 Mbps
Gigabit Ethernet	1 Gbps
10-Gigabit Ethernet	10 Gbps
40-Gigabit Ethernet	40 Gbps
100-Gigabit Ethernet	100 Gbps

Characteristics

An EtherChannel link groups 2 to 8 active links of:

- 100 Mbps, (Fast EtherChannel, FEC),
- 1 Gbps, (Gigabit EtherChannel, GEC),
- 10 Gbps, (10-Gigabit Etherchannel, 10GEC)
- 100 Gbps, (100-Gigabit Etherchannel, 100GEC)

plus possibly 1 to 8 inactive links in reserve that become active when active links are cut.

Is primarily used on the backbone of the local network, between the Access and Distribution layers.

A bit of history

- EtherChannel technology was invented by the Kal-pana company in the early 1990s.
- This company was later acquired by Cisco Systems in 1994.
- In 2000, the IEEE published the 802.3ad standard, which is an open version of EtherChannel.
- In 2008, it therefore transferred it from the 802.3 Ethernet group to the 802.1 group of standards, initially unchanged, except for its name. It is now called 802.1AX....

The advantages

Link aggregation offers the following advantages:

- Increased bandwidth. Aggregated hard links provide higher bandwidth.
- Profitability. A physical network upgrade can be expensive, especially if it requires new cables
- Reliability and availability. If one of the LAG's hard links goes down, the traffic is dynamically and transparently reallocated to one of the other hard links, the bandwidth will simply be reduced, or to a reserved link.
- Better use of physical resources. Traffic can be balanced between hard links....

Terminology

- A combined port group is called **Link Aggregation Group** (LAG), different vendors have their own terms for the concept. He can also be called Team, Bond, ... etc..
- The rule that defines which packets are sent along which link is called the **Scheduling Algorithm**
- The active monitoring protocol that allows devices to include or remove individual links from the LAG is called the **Aggregation Protocol.**
- When an EtherChannel is configured, the resulting virtual interface is called the **port channel**. Physical interfaces are grouped together in this interface....

Several names for one named

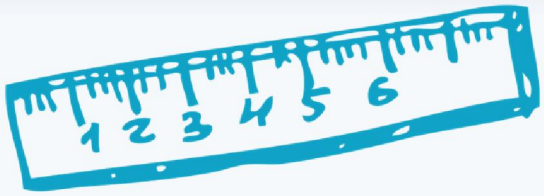
- LAG (Link Aggregation),
- Shortest Path Bridging,
- trunk Ethernet,
- EtherChannel,
- N.I.C teaming,
- port channel,
- port teaming,
- port trunking,
- link bundling,
- multi-link trunking (MLT),
- NIC bonding,
- network bonding, bonding,

Link aggregation:
Multiple physical channels, one logical channel

What are the conditions for ports?

All physical ports that make up a LAG must have:

- Same aggregation protocol enabled
- Same speed
- Same type of duplex
- Same Trunk or Access configuration
- Same VLANs transiting if the mode is Trunk
- Same VLANs if the mode is Access....



EtherChannel is used to aggregate access ports or trunks

Different types

The two main types of aggregation are static (also called manual) and dynamic:

- In static aggregation, the network administrator does the work.
- Dynamic aggregations use a protocol to negotiate parameters between the two connected devices.

Note: Some devices support static aggregations, but not dynamic aggregations with protocol....

Static aggregation

- For the creation of logical static links on the switch, the member ports of a group must be freely defined by the administrator.
- Such a method does not adapt to dynamic changes.
- When creating a LAG link group, a port must be designated as a "master port".
- All ports in a group must be set to the same speed, same duplex, same VLAN, same mode.

Aggregation negotiation

For aggregation negotiation (dynamic aggregations), there are two protocols:

- ✓ **PAgP – P**ort **Ag**regation **P**rotocol

- ✓ **LACP – L**ink **A**gregation **C**ontrol **P**rotocol

PAgP – Port Agregation Protocol

It is a Cisco proprietary trading protocol. By choosing this protocol, it is possible to configure the ports in 2 different modes:

- Auto
- Desirable

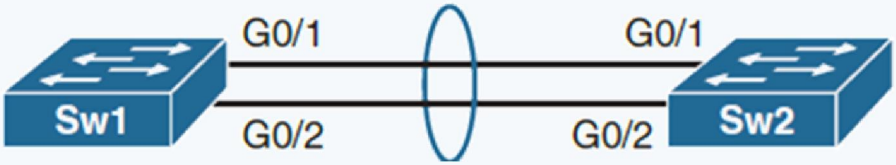
- **Desirable**: The interface initiates negotiations with other interfaces by sending PAgP packets.
- **Auto**: The interface responds to the PAgP packets it receives but does not initiate the PAgP negotiation.

✓ On mode forces the interface to channel without PAgP.

✓ Off mode disables PAgP and prevents ports from forming a port channel

PAgP Mode Settings

Modes must be compatible on either side of the EtherChannel. For example, Sw1 and Sw2 in the figure must be configured with one of the combinations of settings in Table below.



SW1	SW2	Established channel?
On	On	Oui
Auto/Desirable	Desirable	Oui
On/Auto/Desirable	Not configured	Non
On	Desirable	Non
Auto/On	Auto	Non

LACP – Link Agregation Control Protocol

LACP is a standard protocol (802.3AD) very similar to PAGP. The only difference is the names of the wearing modes.
We therefore find two modes of porting:

- Passive
- Active

Active: Corresponds to the Desirable mode of PAGP. The interface initiates negotiations with other interfaces by sending LACP packets.
Passive: Corresponds to PAGP's Auto mode. The interface responds to the LACP packets it receives, but does not initiate the LACP negotiation.

This mode forces the interface to channel itself without LACP....

LACP Mode Settings

As with PAgP, LACP modes must be compatible on either side of the EtherChannel. For example, Sw1 and Sw2 in the figure must be configured with one of the combinations of settings in Table Below.

SW1	SW2	Established channel?
On	On	Oui
Active/Passive	Active	Oui
On/Active/Passive	Not configured	Non
On	Active	Non
Passive/On	Passive	Non

NOTE: For both protocols, ON mode creates the EtherChannel configuration unconditionally, without PAgP or LACP dynamic negotiation....

LACP Settings

There are several LACP parameters contained in the LACP PDUs that are exchanged between the switches.

System Priority: Can be configured automatically or manually. LACP uses the system priority with the MAC address of the device to form the system ID.

Port priority: Each port on the switch must have a port priority. The port priority and port number make up the port identifier.

Administrative Key: Automatically configured by each port. It defines the ability of a port to aggregate with others. Ports that have the same key can be aggregated together.

LACP Settings

- The switch with the lowest system priority is called **Switch master**
- The Switch master is allowed to decide which ports actively participate in EtherChannel.
- Non-active links are paused and activated if any of the active links go down.
- The Switch master allows switching from one link to another in the event of a failure).
- Ports become active based on their priority. A lower number means a higher priority.

MAC address learning

Physical learners

They are switches that learn MAC addresses by using the physical ports in EtherChannel instead of the EtherChannel logical link.

Traffic is forwarded based on the physical port through which the address was learned.

Logical learners

Address learning is based on the aggregated EtherChannel port.

This allows the switch to transmit packets using one of the interfaces on the EtherChannel. Aggregate learning is the default.

LACP

	Active	Passive
Active	Yes	Yes
Passive	Yes	No

PAgP

	Desirable	Auto
Desirable	Yes	Yes
Auto	Yes	No

Procedure

When an EtherChannel link is configured using PAgP,

- PAgP packets are sent between EtherChannel-enabled ports to negotiate the formation of a channel.
- These packets are every 30 seconds.
- PAgP verifies configuration consistency and handles link additions and failures between two switches...

EtherChannel and IEEE 802.1AX

The two protocols are very similar and accomplish the same goal. There are some differences between the two:

- EtherChannel is a proprietary protocol from Cisco, while 802. 1AX is an open standard
- EtherChannel requires precise configuration of the switch, whereas 802.3ad is only an initial configuration.
- EtherChannel supports multiple modes of load distribution across the different links, whereas 802.3ad only has one,
- EtherChannel can be configured automatically by LACP and PAgP, while 802. 1AX can only be done by LACP.

Each switch supports a maximum number of LAGs. A Catalyst 4500 supports a maximum of 64 LAGs.

How is the load **distributed** among the different links?

Load-balancing

The distribution of throughput between the different links depends on the level of functionality of the switch. Three load-balancing algorithms are proposed:

- by MAC address (source and/or destination, Layer 2);
- by IP address (source and/or destination, Layer 3);
- per application port (destination, Layer 4).

By default, load-balacing is done by the **source MAC address**. For an L2 Switch

Options of load-balancing

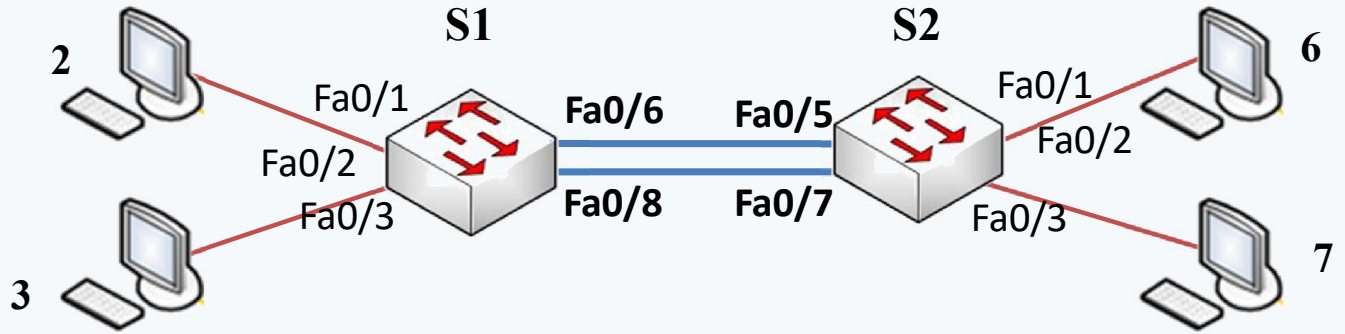
For a Level 2 Switch, the load-balancing algorithms are:

Mode	Criterion
dst-ip	is based on the IP address of the destination host.
dst-mac	based on the MAC address of the destination host
src-dst-ip	based on the source IP address and the destination IP address
src-dst-mac	based on source MAC address and destination MAC address
src-ip	based on the IP address of the source host.
src-mac	based on the source MAC address of the incoming packet.

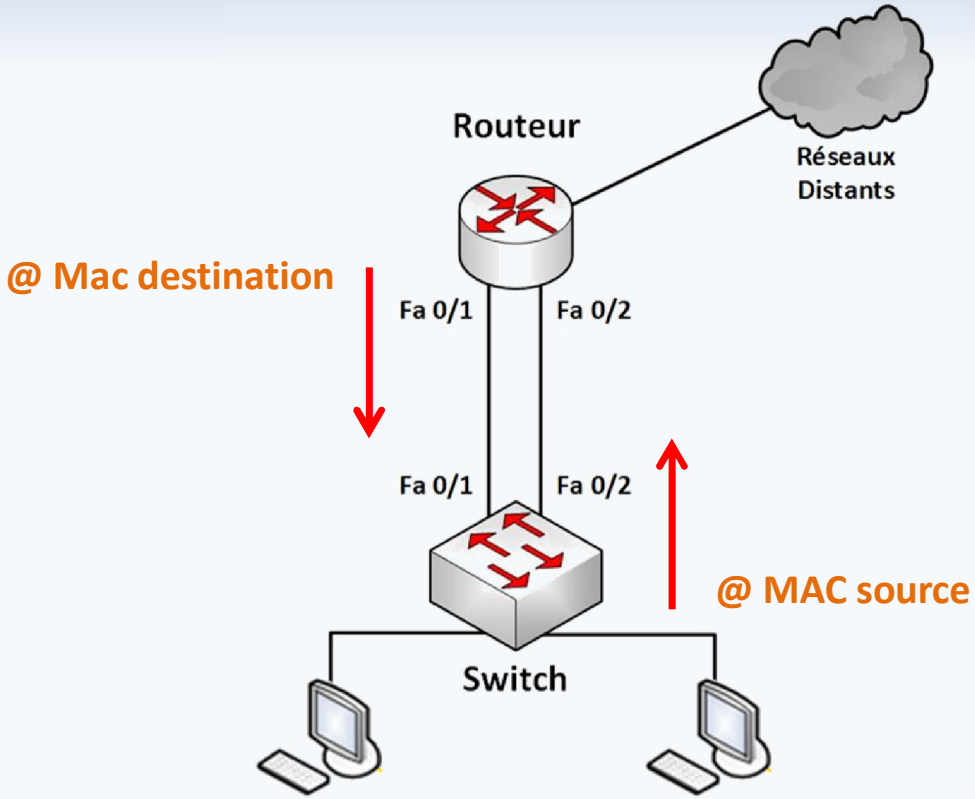
Load-balancing

If we choose to use the source MAC address (the default), the switch will send all frames with the same source Mac address over the same link.

In the case of a destination MAC address, frames to the same destination are transmitted on the same port.



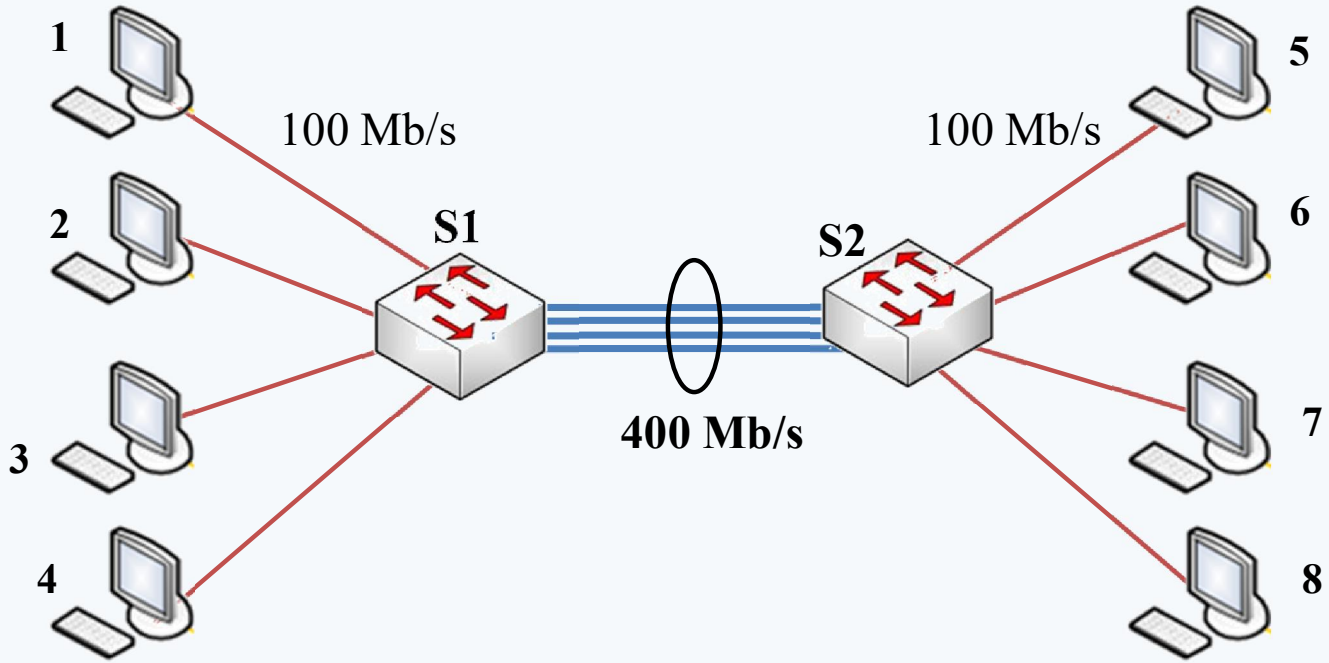
Load Ballancing (Load sharing)



- On the switch, it will be wiser to base yourself on @ MAC source.
- On the router, it is appropriate to use @ Mac destination

Load-balancing based on MAC address

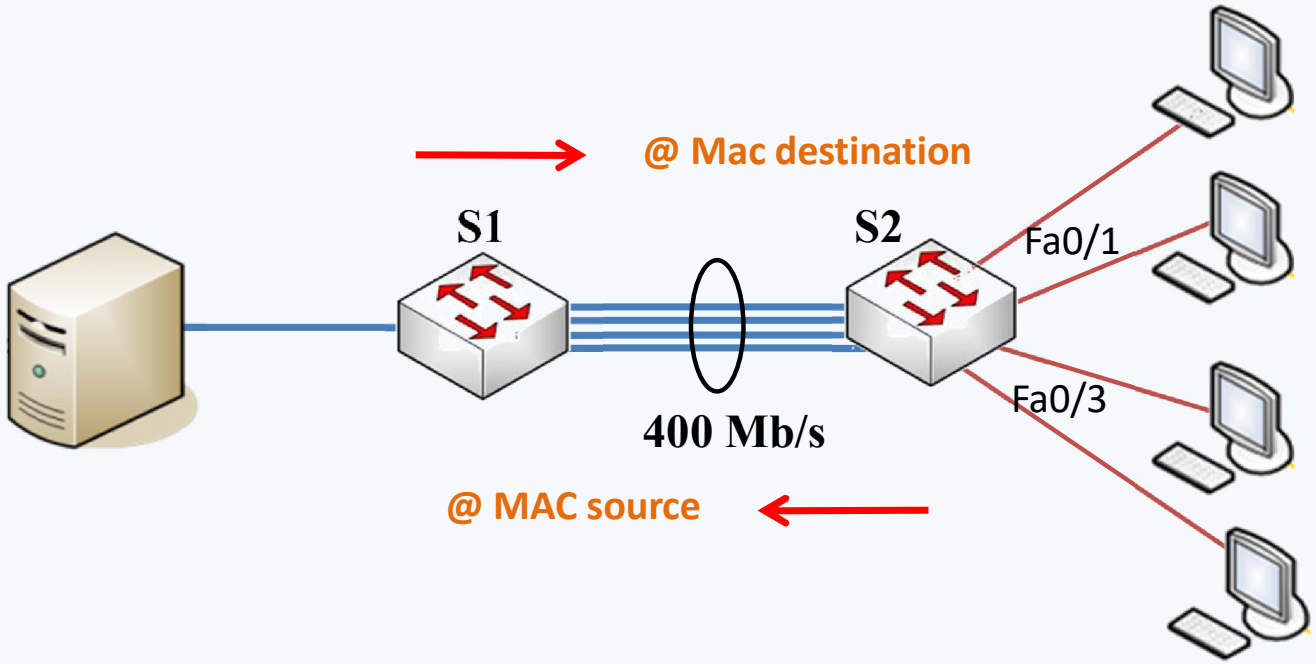
The 4 PCs send traffic to 4 other PCs. The balancing algorithm should be **src-dst-ip** or **src-dst-mac**, so that there is a better chance that each PC will get a different interface for its streams when sending to different neighbors.



Stream balancing is done on the basis of frames, not bits

Load-balancing based on MAC address

If the server is connected and sending traffic to multiple PCs, we should use the algorithm that takes destination hosts into account so that the LACP balancing algorithm has a better chance of giving separate interfaces.



The feed is not evenly distributed across all links

How to choose

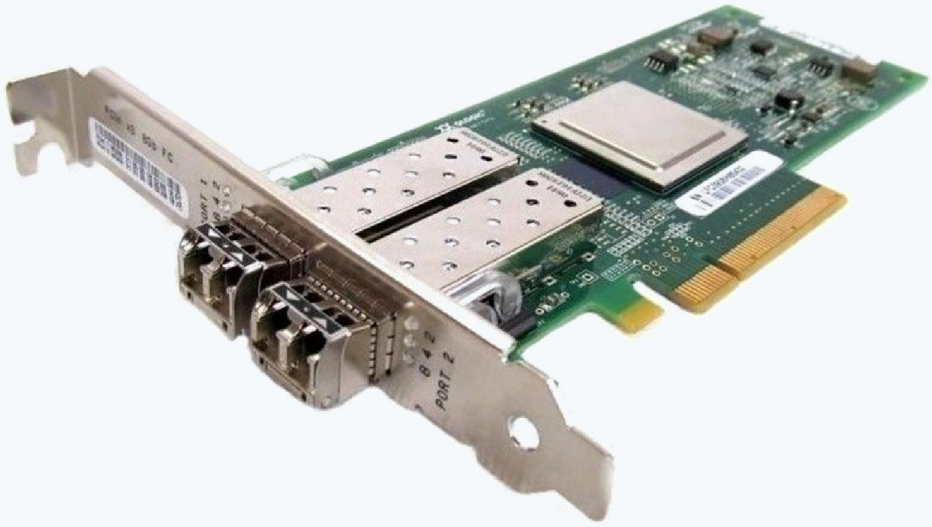
The choice of a particular load balancing method should be based on:

- the position of the switch in the network, and
- The type of traffic that needs to be loaded.

Note: For a Cisco Switch the same mode applies to all Portchannels, on the other hand for Huawei it is possible to configure a mode by LAG....

Fibre Channel

4* 8GB Fibre Channel

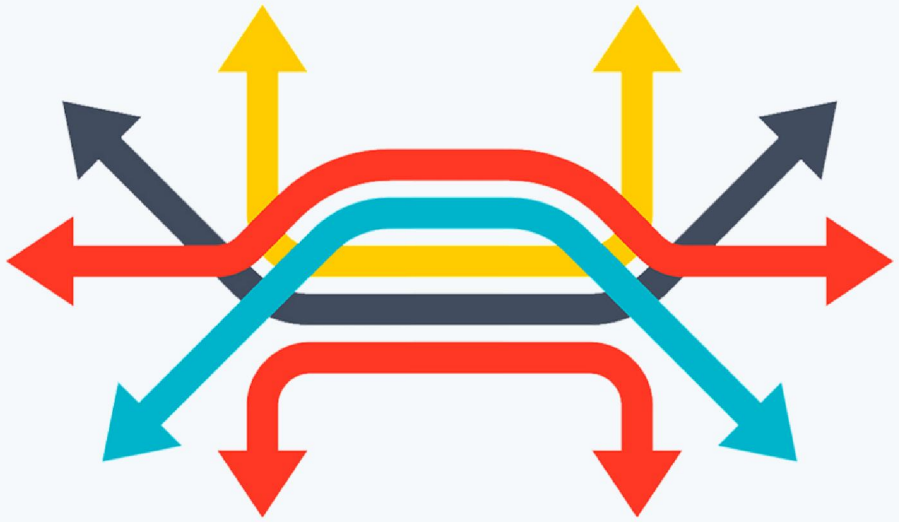


2* 8GB Fibre Channel

Chapter 05

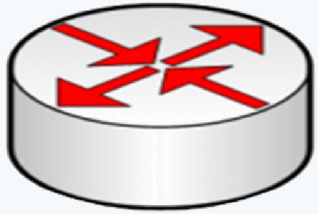
Static Routing

Part: 01

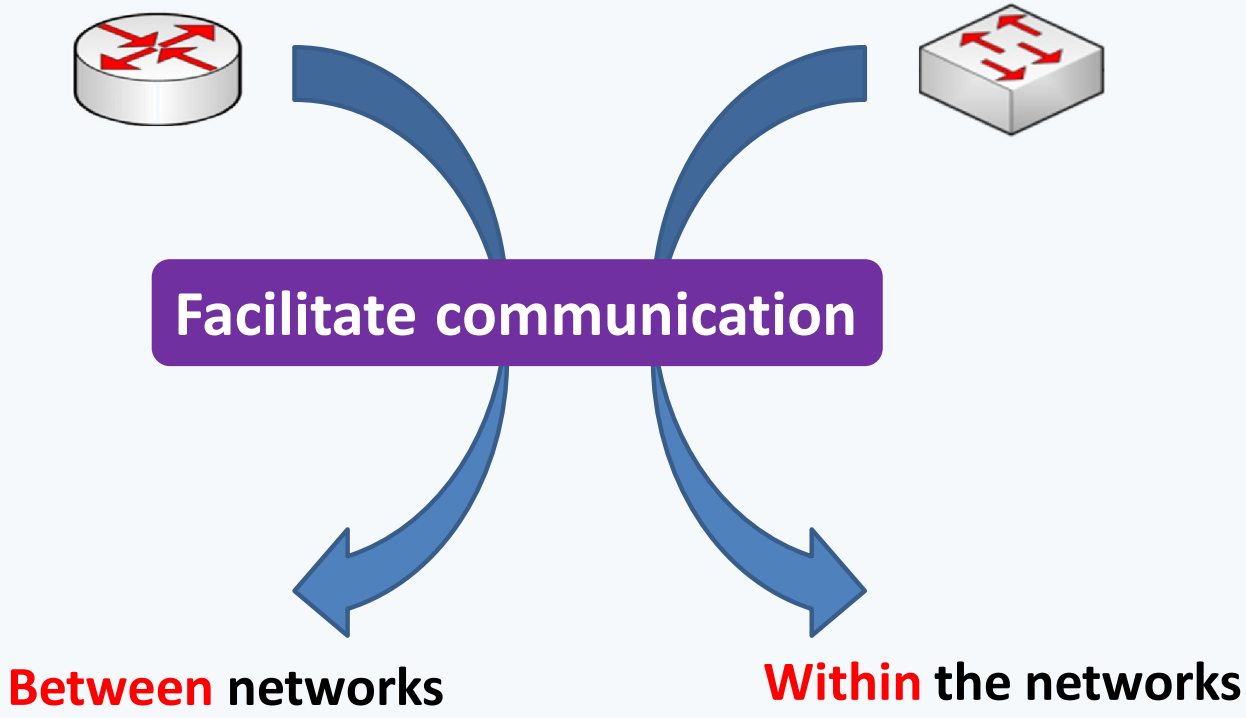


The router

- A router is an interconnecting device that routs packets from one network to another.
- It is a layer 3 equipment compared to the OSI model
- Routers rely on a routing table to identify where a data packet should be transferred.
- The routing function processes IP addresses based on their network address defined by the subnet mask and redirects them according to the routing algorithm and its associated table...




Remember!!!

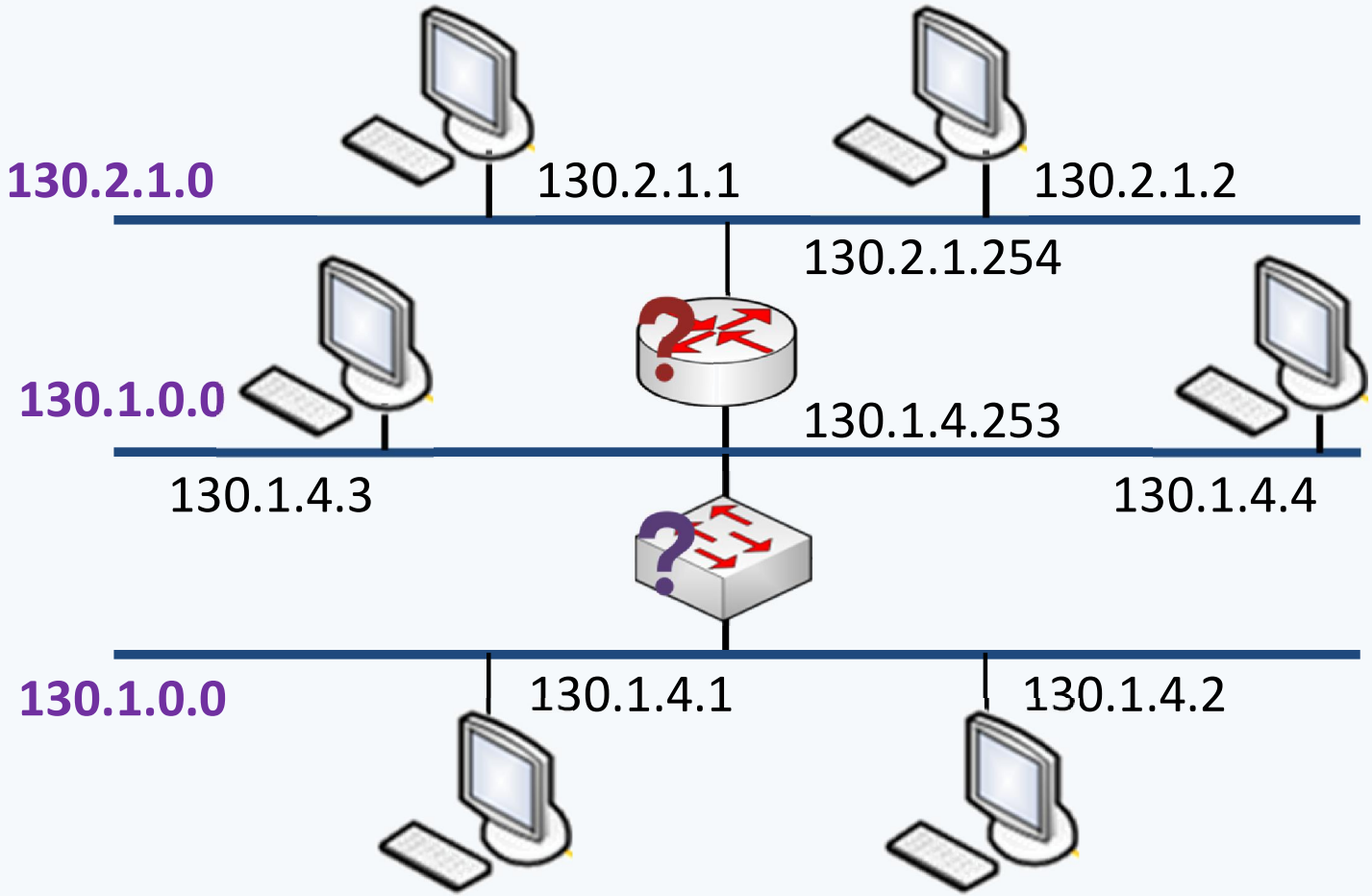

Facilitate communication
Between networks
Within the networks

What is routing?

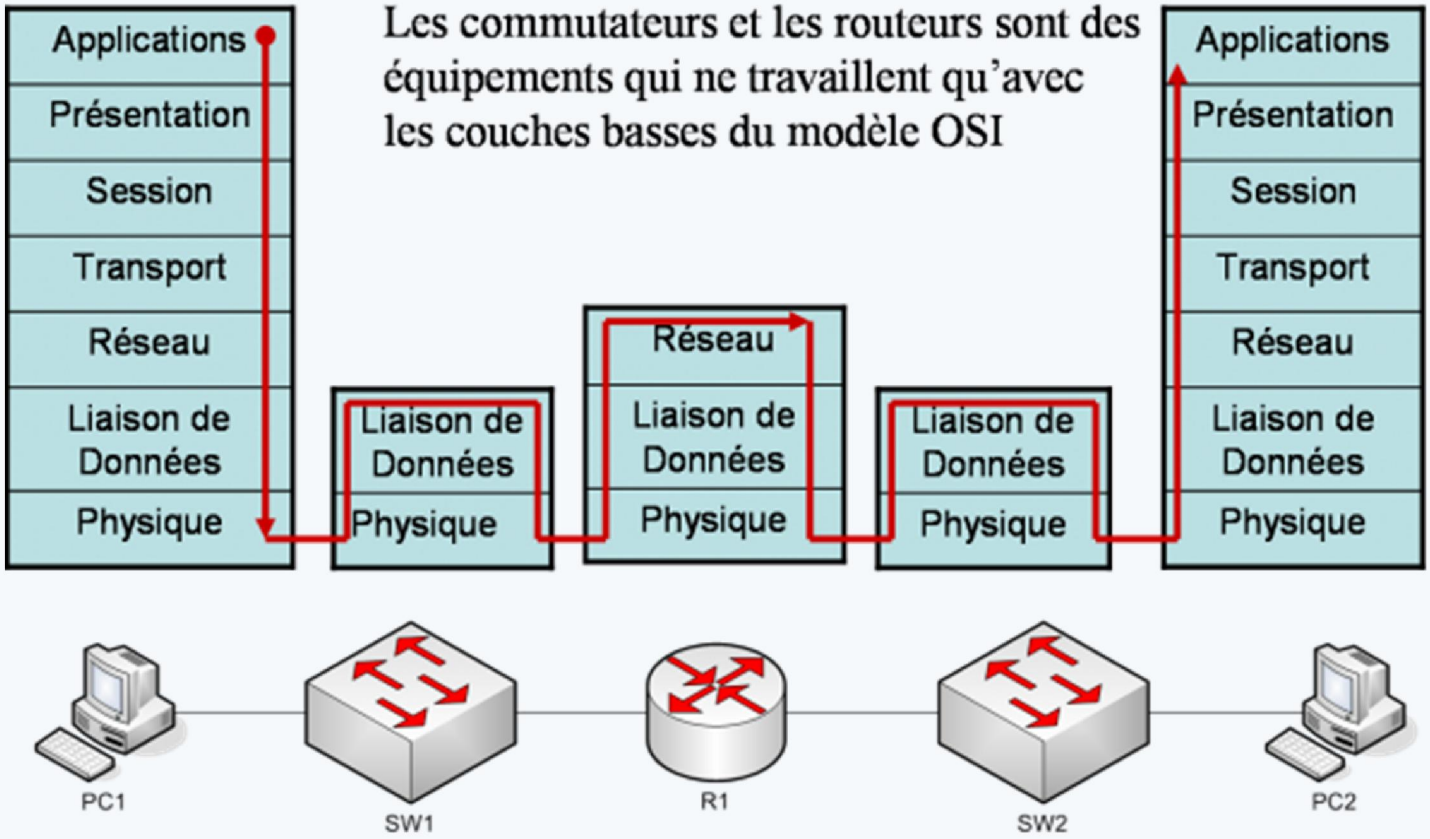
The process by which an element (mail, phone calls, trains, IP packets, ...) will be routed from one place to another.
An element doing routing must know:

- The destination,
- From what source he can learn the paths to the desired destination,
- Possible routes to reach the destination,
- The best route(s) to reach the destination,
- A way to update the routes....

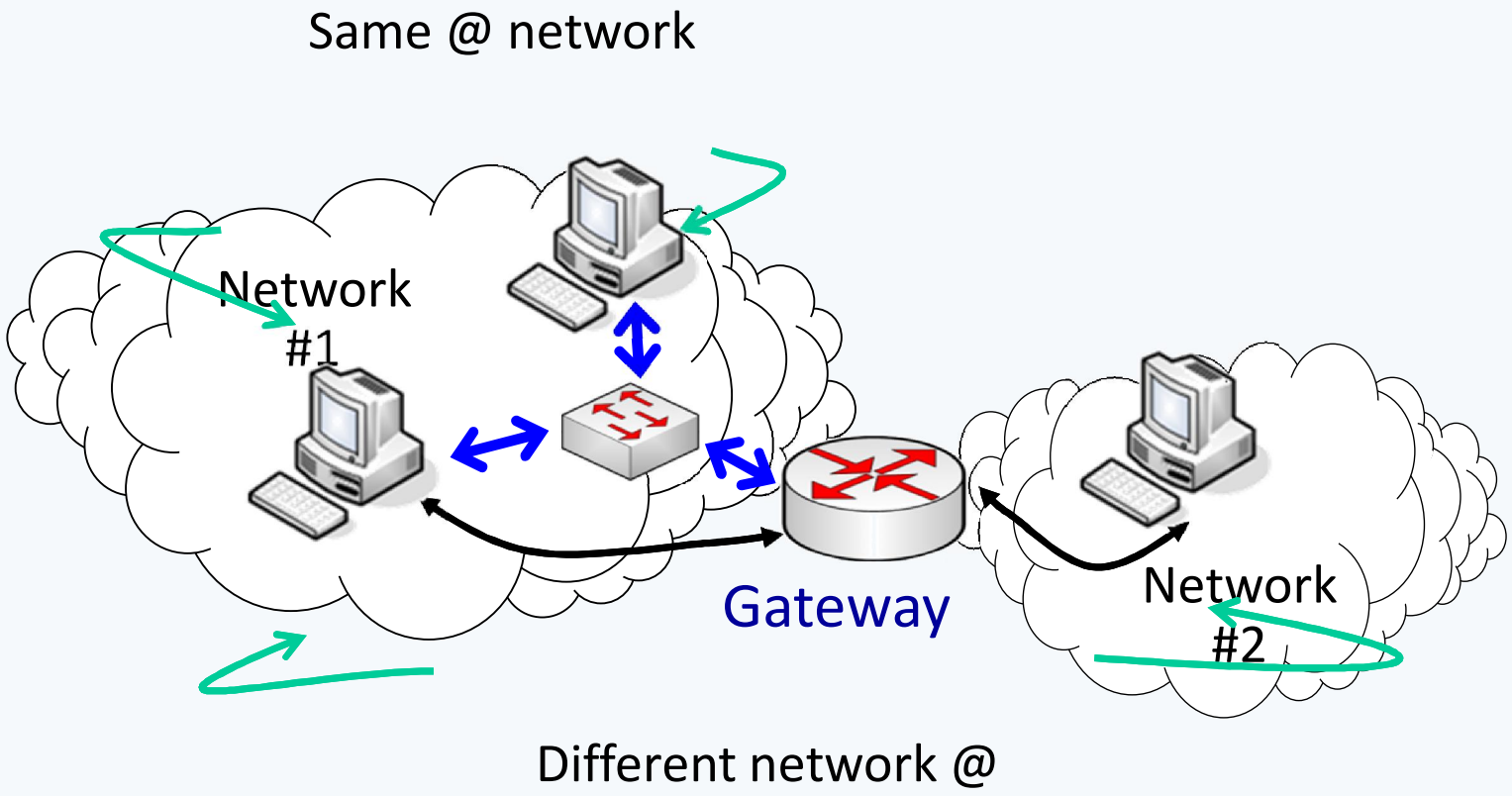
Switching or routing?



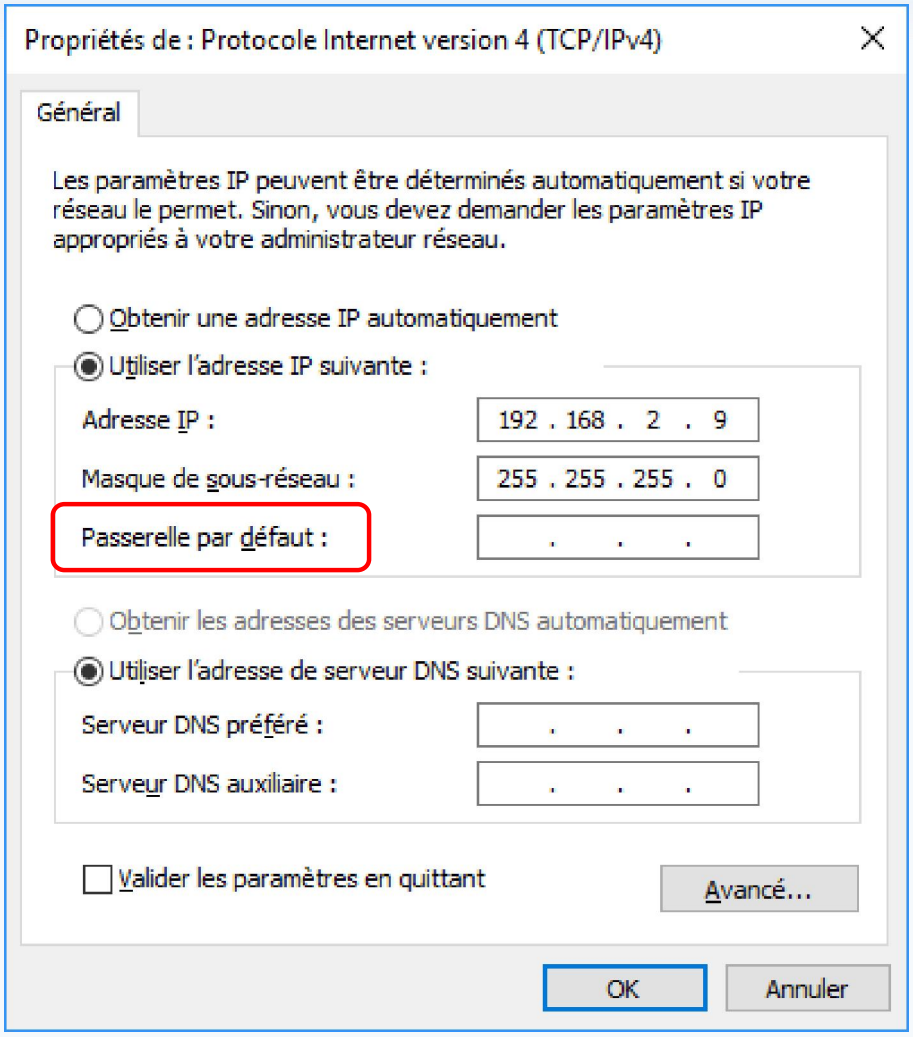
The Switch / The Router



Décision de routage

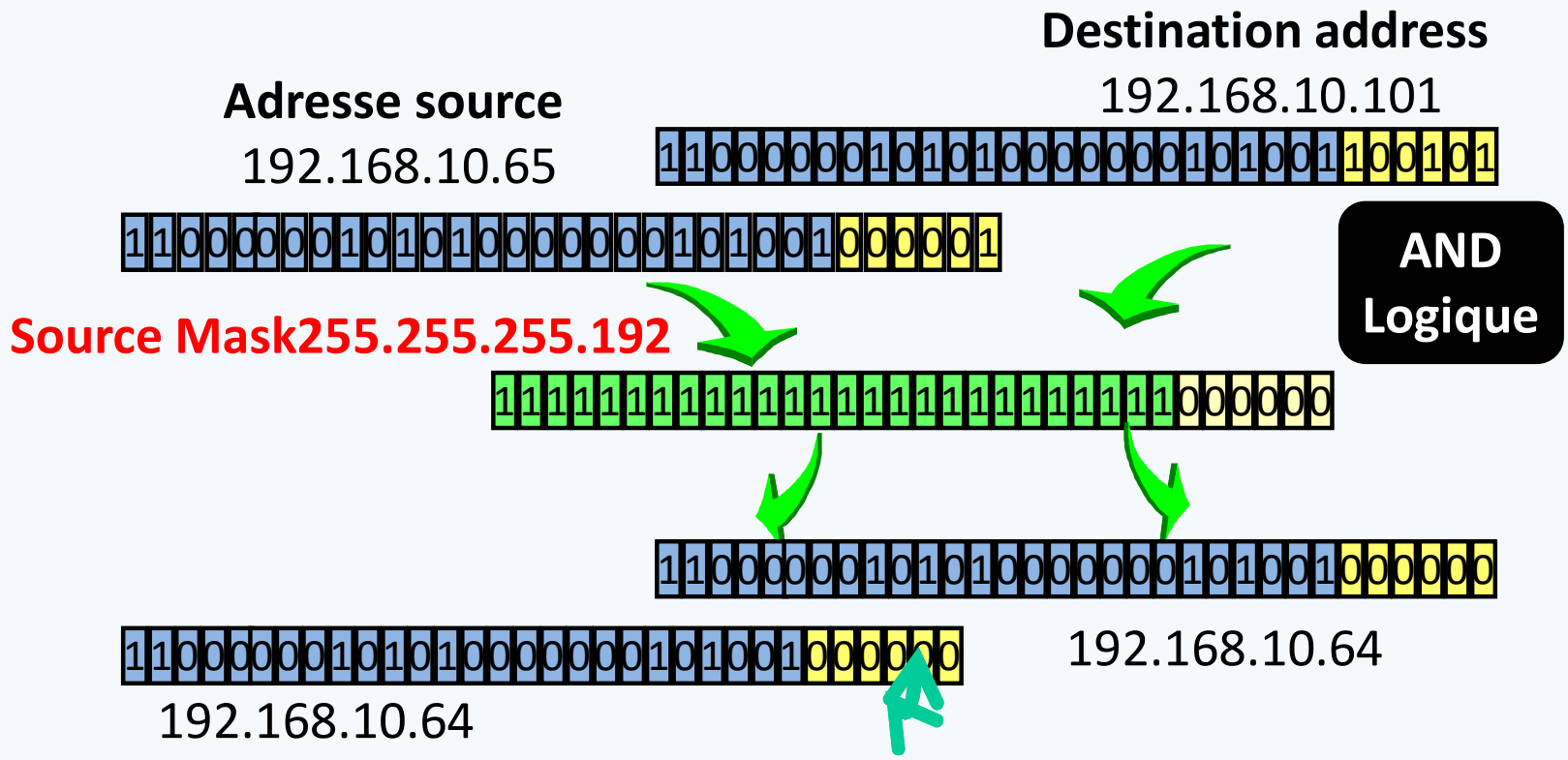


IP settings


Propriétés de : Protocole Internet version 4 (TCP/IPv4)
Général
Les paramètres IP peuvent être déterminés automatiquement si votre réseau le permet. Sinon, vous devez demander les paramètres IP appropriés à votre administrateur réseau.
Obtenir une adresse IP automatiquement
Utiliser l'adresse IP suivante :
Adresse IP : 192 . 168 . 2 . 9
Masque de sous-réseau : 255 . 255 . 255 . 0
Passerelle par défaut :
Obtenir les adresses des serveurs DNS automatiquement
Utiliser l'adresse de serveur DNS suivante :
Serveur DNS préféré :
Serveur DNS auxiliaire :
Valider les paramètres en quittant
Avancé...
OK
Annuler

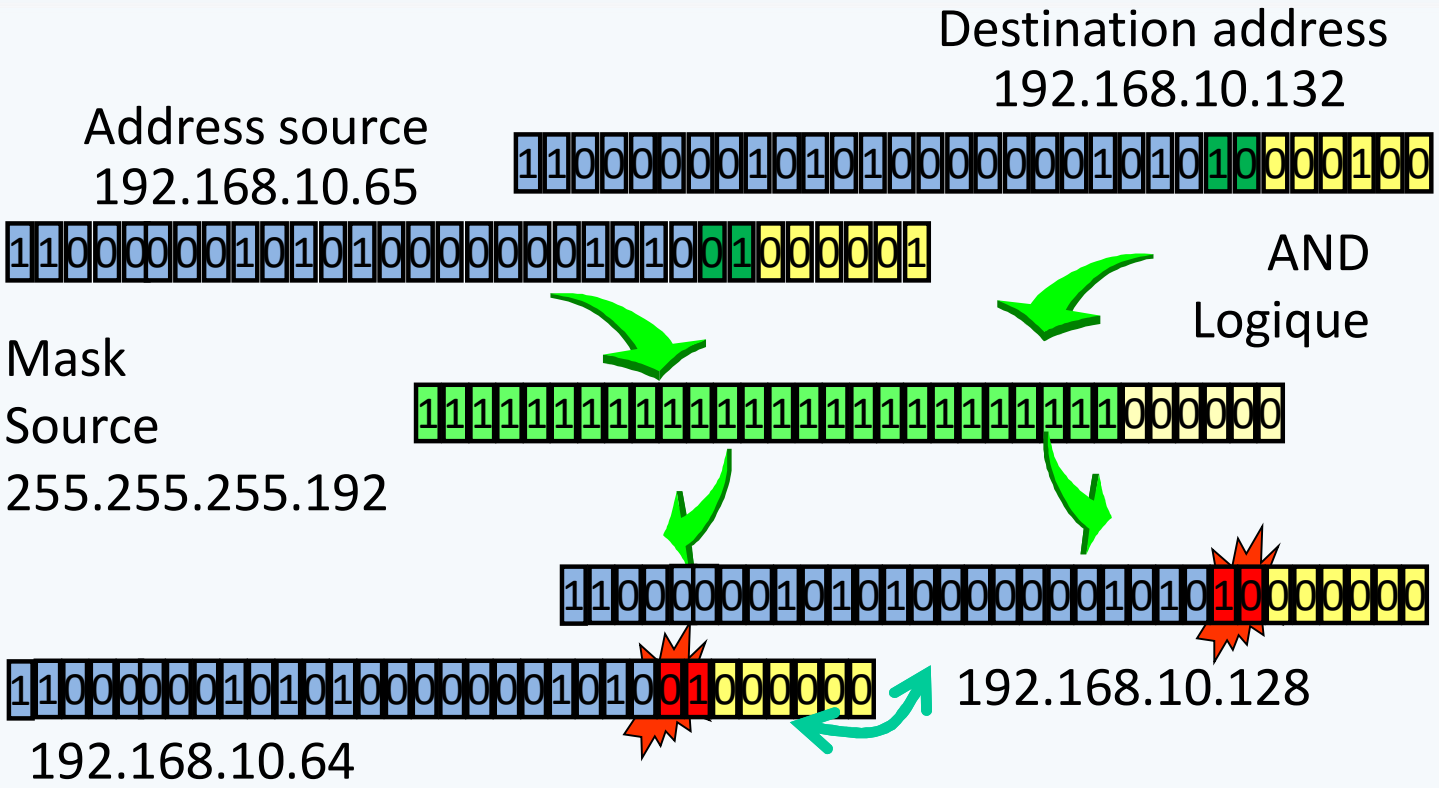
How will a PC know that IP @ source and IP @ destination are part of the same network???

Case of the the same network

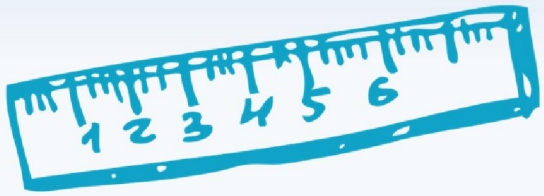


same prefix ➔ same network
Direct communication

Case of different networks



Different prefix ➔ Different network
Communication through the router



A PC1 checks if IP @ of PC2 is part of the **same subnet** through an **AND** between its own Mask and IP @ of PC2

The gateway is the **IP address** of the **router interface** that will take the data out of the local network

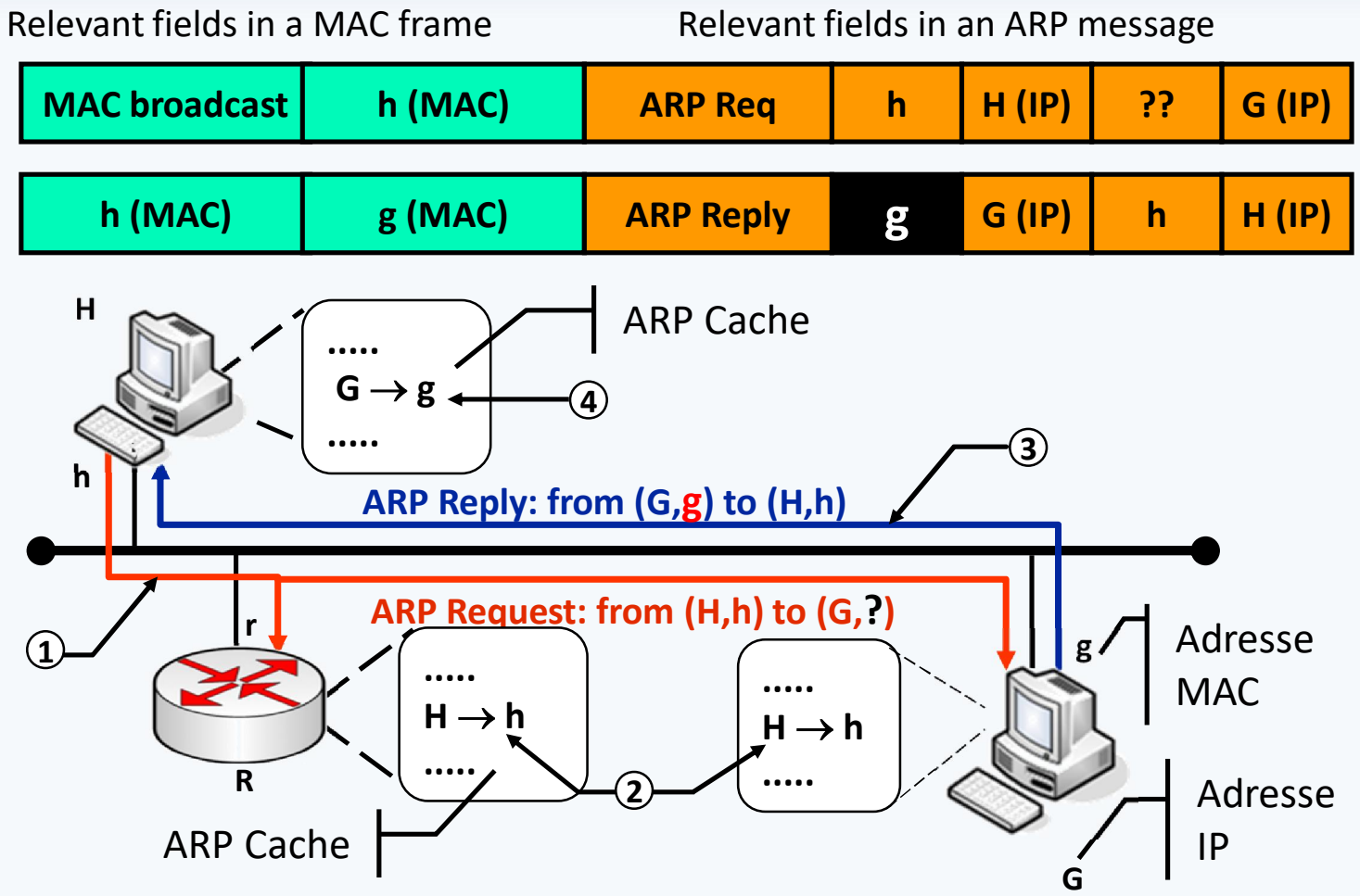
A.R.P

Address Resolution Protocol

General characteristics

- Broadcast-based solicitation protocol
- Mapping between a Layer 3 address and a Layer 2 address
- Encapsulation Ethernet Ethertype OxO8O6
- Stored in an ARP cache
- When a Layer 3 Packet Needs to Be Sent
  - If the mapping is in the cache => send the packet
  - Otherwise, an ARP request is generated...

Address Resolution



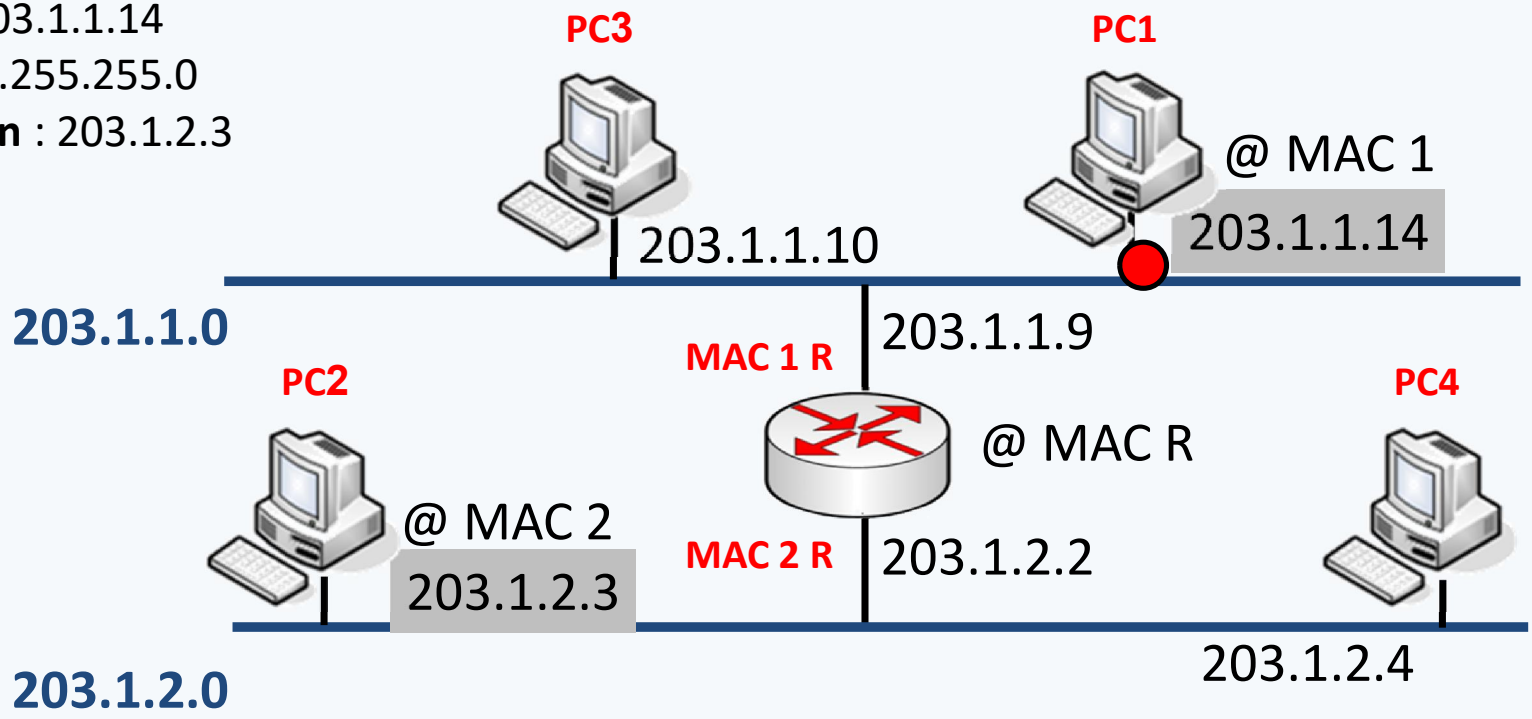
Packet forwarding by the router

Searching for a network address that matches the packet's destination IP address results in one of three path determinations:

1. Directly Connected Network: If the packet's destination IP address belongs to a device on a network directly connected to one of the router's interfaces, that packet is forwarded directly to that device.
2. Remote network: If the packet's destination IP address belongs to a remote network, the packet is forwarded to another router.
3. No route determined: If the packet's destination IP address does not belong to a connected or remote network and the router does not have a default route, the packet is discarded.

Case of a directly connected network

Source : 203.1.1.14
Mask: 255.255.255.0
Destination : 203.1.2.3
Mask : ?



PC 1 to The router

@ MAC destination **MAC1 R**
@ MAC source **MAC PC1**
@ IP destination **203.1.2.3**
@ IP source **203.1.1.14**

Router to PC 2

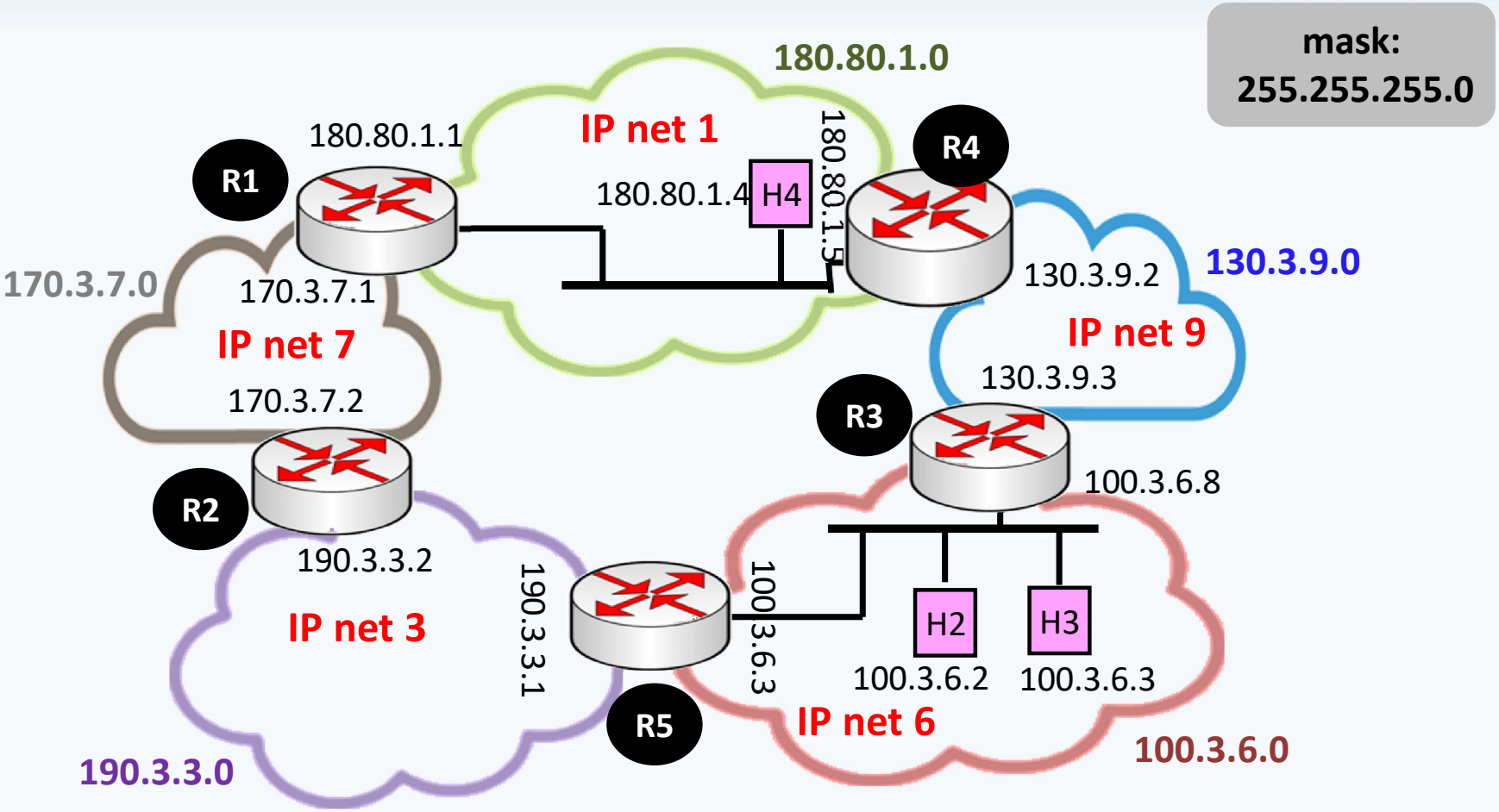
@ MAC destination **MAC PC2**
@ MAC source **MAC2 R**
@ IP destination **203.1.2.3**
@ IP source **203.1.1.14**

Explanation

1. By using the AND operation on the @IP of the destination and the subnet mask of PC1, it determines that the source and destination IP addresses are on different networks.
2. PC1 checks its ARP table for the default gateway IP address and its associated MAC address. It then encapsulates the packet in an Ethernet header and forwards it to the router.
3. The router examines the destination MAC address, which is the MAC address of one of the receiving interfaces. R1 uncaps the frame and reads @ IP destination.
4. The destination network is directly connected to R1. R1 must determine the destination MAC address corresponding to the destination IP address of the packet.
5. R1 looks up the packet's destination MAC address in its ARP cache. If the entry is not in the ARP cache, R1 sends an ARP request through its second interface.

The router has a @MAC as being a piece of equipment, **BUT** also each interface has its own @MAC.

Case of a remote network



E.g., R5 routes a packet from H2 to H4

Routing Table

A routing table will consist of lines with quadruplets: address, mask, gateway, and interface.

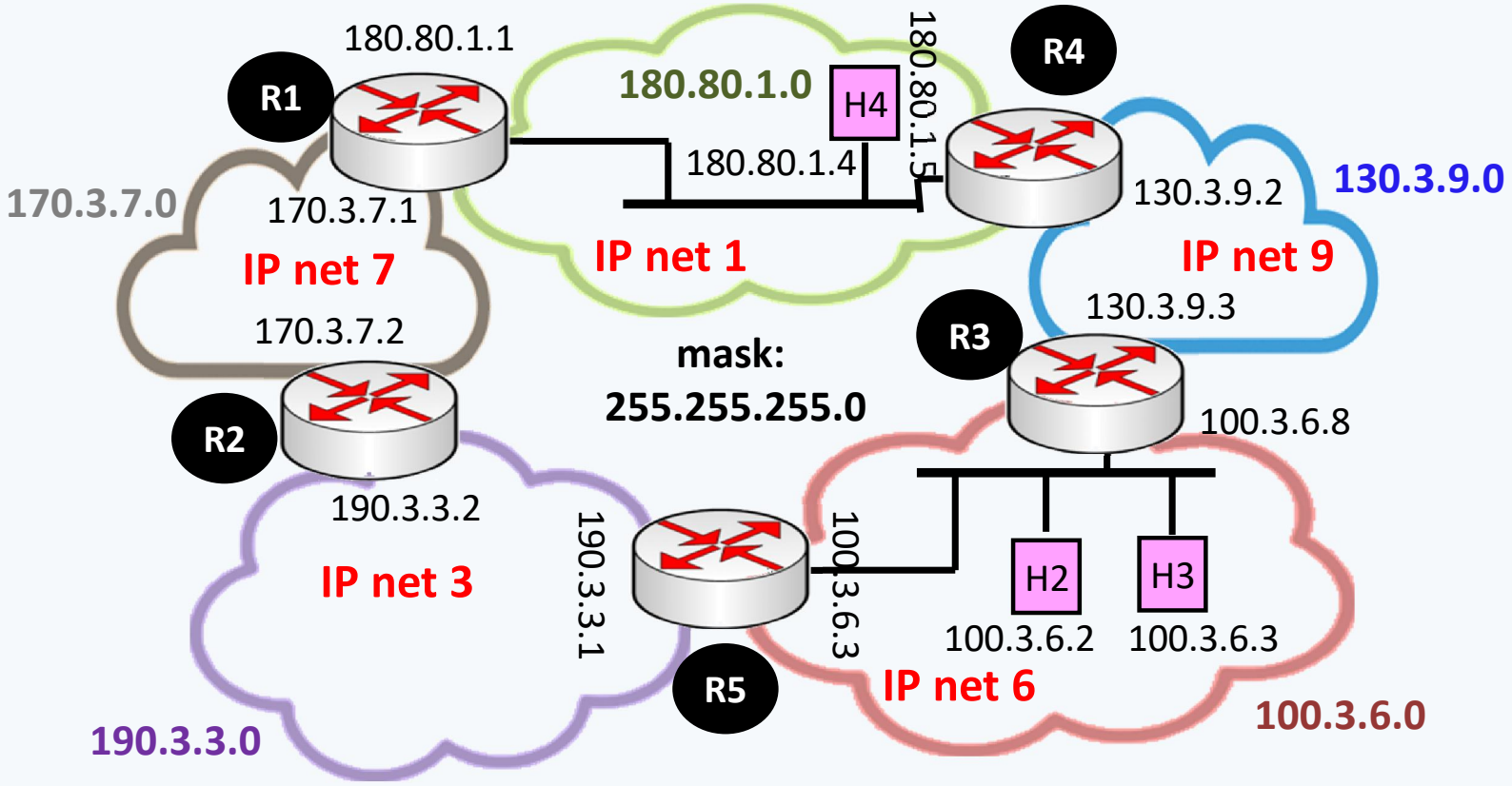
Destination	Mask	Next hop
180.80.1.0	255.255.255.0	190.3.3.2
170.3.7.0	255.255.255.0	190.3.3.2
130.3.9.0	255.255.255.0	100.3.6.8

This table must be periodically updated

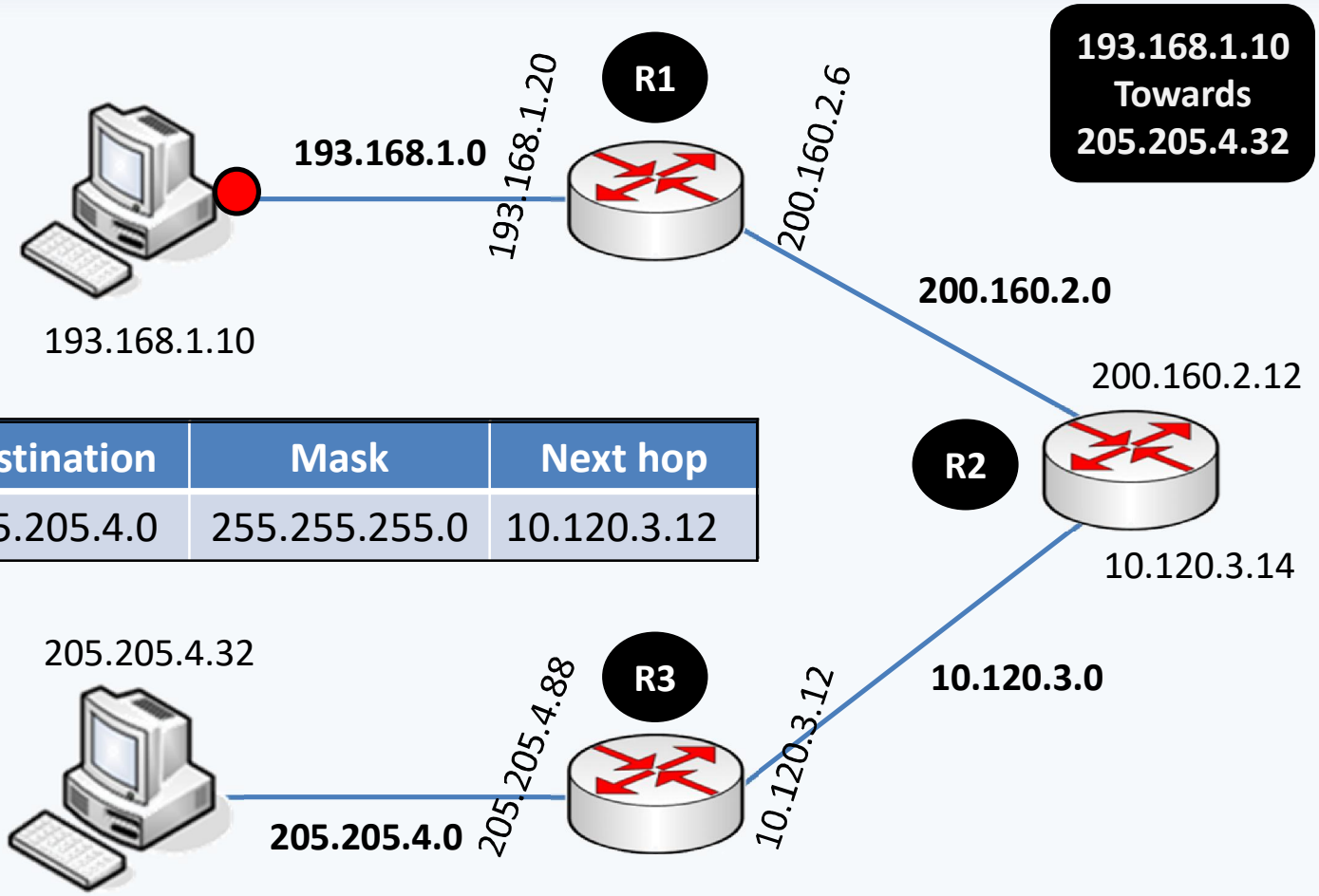
- Manually: STATIC Routing
- Automatically: DYNAMIC routing

R5

Destination	Mask	Next hop
180.80.1.0	255.255.255.0	190.3.3.2
170.3.7.0	255.255.255.0	190.3.3.2
130.3.9.0	255.255.255.0	100.3.6.8



Example 02: Case of a remote network



Destination	Mask	Next hop
205.205.4.0	255.255.255.0	10.120.3.12

Concept of static routing

Routing Methods

A router can get routes from three basic sources:

1. Directly connected routes: Automatically entered into the routing table when an interface is enabled with an IP address
2. Static routes: Manually configured by the network administrator and entered into the routing table if the egress interface of the static route is active.
3. Dynamic routes: Learned by routers by sharing routes with other routers that use the same routing protocol....

Routing modes

There are two very distinct routing modes, they are:

- Static Routing
- Dynamic routing

Dynamic routing certainly has several advantages over static routing; However, networks still use static routing. In fact, networks typically use a combination of static and dynamic routing....

Static Routing

- In this routing, administrators will **configure routers one by one** within the network in order to enter the routes to take to go on which network.
- Allows the administrator to choose the path that he or she considers best to go from network A to another networkB.
- In concrete terms, a router will be a bridge between two networks and the next router will be another bridge between two other networks....

Adventages of Static Routing

Bandwidth saving: No information passes between routers to keep them up to date, bandwidth is not cluttered with informational messages.

Security: Unlike dynamic routing protocols, static routing does not broadcast information over the network since routing information is directly captured.

Knowing the path in advance: The administrator who has configured the entire topology will know exactly where the packets are going, so this can make it easier to understand an incident on the network.

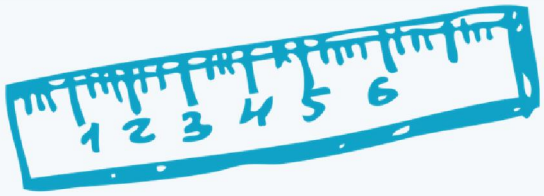
Disadvantages of static routing

- The configuration of large networks can become quite long and complex, as it is necessary to know the entire topology to capture the information.
- Can become a source of error and additional complexity as the network grows.
- It does not automatically adapt to changes and failures
- Each time the network evolves, each router must be aware of the evolution by a manual update of the administrator's by...

Using static routing

Use static routes:

- In a small network that only requires simple routing
- In a hub-and-spoke network topology
- When you want to create a fast ad hoc route
- As a backup in case the main road fails
- Do not use static routes:
- In a large network
- When you expect the network to grow in size...

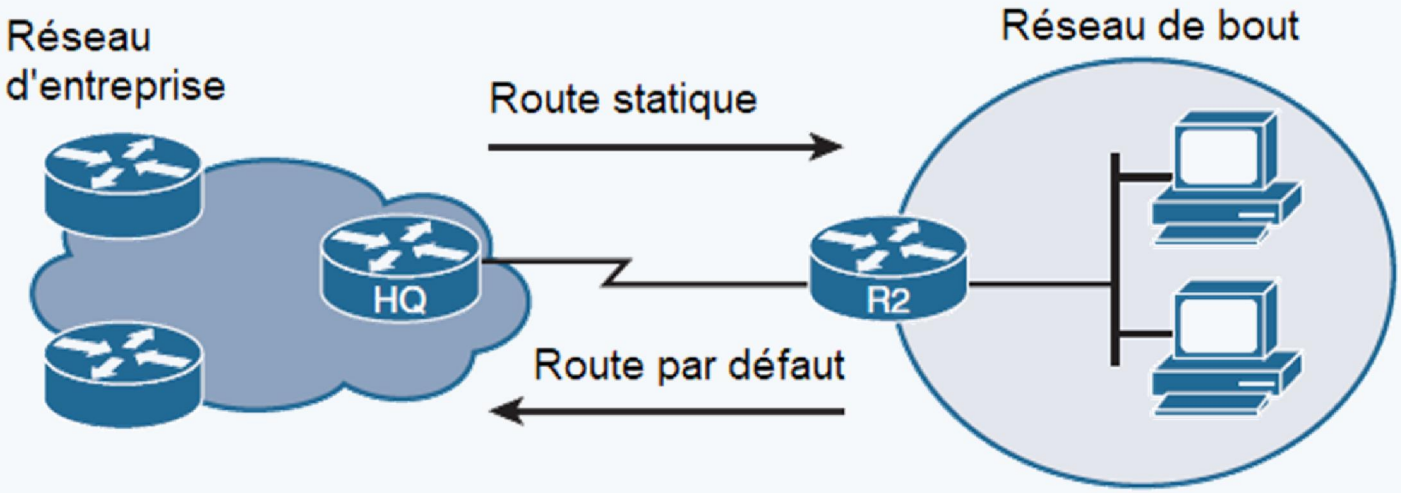


End Routers => **Static Routing**

Core Routers => **Dynamic Routing**

Example

Static routes are commonly used when routing from a larger network to a truncated network (a network that is accessed by a single link).



Administrative Distance

Administrative distance is used by routers to determine which route is best. Each road is associated with an administrative distance number, and the lower this number, the more reliable the route is considered.

Routing protocol	Administrative Distance
Directly connected	0
Static Route	1
OSPF	110
RIP	120
Unknown	255

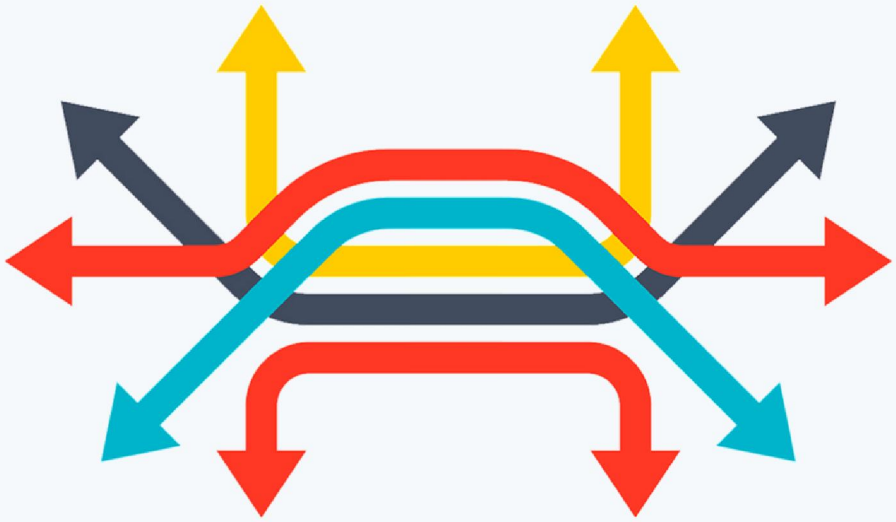
Administrative distance default values

À suivre …..

Chapter 05

Static Routing

Part: 02



Reminder

Static routing consists of constructing, in

- **each** node, a table indicating, for
- **each** destination, the address of the next node.

This table is built by the network administrator when configuring the network and after **every** change in topology....

The routing table

The routing table operates in random access memory (RAM) and includes information such as:

- Directly connected networks for any network directly connected to an interface.
- Remote networks that can be reached for any network that is not directly connected to the router.

Detailed information about these destinations includes the network address, its mask, the address of the next hop (router) to the destination, and the router's egress interface.

Directly connected routes

A network is added to the routing table by activating a router interface.

- Each interface on a router belongs to a different network block
- An interface is activated with the "no shutdown" command
- The "directly connected" network will be indicated by a "C" code in the routing table

For a static or dynamic route to be installed in the routing table, **at least a directly connected network** is required (to forward packets to a gateway)....

Static Route

- It is a route that has been manually configured.
- It is indicated by an "**S**" in the routing table
- The table must contain at least one directly connected network for a static route to function.
- A static route indicates:
  - The destination: the network to be reached and its mask
  - The direction: the IP address of the gateway or the egress interface...

Example of a routing table

```
LAB-B#show ip route
Les codes : C - connecté, S - statique, I - IGRP, R - RIP, M -
mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate
default
U - per-user static route
Gateway of last resort is not set
R 204.204.7.0/24 [120/1] via 199.6.13.2, 00:00:09, Serial0
R 223.8.151.0/24 [120/1] via 199.6.13.2, 00:00:09, Serial0
C 201.100.11.0/24 is directly connected, Serial1
C 219.17.100.0/24 is directly connected, Ethernet0
R 192.5.5.0/24 [120/1] via 201.100.11.1, 00:00:04, Serial1
C 199.6.13.0/24 is directly connected, Serial0
R 210.93.105.0/24 [120/2] via 199.6.13.2, 00:00:09, Serial0
```

Static routes using the `Next-Hop` parameter

For a Cisco router, for example, a static route entry is written as a routing table entry.
R2(config)# ip route <network> <mask> <address> [AD]

where:

- **network** : is the address of the network to be reached
- **mask** : is the mask of the network to be joined
- **address** : is the address of the next router (next-hop) directly connected to reach the network
- **AD** : Optional administrative distance (1, default)

Example :
R1(config)# *ip route* 198.130.6.0 255.255.255.0 172.16.2.1

Static routes using the `Next-Hop` parameter

- When using the next-hop parameter, the router must have a route in the table to the network to which the next-hop address belongs.
- Configuring a next-hop address requires the router to perform a **recursive lookup** to find the egress interface before it can send the packet through the interface
- A **recursive static route** uses the next-hop router to send packets to their destination.
- It requires two searches of the routing table.
- It must first look up the destination network, then the network direction for the next-hop router and the egress interface....

Static routes using the« Exit Interface »

- To avoid recursive lookup and to have a router immediately send packets to the egress interface, the static route is configured using the "Exit Interface" parameter instead of the next-hop parameter (ip-address).
- The routing table shows that the routes are directly connected, technically this is not true.
- An advantage of using this method is that the static route does not depend on the stability of the IP address of the next hop.
- This is the best practice

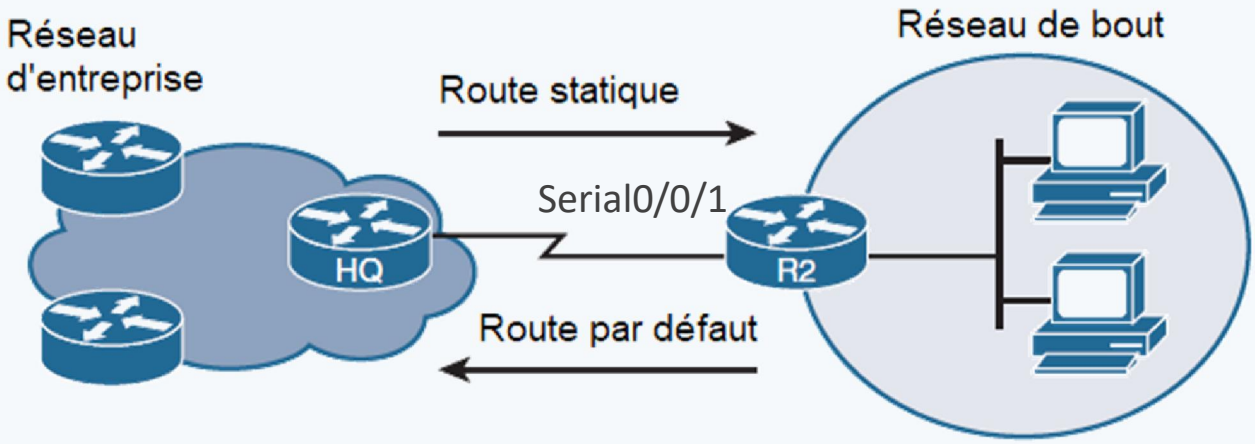
Example

R2(config)# ip route 182.16.3.0 255.255.255.0 serial 0/0/0

Default IPv4 Route

- A default route is a special type of static route,
- When a route has no specific match in the route table, the default route is a match.
- Example:

R2(config)# ip route 0.0.0.0 0.0.0.0 Serial0/0/1



Administrative Distance

The administrative distance is used by routers to determine what the best route is, the lower that number, the more reliable the route is considered to be.

Routing protocol	Administrative Distance
Directly connected	0
Static Route	1
OSPF	110
RIP	120
EIGRP	90
Unknown	255

Administrative distance default value

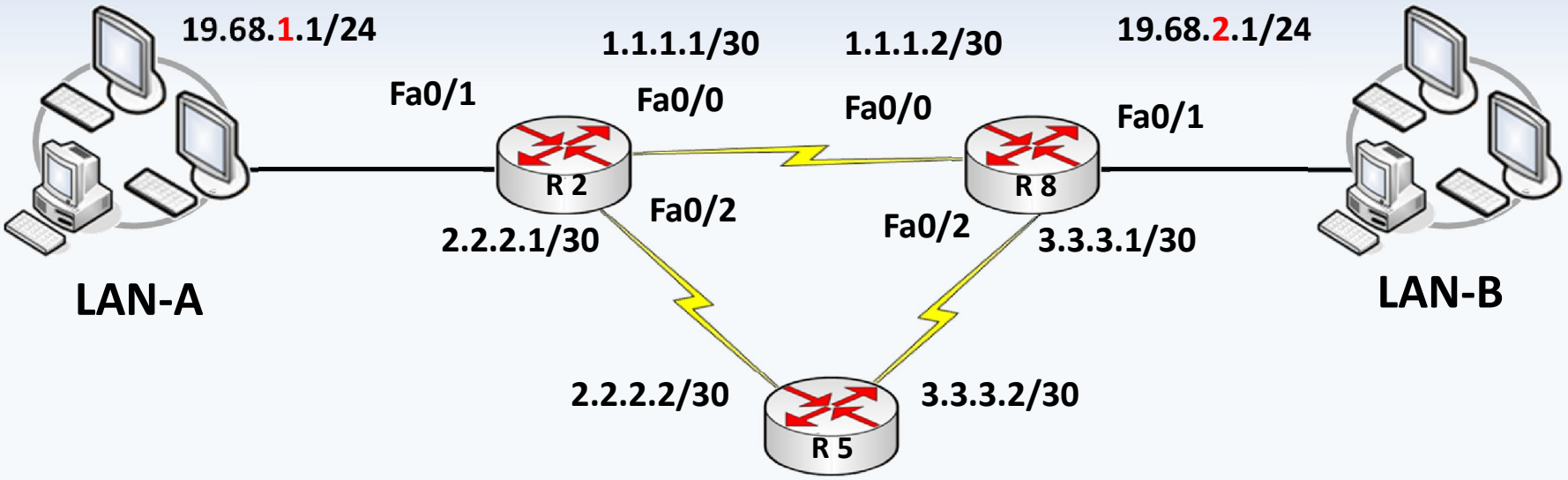
Floating static route

- It is a static route that will take over in the event of a break in the best link. It can be used as a backup mechanism.
- It is configured with a higher administrative distance than a route learned otherwise

Exemple :

R1(config)# *ip route* 195.19.6.0 255.255.255.0 172.16.2.1 5

Floating Static Route (2)



R2(config)# *ip route* 19.68.2.0 255.255.255.0 2.2.2.2
R5(config)# *ip route* 19.68.2.0 255.255.255.0 3.3.3.1
R2(config)# *ip route* 19.68.2.0 255.255.255.0 1.1.1.2 5

- The route from R2 to LAN-B directly through R8 is a **floating route**.
- The route from R2 to LAN-B through R5 and then R8 is a **normal route**.

Route Aggregation (Summary Routes)

- A summary route is a route that represents multiple routes. In general, a set of contiguous roads
- These routes share the same gateway or egress interface.

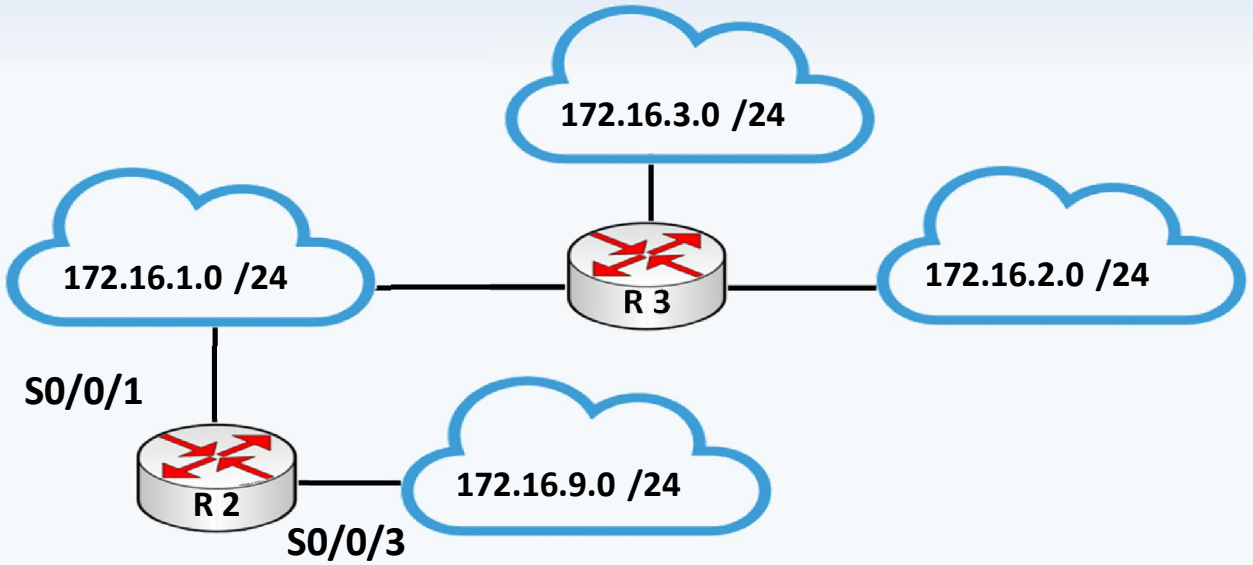
Example:

  - 130.0.0.0/16,
  - 130.1.0.0/16,
  - 130.2.0.0/16,

- These routes can be grouped in the address 130.0.0.0
- The mask that can group them together is: 255.0.0.0

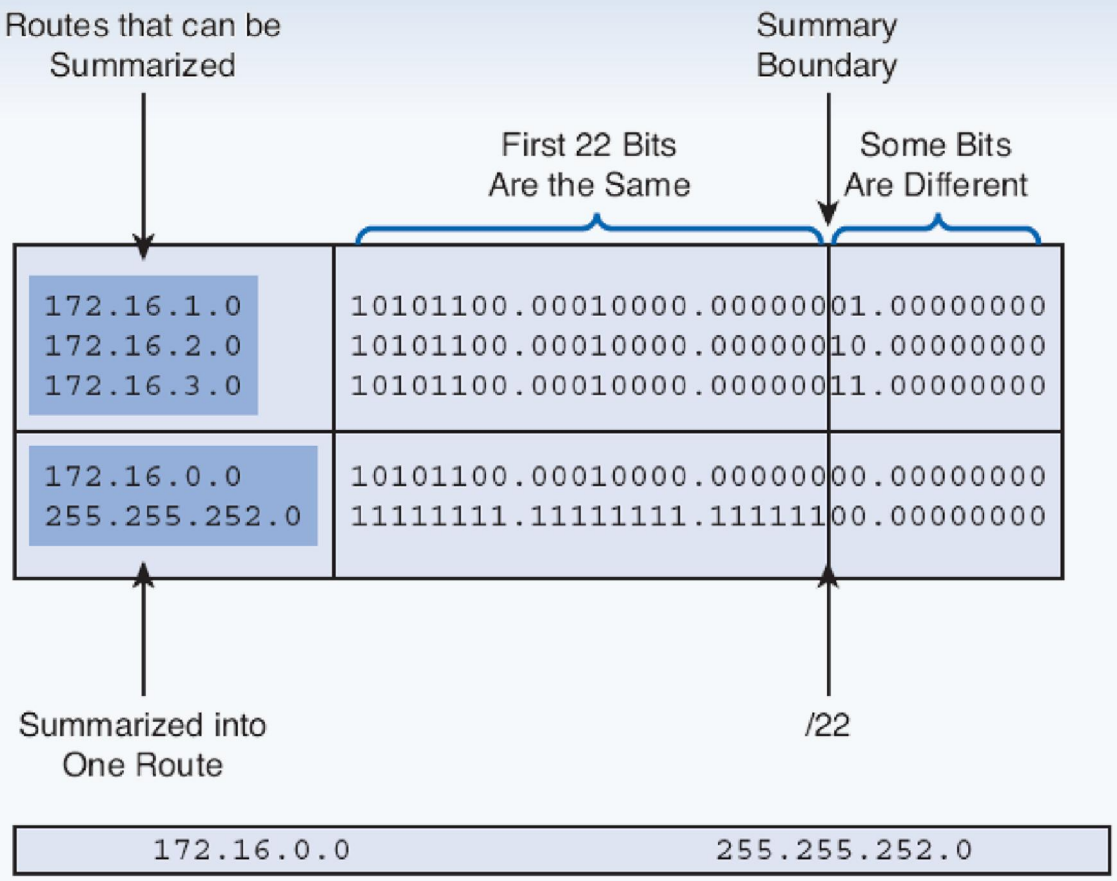
Route aggregation

Example:



- Three routes relay the data via R2's Serial 0/0/1 interface.
- These routes can be aggregated with the address: 172.16.0.0
- The largest mask that groups them together: 255.255.252.0

Calculating an aggregated route



Calculating an aggregated route

- **Step 1**. Write in binary the networks to be grouped.
- **Step 2**. Starting from the left, identify the bits Similar.
- **Step 3**. When one of the bits differs, you have the largest mask that groups them together.
- **Step 4**. The network address is made up of the n bits common, followed by 0s
- **Step 5**. Rewrite the address in dotted decimal notation....

Route aggregation

Aggregation makes it possible to:

- Reduces the number of entries in the routing table
- Increases performance when searching for matches.
- Makes the routing process more efficient.
- A single static route can represent hundreds of networks.

In 2007, more than 200,000 routes in routers at the heart of the internet.

In most cases they are aggregated routes...

Example

Address	First byte	Second Byte	Third Byte	Fourth Byte
193.168.98.0	11000001	10101000	01100010	00000000
193.168.99.0	11000001	10101000	01100011	00000000
193.168.100.0	11000001	10101000	01100100	00000000
193.168.101.0	11000001	10101000	01100101	00000000
193.168.102.0	11000001	10101000	01100110	00000000
193.168.105.0	11000001	10101000	01101001	00000000

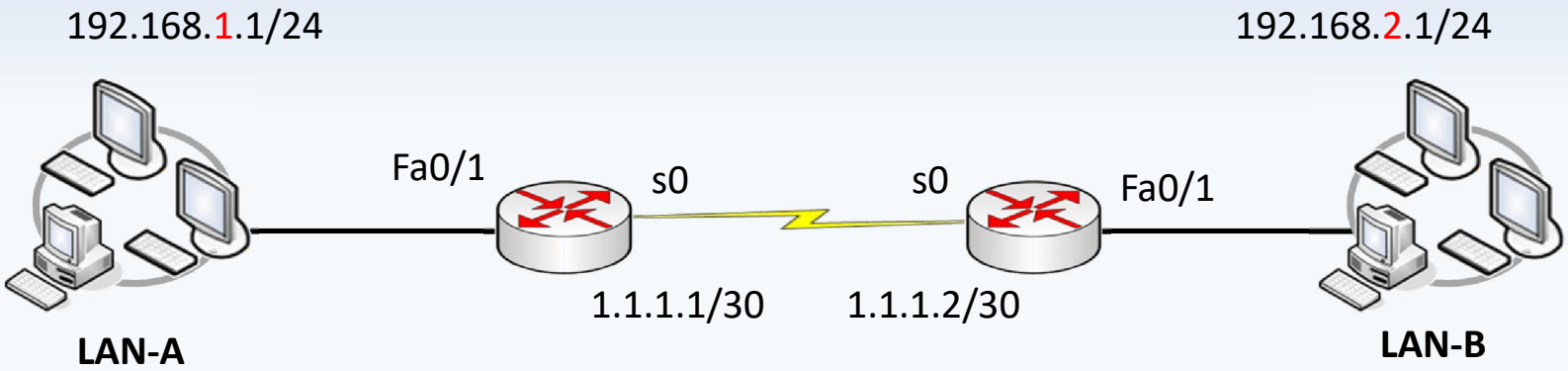
- The route summarized is: 193.168.96.0/20.
- The mask is :255.255.240.0.

Chapter 06

Dynamic routing

Part 01

Towards dynamic routing



	Destination	P
C	192.168.1.0/24	e1
C	1.1.1.0/30	s0
R	192.168.2.0/24	1.1.1.2

	Destination	P
C	192.168.2.0/24	e1
C	1.1.1.0/30	s0
S	192.168.1.0/24	1.1.1.1

C : Connected Road

S: Static Route

R : Rip (protocole de routage dynamique)

Routing Protocols

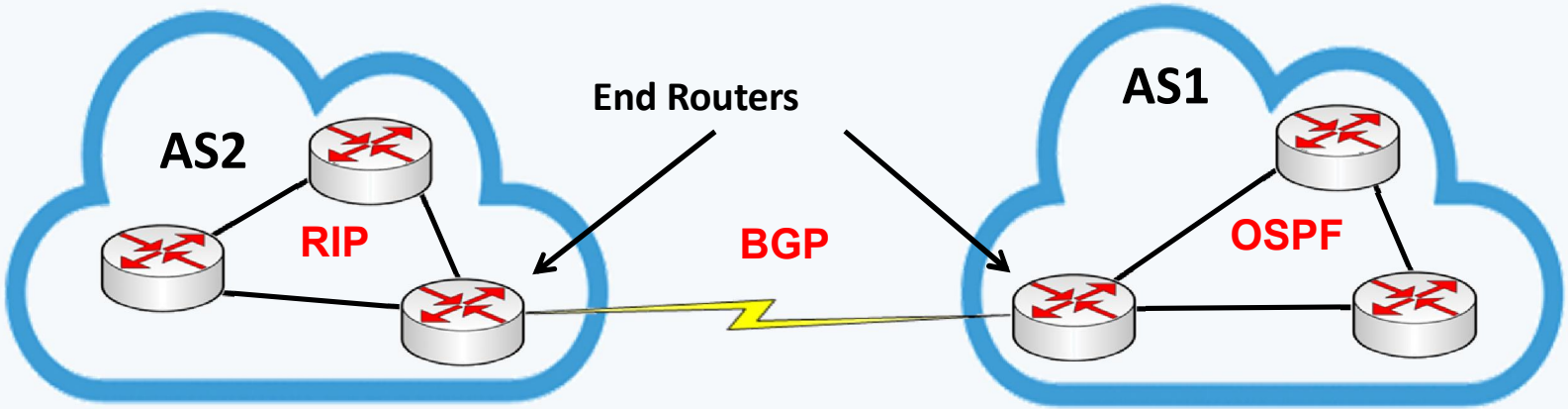
Routing protocols specify a method for **dynamically** sending and receiving routing updates between routers.
They essentially solve three problems:

- it discovers the other routers in the network,
- it builds the routing tables,
- It keeps the routing tables up to date.

The global routing domain (**Internet**) has been subdivided into Autonomous System (AS) routing domains. This division leads to a distinction between two families of routing protocols (IGP and EGP)

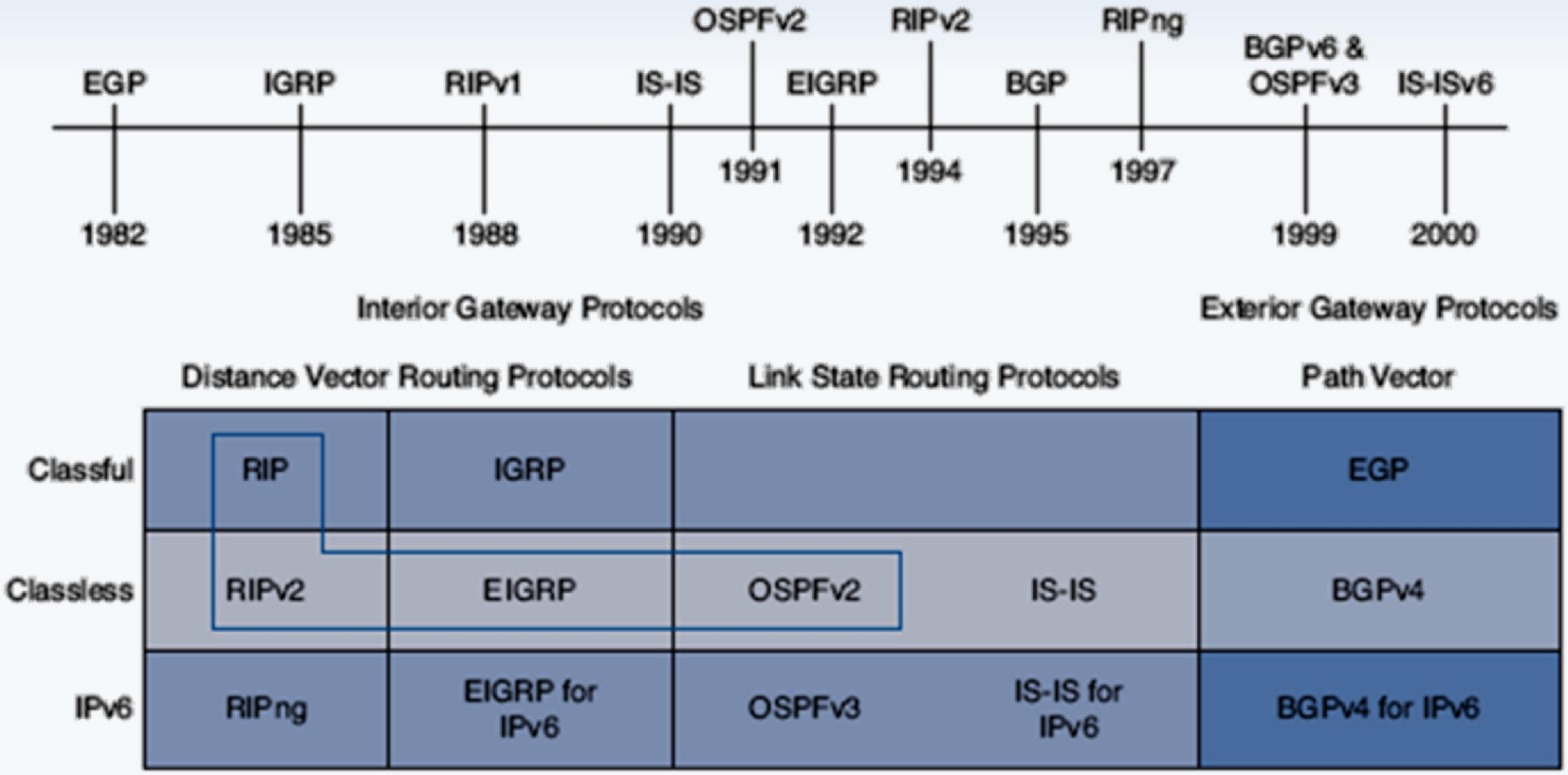
Routing Protocols

- Intra-domain routing protocols, for intra-domain routing (IGP).
- Exterior Gateway Protocol (EGP) they route packets in the inter-network.



- IGP : RIP, EIGRP, OSPF
- EGP : BGP

Routing Protocols

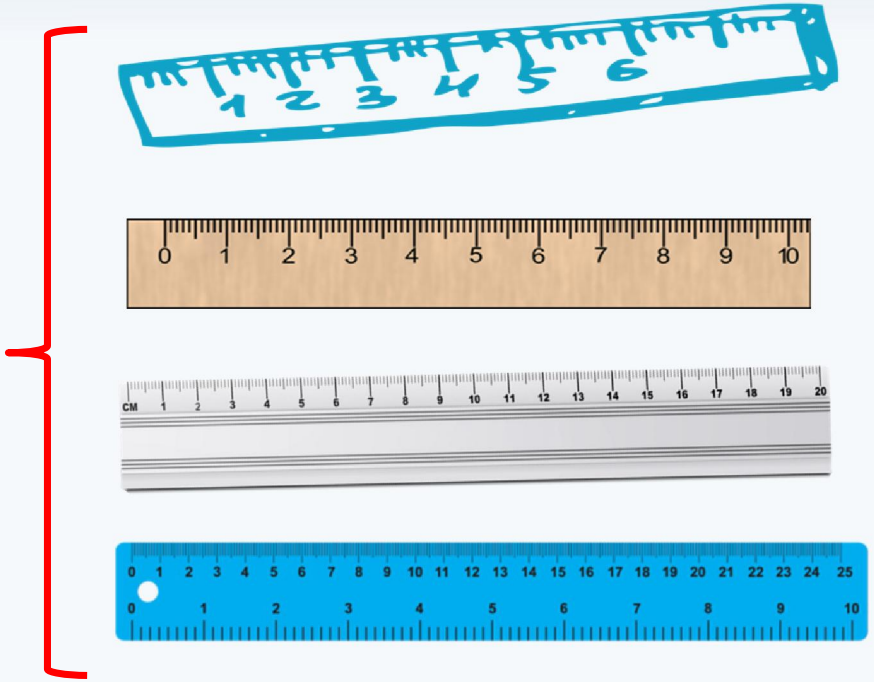


PGI Protocols

IGP routing protocols can be classified into two categories:

- Distance vector routing
  - The RIPv1 protocol
  - The RIPv2 protocol,
  - The EIGRP protocol
- Routing by Link State Information
  - The OSPF protocol....

A protocol is a sequence of



Distance vector protocol

- Uses a routing algorithm that adds distances to find the best routes (Bellman-Ford).
- Routers send their entire routing table to neighbors.
- These protocols are susceptible to routing loops.
- No router performs any particular function. We will speak of "flat" knowledge or **non-hierarchical routing**.
- They are slowly converging.
- The RIP and the IGRP are cited. EIGRP is an advanced distance vector protocol, fully powered by Cisco Systems that does not have these disadvantages....

Distance vector routing

RIP v.1

Routing Information Protocol

Vector-distance routing, RIP

- Derived from the work of Bellman-Ford, developed by the University of California for Unix and initially used in Arpanet.
- The first version was standardized in 1988
- Easiest to set up and understand
- RIP distinguishes between two types of equipment:
  - Active routers: they periodically broadcast their routes to the other nodes
  - Passive routers: Listen and simply update their tables based on the information received.

Vector-distance routing, RIP

- It uses hop count (the number of routers traversed) as the only metric for path selection;
- RIP belongs to the family of distance vector protocols, since it calculates the distance, in the number of routers traversed, between origin and destination.
- RIP does not guarantee that the selected path will be the fastest.
- A short but congested path can be a poor choice compared to a longer but totally clear path.
- Routing messages are encapsulated in a **UDP** segment using port 520 and are **broadcast every 30s**.

Administrative Distance

The administrative distance is used by routers to determine what the **best route** is, the **lower** that number, the more reliable the route is considered to be.

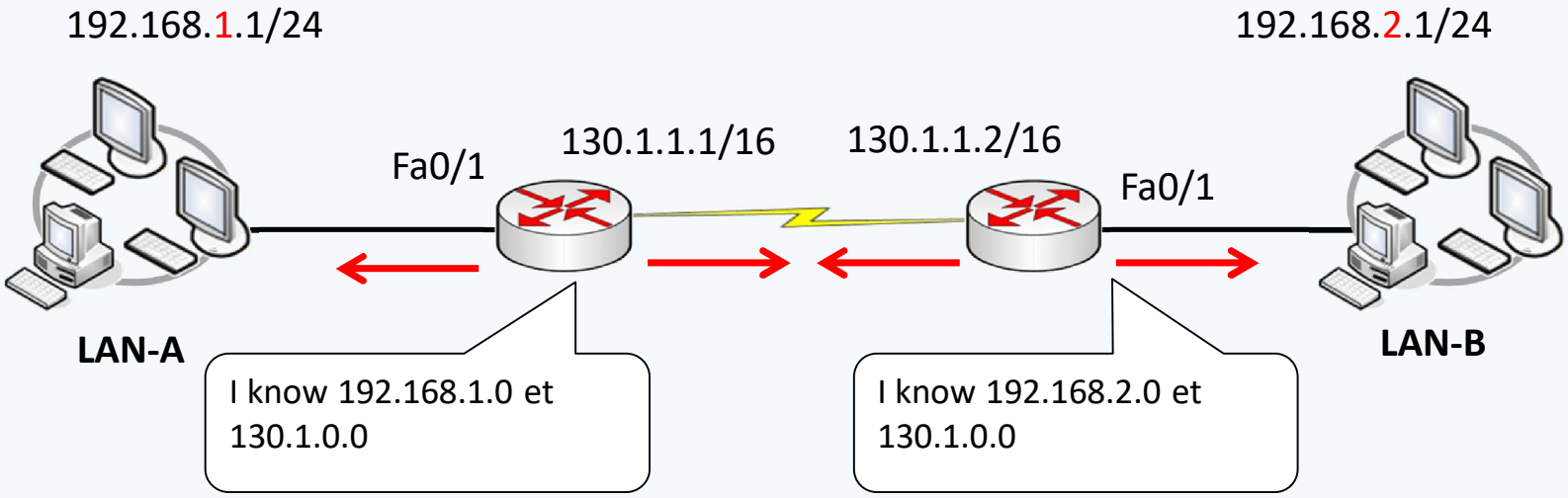
Routing protocol	Administrative Distance
Directly connected	0
Static Route	1
OSPF	110
RIP	**120**
EIGRP	90
Unknown	255

Administrative distance default value

The purpose of the protocol is: to discover networks that are not directly connected, in order to route the packets afterwards

Principle of the RIP protocol

A router periodically announces to **its neighbors** the routes it knows.

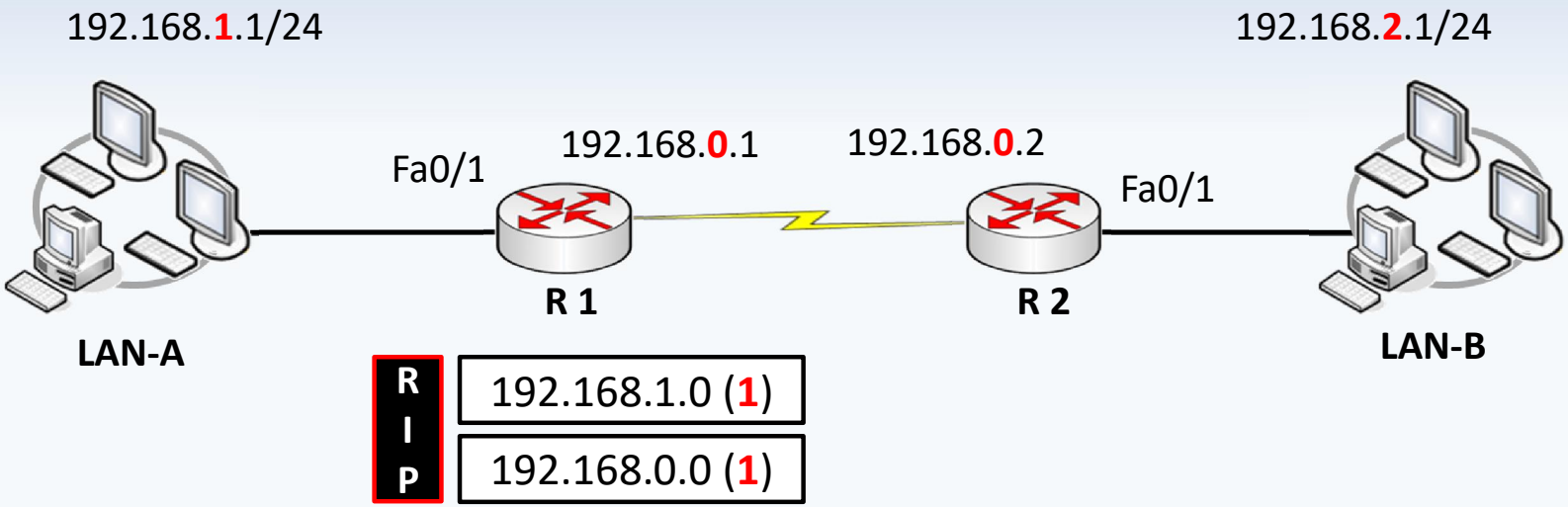


Vector-distance routing, RIP

- Initially, each router only knows the cost of its own links to its direct neighbors. This is the initial vector
- Each router will swap its initial vector with all of its neighbors
- After a certain number of iterations, each router will know the cost to each destination,
- Works well on small networks
- A router operating in active mode sends a broadcast message every 30 seconds to indicate that it knows a route and the cost of the route.

The initial vector presents the directly connected networks

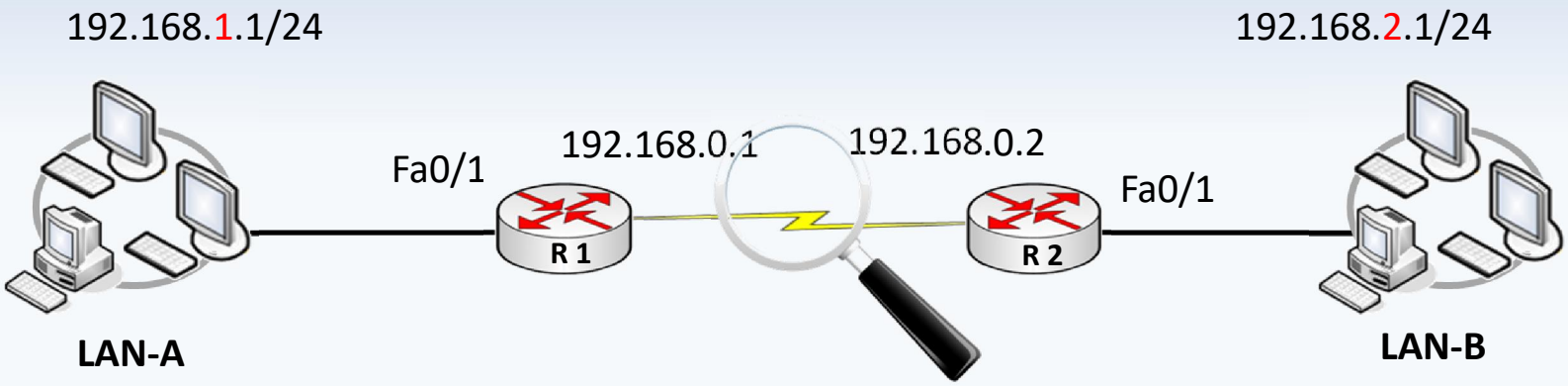
Principle



	Destination	P	M
C	192.168.1.0/24	e1	0
C	192.168.0.0/24	s0	0

	Destination	P	M
C	192.168.2.0/24	e1	0
C	192.168.0.0/24	s0	0
R	192.168.1.0/24	192.168.0.1	1

Zoom on RIPv1



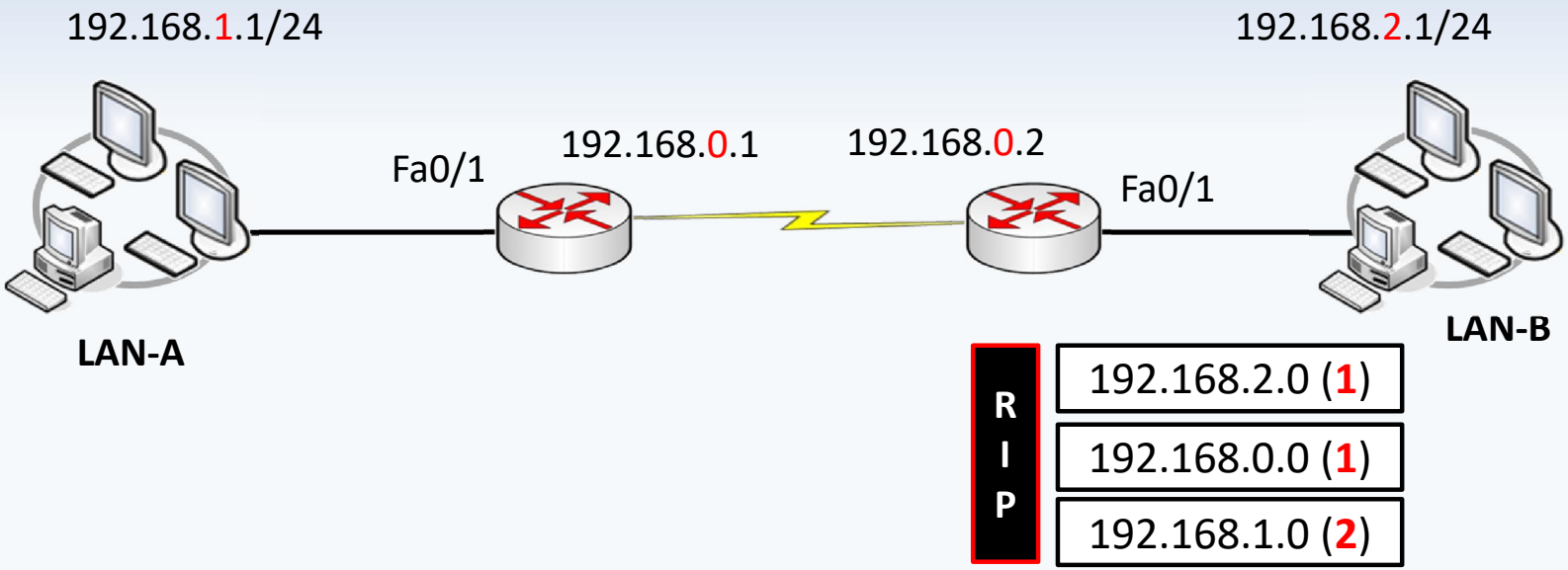

```
⊞ Frame 1014: 56 bytes on wire (448 bits), 56 bytes captured (448 bits)
⊞ Point-to-Point Protocol
⊞ Internet Protocol, Src: 192.168.0.1 (192.168.0.1), Dst: 255.255.255.255 (255.255.255.255)
⊞ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
⊟ Routing Information Protocol
    Command: Response (2)
    Version: RIPv1 (1)
  ⊟ IP Address: 192.168.1.0, Metric: 1
      Address Family: IP (2)
      IP Address: 192.168.1.0 (192.168.1.0)
      Metric: 1
```

Un message RIPv1 de R1 vers R2 capturé par *Wireshark*

Initialization

- Each RIP interface sends a Request message, requests entire routing tables.
- A Response message is sent from RIP neighbors.
  - If New Route: Installs it in the routing tables.
  - If an existing road: replaces it if the metric is better....

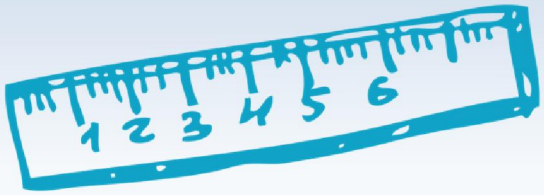
Convergence of the RIPv1



	Destination	P	M
C	192.168.1.0/24	e0	0
C	192.168.0.0/24	s0	0
R	192.168.2.0/24	192.168.0.2	1

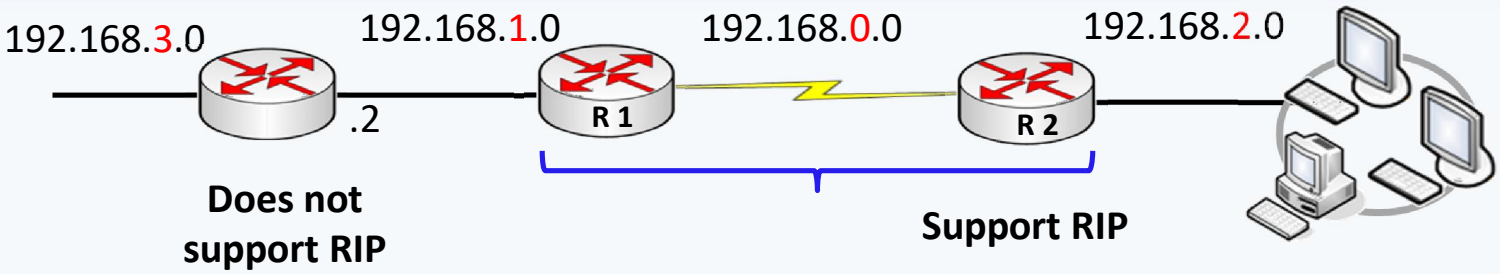
	Destination	P	M
C	192.168.2.0/24	e0	0
C	192.168.0.0/24	s0	0
R	192.168.1.0/24	192.168.0.1	1

Convergence of the RIP



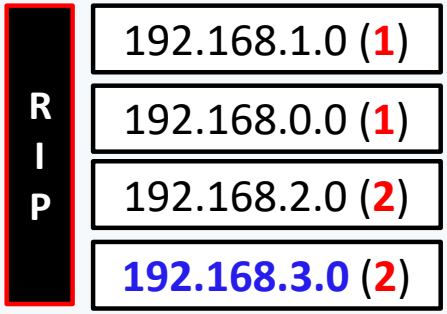
Convergence is when **all** routers have the **same routing information**

Redistribution



R 1

	Destination	P	M
C	192.168.1.0/24	E0	0
C	192.168.0.0/24	S0	0
R	192.168.2.0/24	192.168.0.2	1
S	192.168.3.0	192.168.1.2	1



MSG RIP from R 1 to R2

The use of a routing protocol to advertise routes learned by another routing protocol, or static routes, is called **redistribution**.

Number of iterations

- There will be convergence when there is no more updating, i.e. each node knows the whole network.
- After each iteration, each node exchanges its distance vector with its direct neighbors
- Example:

After how many iterations does the RIP algorithm converge?

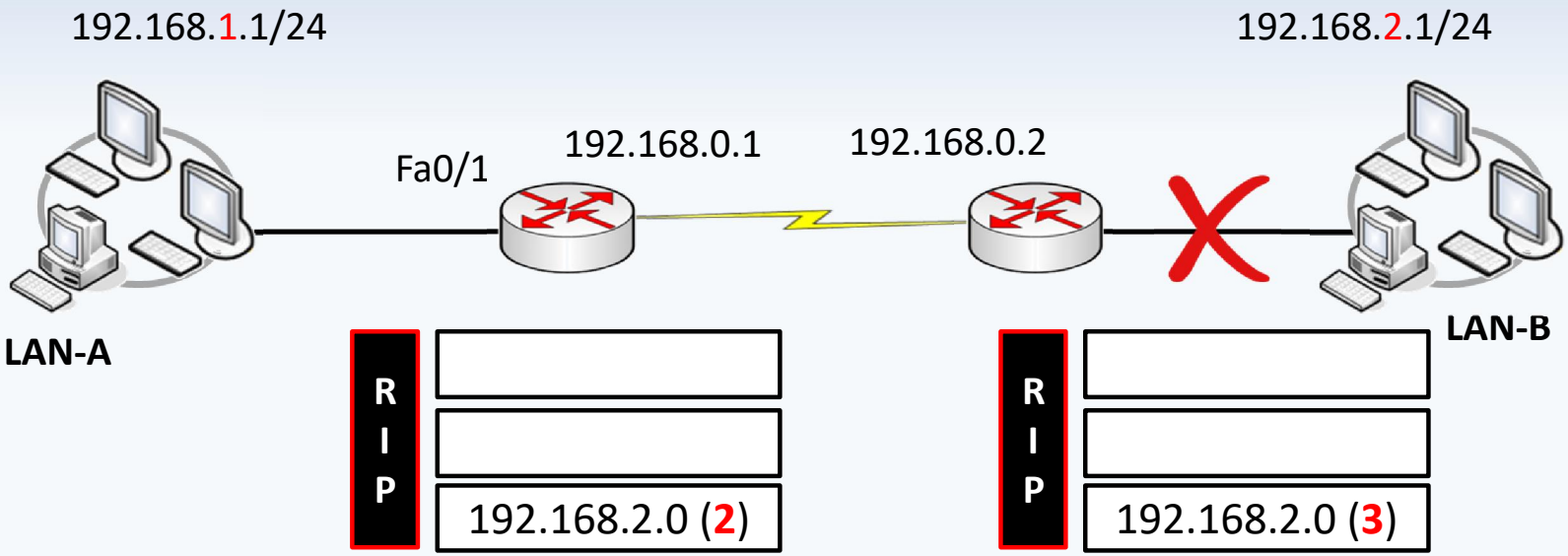
Converges after 3 iterations

1st iteration

2nd iteration

3rd iteration

Routing loop



!! Routing Loop

Métrique →∞

	Destination	P	M
R	192.168.2.0/24	192.168.0.1	2
C	192.168.0.0/24	s0	0
R	192.168.1.0/24	192.168.0.1	1

Routing Loop (2)

Definition :

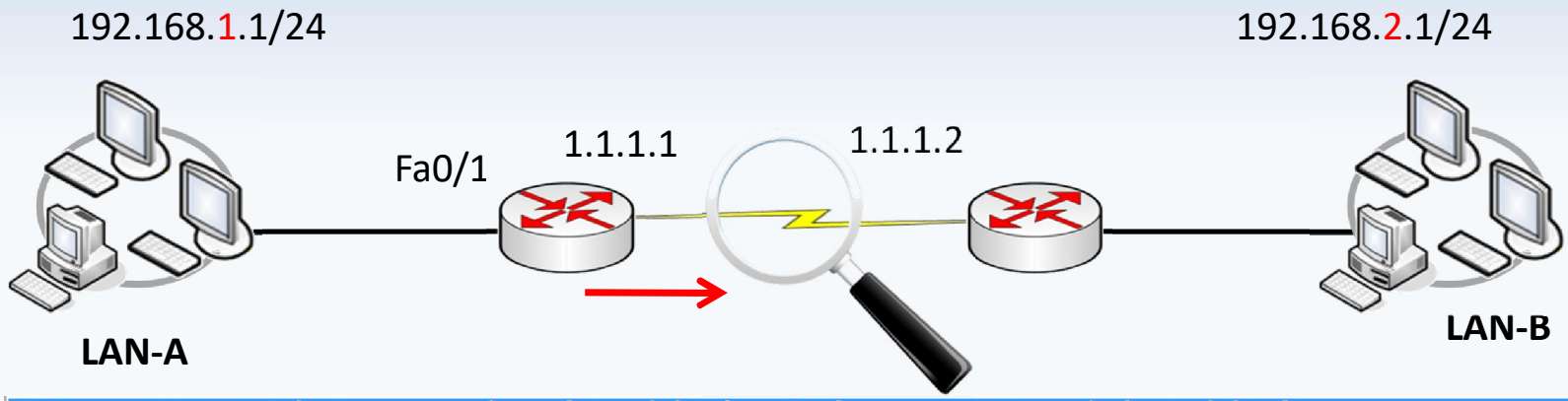
A routing loop is a route, broadcast for packets that never reach their destination: they repeatedly pass through the same series of nodes in the network.

Solution 1: Limit the metric

To avoid loops and limit convergence time, a router's visibility is limited to **15 hops**, a metric of 16 represents an unreachable route.

- Max metric = 15
- Metric = 16 => route not reachable

Solution 2: Split horizon




```
Frame 18: 56 bytes on wire (448 bits), 56 bytes captured (448 bits)
Point-to-Point Protocol
Internet Protocol, Src: 1.1.1.1 (1.1.1.1), Dst: 255.255.255.255 (255.255.255.255)
User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
Routing Information Protocol
    Command: Response (2)
    Version: RIPv1 (1)
  IP Address: 192.168.1.0, Metric: 1
      Address Family: IP (2)
      IP Address: 192.168.1.0 (192.168.1.0)
      Metric: 1
```

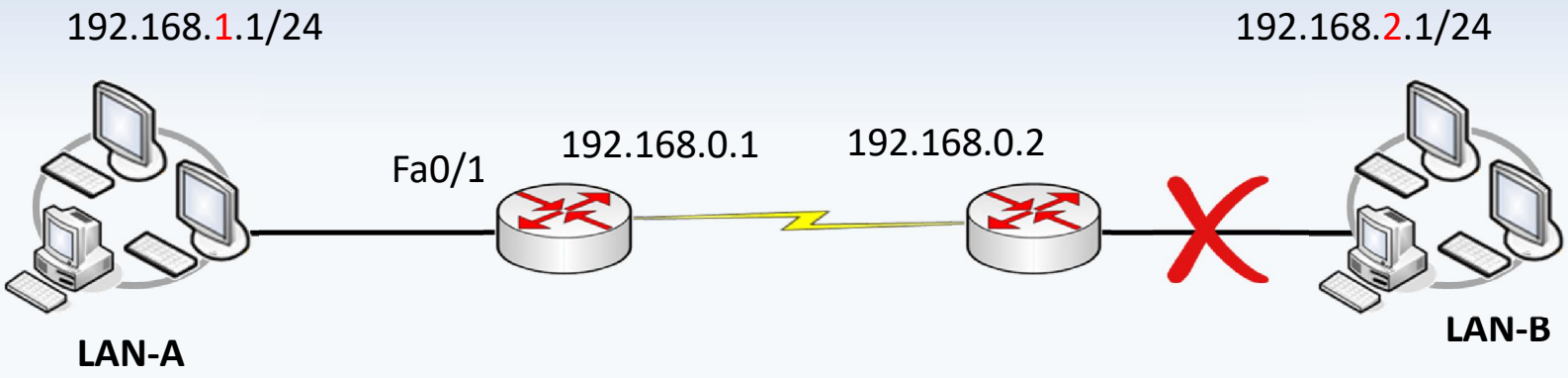
Never announce a route on an interface through which you have already learned this route.

Solution 2: Split horizon

Definition :

Split horizon *is a technique to accelerate network convergence. It consists of not transmitting routing information to the place from which it originates.*

Solution 3: Road Poisoning



	Destination	P	M
C	192.168.2.0/24	e0	**16**
C	192.168.0.0/24	s0	0
R	192.168.1.0/24	192.168.0.1	1

RIP 192.168.2.0 (**16**)

Solution 3: Road Poisoning

Definition :

If a router detects the loss of a connection to an adjacent network, it will immediately update the metric to that network to 16. Since this value is not allowed by RIP, routers will consider the route invalid.

important

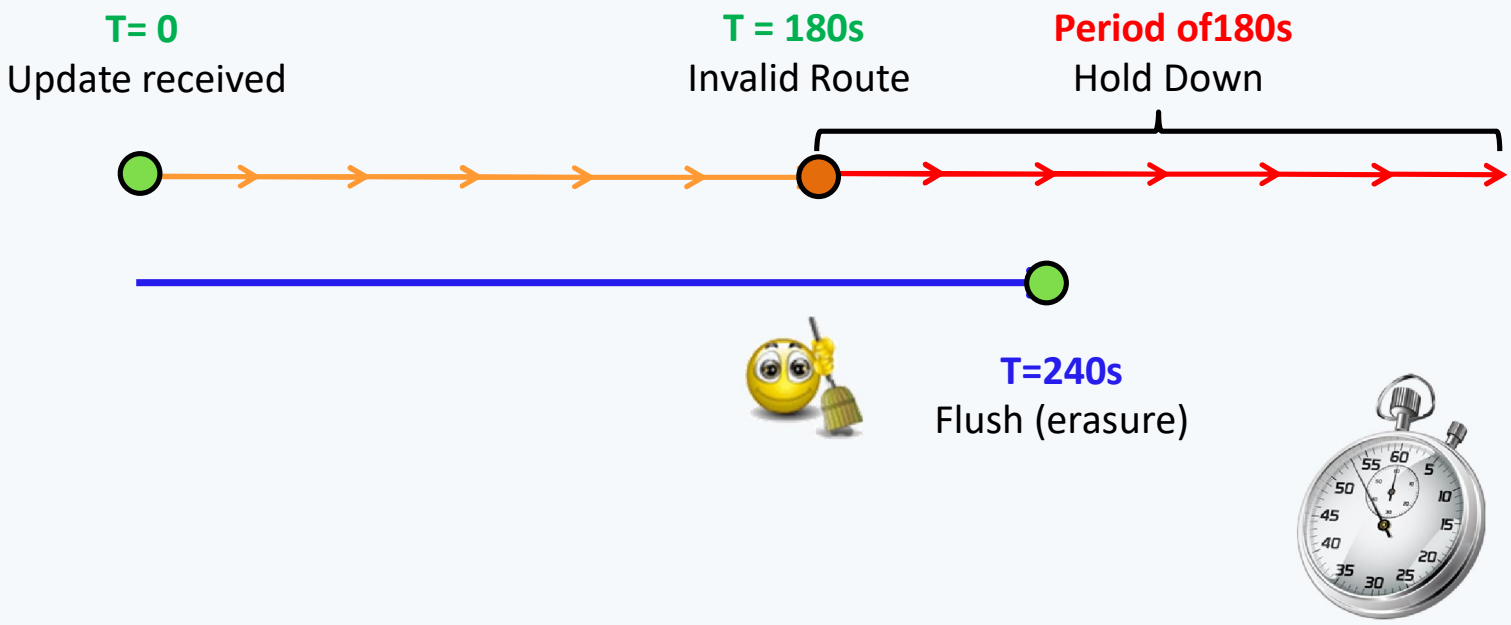
Solution 4: Using Timers

- **Invalid timer**: The time interval **after which** a route is marked and declared invalid if no updates are received. The default **value is 180s**. However, the route is saved until the end of the Flush timer.
- **Hold Down**: The time interval **during which** the router will reject all messages indicating that the route is reachable. However, the route is still used to forward packets. When the timer expires, routes announced by other sources are accepted and the route is no longer inaccessible. The default value **is 180** seconds.
- **Flush**: The time interval **after which** if there is no update, the route will be deleted. The default value is **240 sec**.

Example of a routing table

```
LAB-B#show ip route
Les codes : C - connecté, S - statique, I - IGRP, R - RIP, M -
mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, * - candidate
default
U - per-user static route
Gateway of last resort is not set
R 204.204.7.0/24 [120/1] via 199.6.13.2, 00:00:09, Serial0
R 223.8.151.0/24 [120/1] via 199.6.13.2, 00:00:09, Serial0
C 201.100.11.0/24 is directly connected, Serial1
C 219.17.100.0/24 is directly connected, Ethernet0
R 192.5.5.0/24 [120/1] via 201.100.11.1, 00:00:04, Serial1
C 199.6.13.0/24 is directly connected, Serial0
R 210.93.105.0/24 [120/2] via 199.6.13.2, 00:00:09, Serial0
```

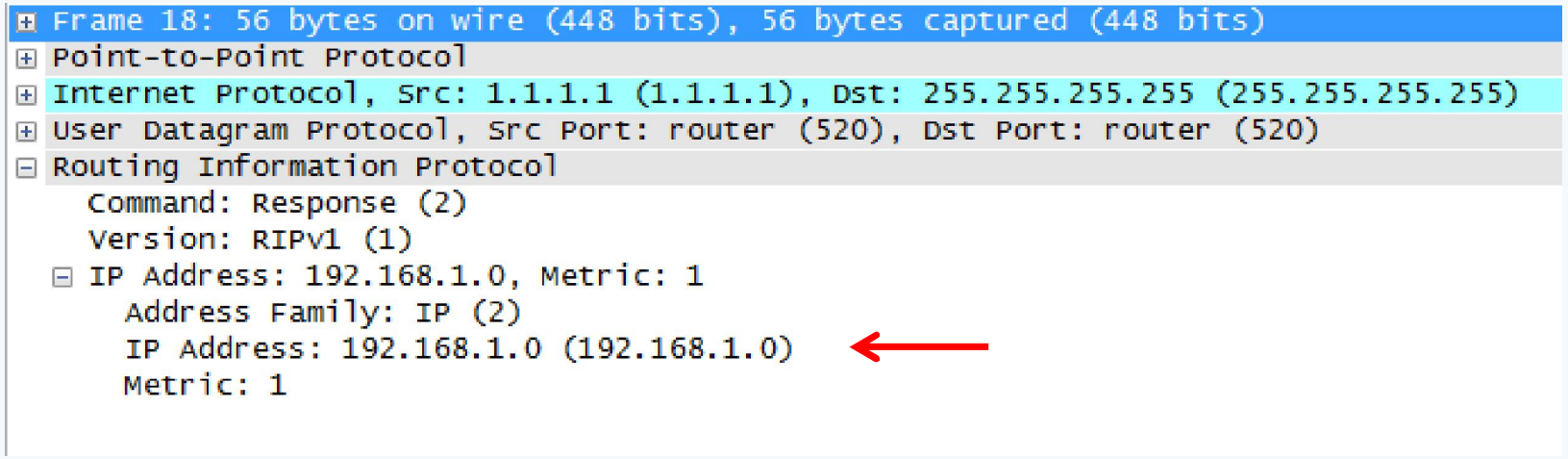
Temporization RIP



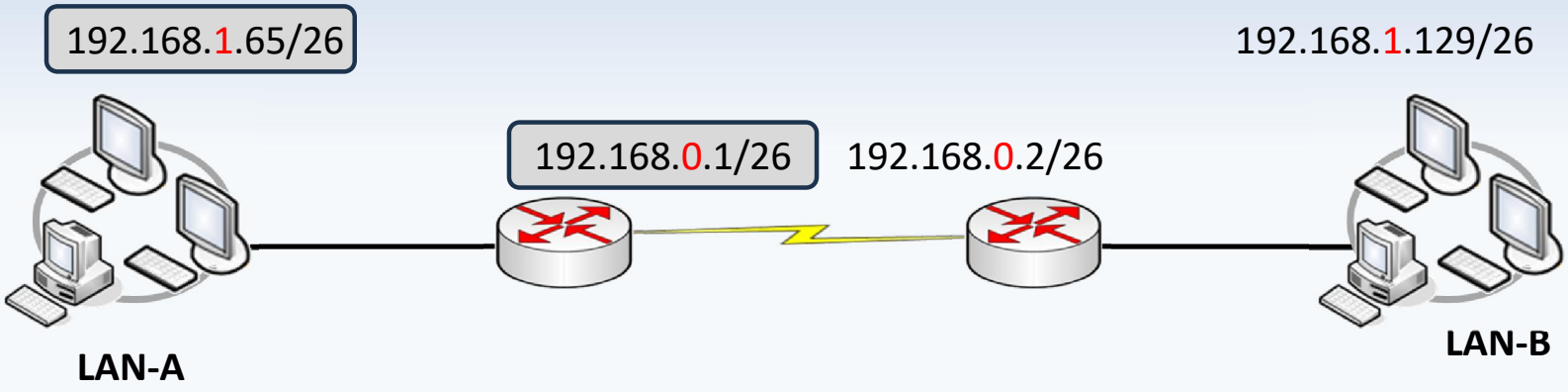
- **Administrative Distance:** Routing protocol preference. Fixed value for **each protocol.**
- **Metric:** The preference of the route within the **same routing protocol.**

RIPv1 and Classes

- RIPv1 does not send the IP address mask in the RIP message.
- It usually applies the default mask, Example 255.255.255.0 for 192.168.1.0.



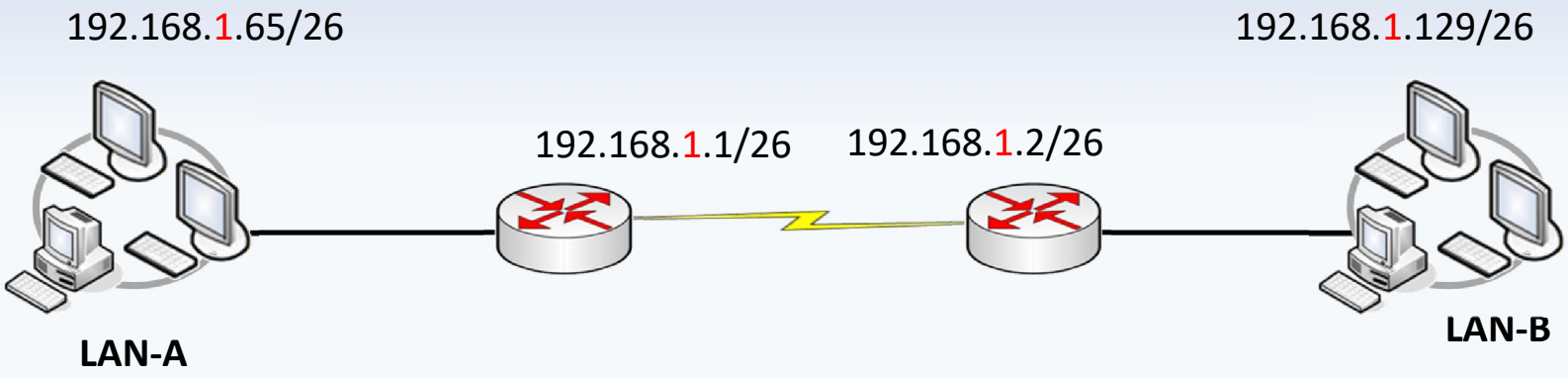
The algorithmRIP (1)



Emission The route to be advertised and the source interface have the same original network?
No → Announce the original network (summary).

RIP	192.168.1.**0** (1)

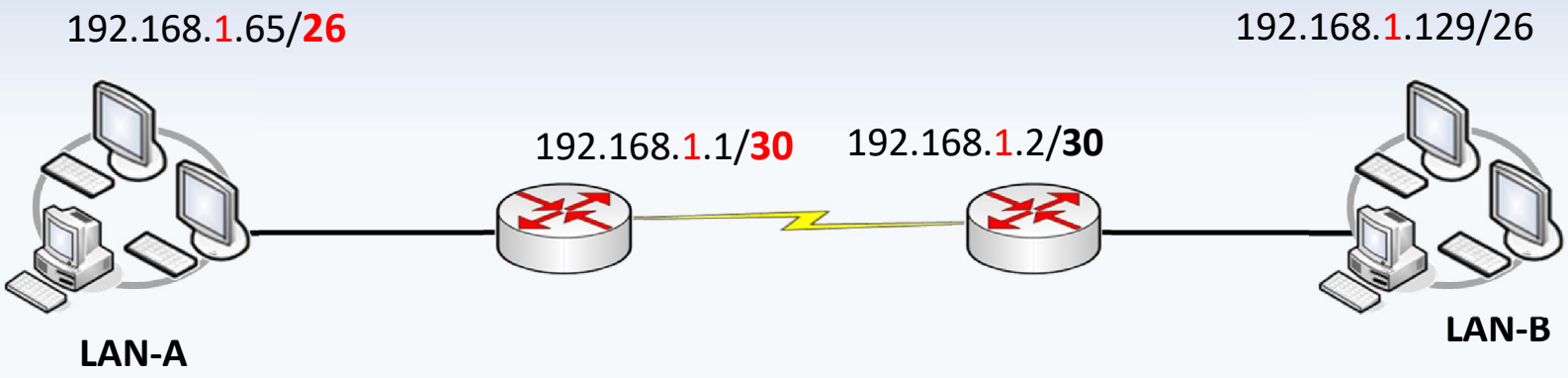
The algorithmRIP (2)



Emission The route to be advertised and the source interface have the same original network?
Yes
same mask?
Yes, → Advertise the subnet, not the summary

RIP	192.168.1.**64** (1)

L' algorithme RIP (3)



Emission The route to be advertised and the source interface have the same original network?

Yes

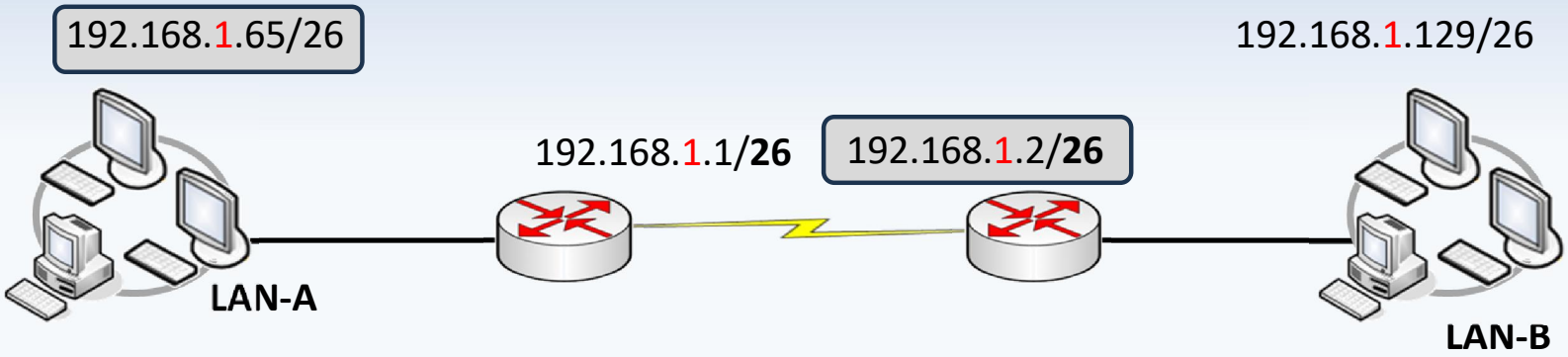
same mask?

No → does not send the RIP message

RIPv1 is not adequate for VLSM networks

VLSM: **V**ariable **L**ength **S**ubnet **M**ask

L’ algorithme RIP (4)

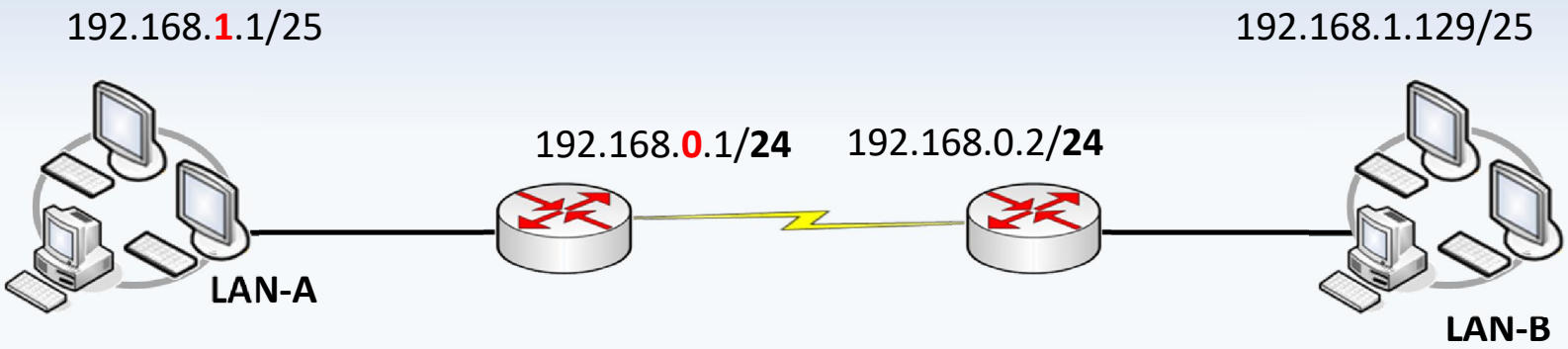


Reception The route to be advertised and the receiving interface have the same original network?
Yes → Apply the receiving interface mask.

RIP 192.168.1.**64** (1)

	Destination	P	M
C	192.168.1.128/26	e0	0
C	192.168.1.0/26	s0	0
R	192.168.1.**64/26**	192.168.1.1	1

The algorithmRIP (5)



Reception

The route to be advertised and the receiving interface have the same original network?

No

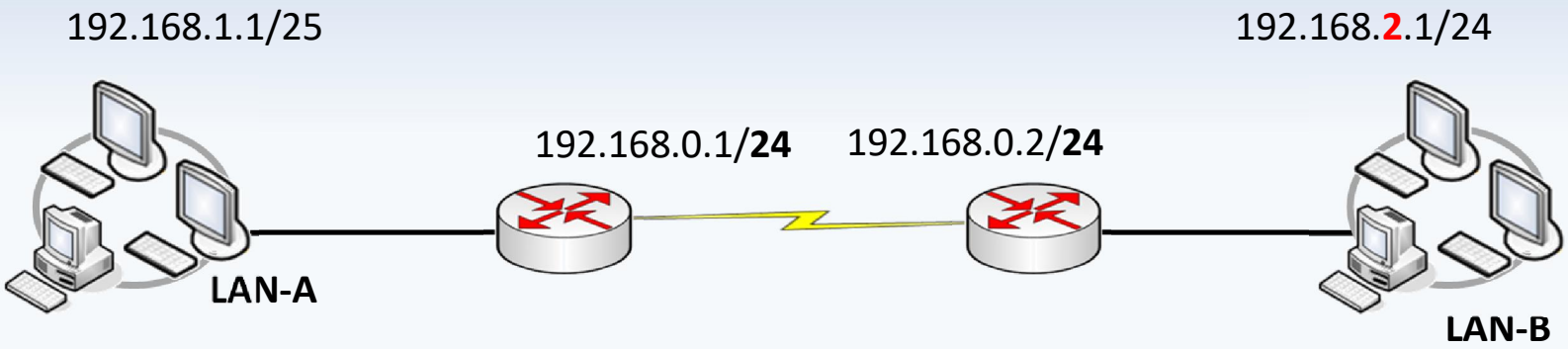
In the route table a subnet of the advertised route?

Yes, → Ignore the message

RIPv1 is not adequate for discontinuous networks

	Destination	P	M
C	192.168.1.128/25	e0	0
C	192.168.0.0/24	s0	0

The algorithmRIP (6)



Reception

The route to be advertised and the receiving interface have the same original network?

No

In the route table a subnet of the advertised route?

No, → Apply the default mask

RIP | 192.168.1.**0** (1)

RIPv1 is classful

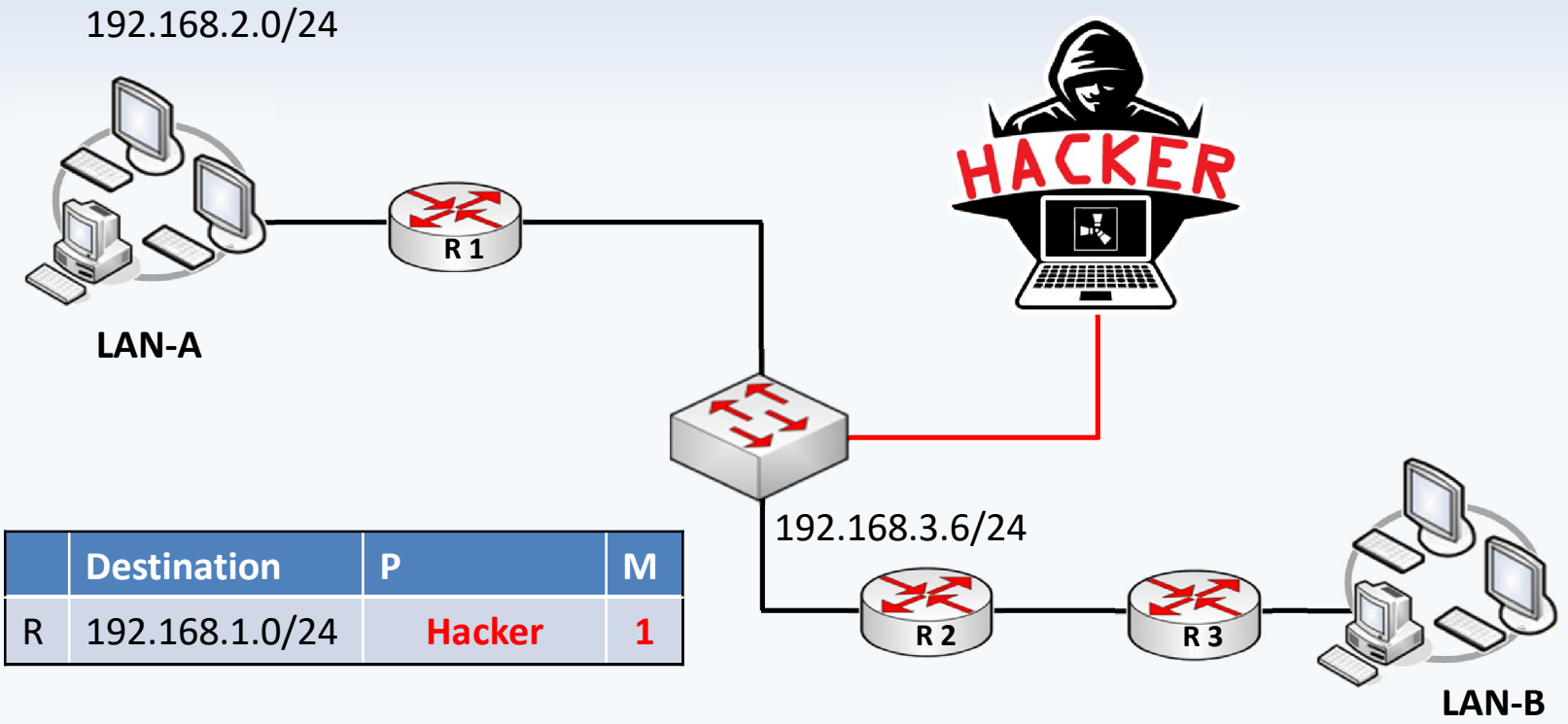
	Destination	P	M
C	192.168.2.0/24	e0	0
C	192.168.0.0/24	s0	0
R	192.168.1.1/24	192.168.0.1	1

- **Routed protocols** establish addresses to identify networks and computers within each network.
- **Routing protocols** construct routing tables of network addresses in order to identify paths between networks.

The RIP algorithm (summary)

- RIPv1 does not send subnet mask information in the update.
- A RIP v1-enabled router uses either the subnet mask configured on a local interface or the default subnet mask depending on the address class.
- Because of this limitation, RIPv1 networks cannot be discontinuous, nor can VLSM or over-networking.
- When a router is turned off, it sends its table with all its links to its neighbors with a metric of 16.
- When a new route is advertised, if the router already contains an entry of the same cost, it ignores this information....

The lack of authentication in RIP v1



	Destination	P	M
R	192.168.1.0/24	Hacker	1

RIP	192.168.1.0/**24** (2)	R2 vers R1
RIP	192.168.1.0/**24** (1)	H vers R1

The Disadvantages of RIP v1

- The distribution of tables every 30 s induces a lot of traffic and a significant convergence time (stabilization of the tables) which can be several minutes.
- There is no acknowledgment of receipt of messages. If a router does not receive any messages for 180 seconds, the silent route is declared unreachable.
- The messages are not authenticated. It is then possible for an attacker to generate RIP messages with costs such that all routes pass through the same router (congestion).

Classful protocols do not send the mask when updating the routing information. When he proposed, they didn't need to send the mask.

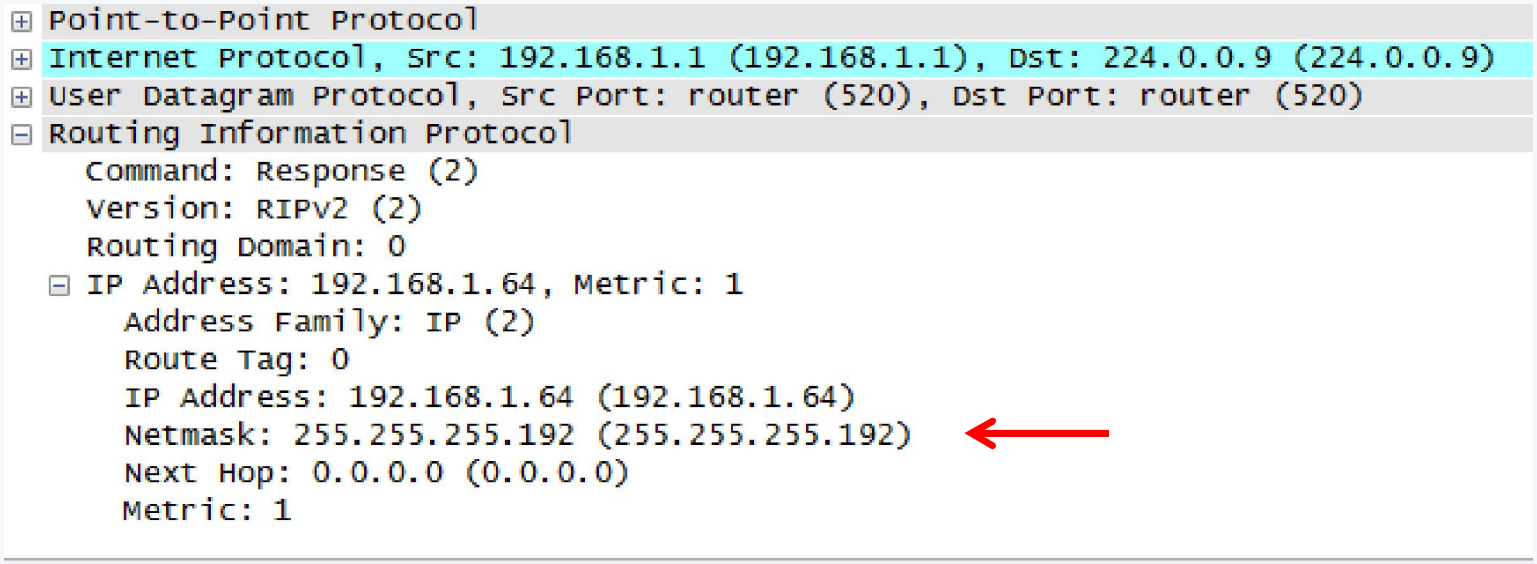
Distance vector routing

RIP v.2

Routing Information Protocol

RIPv2 and Classes

RIP v2 addresses some of the drawbacks of RIP v1 while remaining compatible with it. RIP v2 allows the Subnet Mask Field to be broadcast.



RIPv1 and RIPv2

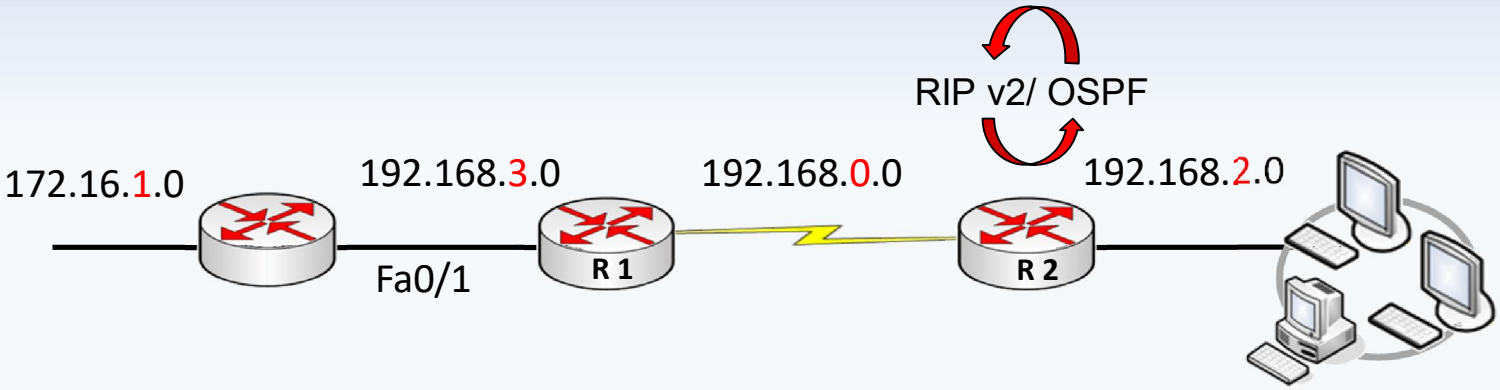
RIPv1

```
⊞ Frame 49: 56 bytes on wire (448 bits), 56 bytes captured (448 bits)
⊞ Point-to-Point Protocol
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168.0.1), Dst: 255.255.255.255 (255.255.255.255)
⊞ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
⊟ Routing Information Protocol
    Command: Response (2)
    Version: RIPv1 (1)
  ⊟ IP Address: 192.168.1.0, Metric: 1
      Address Family: IP (2)
      IP Address: 192.168.1.0 (192.168.1.0)
      Metric: 1
```

RIPv2

```
⊞ Frame 27: 56 bytes on wire (448 bits), 56 bytes captured (448 bits)
⊞ Point-to-Point Protocol
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168.0.1), Dst: 224.0.0.9 (224.0.0.9)
⊞ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
⊟ Routing Information Protocol
    Command: Response (2)
    Version: RIPv2 (2)
  ⊟ IP Address: 192.168.1.0, Metric: 1
      Address Family: IP (2)
      Route Tag: 0
      IP Address: 192.168.1.0 (192.168.1.0)
      Netmask: 255.255.255.0 (255.255.255.0)
      Next Hop: 0.0.0.0 (0.0.0.0)
      Metric: 1
```

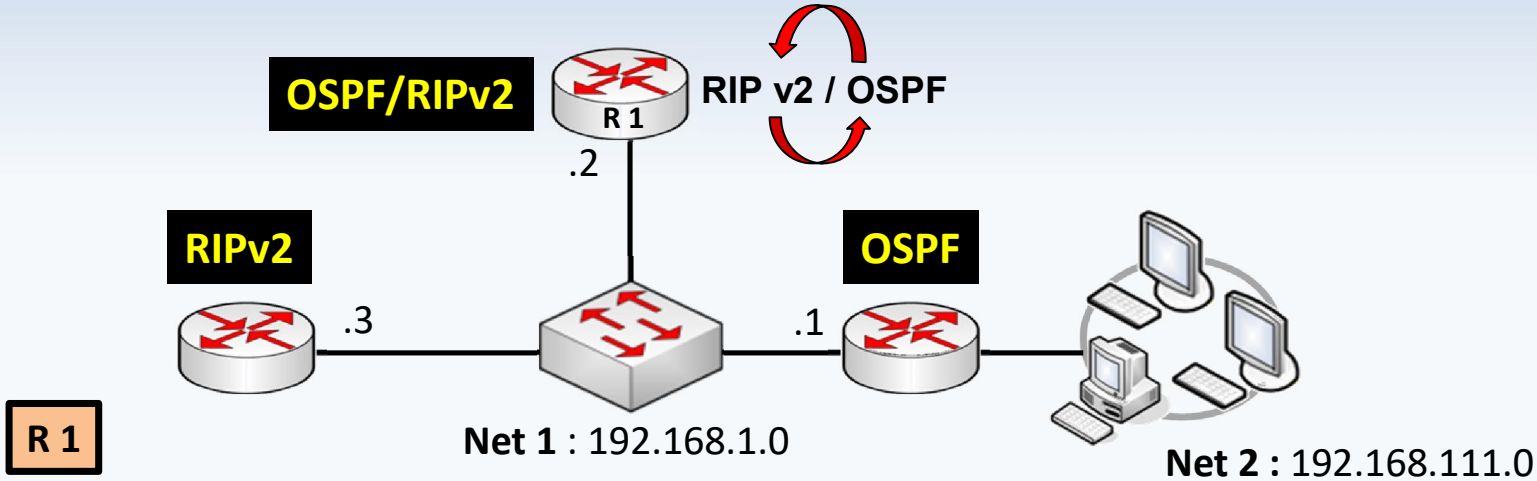
The notion of Tag in RIPv2



- Marking a routes with a number
- value: 0 → 65,535
- During redistribution


```
Routing Information Protocol
    Command: Response (2)
    Version: RIPv2 (2)
  IP Address: 172.16.1.0, Metric: 1
      Address Family: IP (2)
      Route Tag: 5
      IP Address: 172.16.1.0 (172.16.1.0)
      Netmask: 255.255.255.0 (255.255.255.0)
      Next Hop: 0.0.0.0 (0.0.0.0)
      Metric: 1
```

Next Hop



R 1

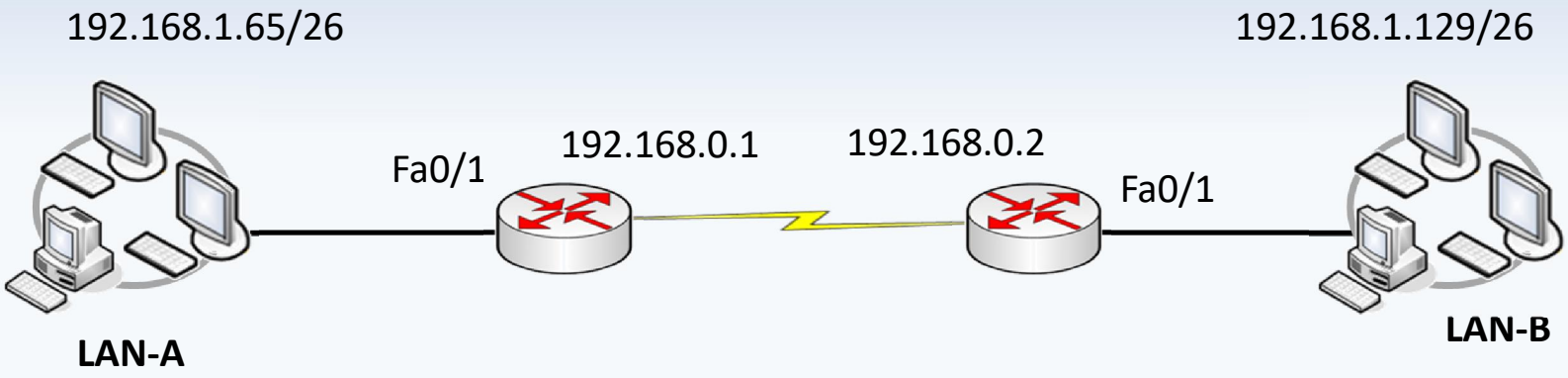
```
Routing Information Protocol
    Command: Response (2)
    Version: RIPv2 (2)
  IP Address: 192.168.55.0, Metric: 1
      Address Family: IP (2)
      Route Tag: 0
      IP Address: 192.168.55.0 (192.168.55.0)
      Netmask: 255.255.255.0 (255.255.255.0)
      Next Hop: 0.0.0.0 (0.0.0.0)
      Metric: 1
  IP Address: 192.168.111.0, Metric: 1
      Address Family: IP (2)
      Route Tag: 0
      IP Address: 192.168.111.0 (192.168.111.0)
      Netmask: 255.255.255.0 (255.255.255.0)
      Next Hop: 192.168.1.1 (192.168.1.1)
      Metric: 1
```

Characteristics

RIPv2 makes the following changes to RIPv1:

- Support for classless routing.
- The broadcast of the netmask in routing updates.
- The support ofV.L.S.M.
- Routing updates are delivered via multicast. (With RIPv1 the updates are done in broadcast)
- Authentication (not confidentiality) of the routing update source by plaintext (active by default), or ciphertext.
- RIP-v2 compensates for many of the shortcomings of RIP-v1, its range, still limited to 15 hops, makes it suitable for small and medium networks.
- The use of route tags to differentiate between routes learned by other routing protocols and redistributed in RIP

Convergence of the RIPv2



RIP 192.168.1.64**/26** (1)

	Destination	P	M
C	192.168.1.64/26	e0	0
C	192.168.0.0/24	S0	0
R	192.168.1.128/**26**	192.168.0.2	1

RIP 192.168.1.128**/26** (1)

	Destination	P	M
C	192.168.1.128/26	e0	0
C	192.168.0.0/24	s0	0
R	192.168.1.64/**26**	192.168.0.1	1

Convergence of the RIP v2

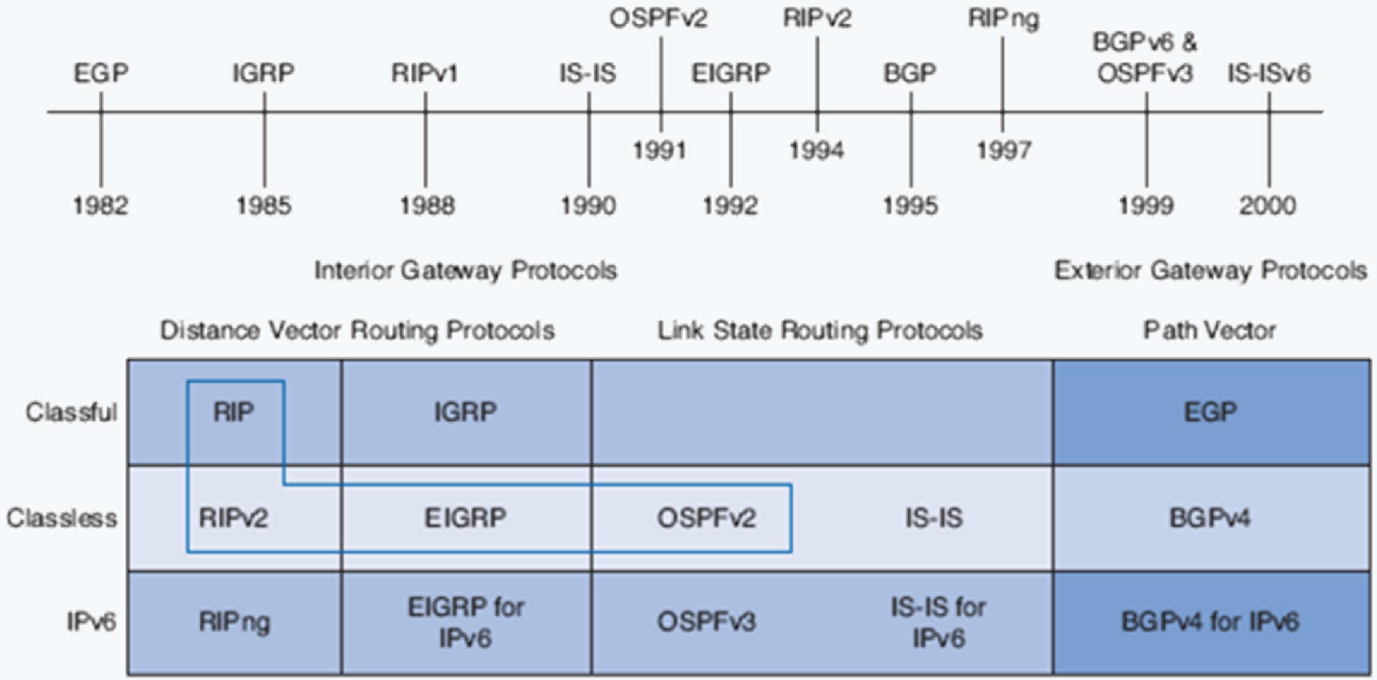
If two routes have the **same AD** as well as the same metrics, the routing protocol will **balance the load**, which means that data will be sent over **each link**.

Chapter 06

Dynamic routing

Part: 02

Routing Protocols



Comparison of routing protocols

Distance vector	Link Status
Algorithme Bellman-Ford (RIP)	Algorithme Dijkstra (OSPF)
Easy to set up	Required skills
Sharing route tables	Link sharing
Flat Networks	organized into zones
Slower convergence	Fast convergence, load balancing
Constrained topologies	Topologies complexes et larges
Bandwidth-hungry	Relativement discret
Low consumption of RAM and CPU	RAM and CPU consuming
Regular Broadcast/Multicast updates	Immediate updates
No signalling	Reliable and connected signalling

Link State Routing

The protocol O.S.P.F

Open Shortest Path First

Presentation

- The **O**pen **S**hortest **P**ath **F**irst (OSPF) protocol was developed by the IETF to address the need for a non-proprietary, highly functional Internal Gateway Protocol (IGP) in the TCP/IP protocol stack.
- The current version of OSPFv2 is described in RFC 2328 (1998).
- A version 3 is defined in RFC 5340 which allows the use of OSPF in an IPv6 network (2008) and even to embed IPv4 routes.
- Strategy: Send information about its neighbors to all nodes (not just neighbors): Distribution of the network topology and cost of each link to all routers..

Presentation

- OSPF is a link-state routing protocol.
- Uses a single parameter as a metric, which is the cost of the interface (**Cost**)
- Uses the addresses of (**multicast**) 224.0.0.5 and 224.0.0.6
- Its main competitor in homogeneous enterprise infrastructures is EIGRP, which owns Cisco.
- Routers collect the full cost of links from all possible paths. The best routes are then integrated into the routing table.
- They converge very quickly.
- This is called hierarchical routing,...

Principle

The operation of the OSPF goes through the following steps:

1. **Discovery** of OSPF neighbors, through "hello" messages and establishment of **contiguities**,
2. Synchronization of databases (LSDB, Link State Data Base) through the exchange of **L.S.A,** 'Link-State Advertisement'", Each router receives an LSA message, it in turn broadcasts it to its neighbors (LSA flood), until all routers have all the LSAs,
3. Build the topology table from the received LSAs, transform its LSDB into a directed graph,
4. Run the SPF algorithm ("Dijkstra")
5. Build the route table: From the SPF tree, the best paths are inserted into the route table
6. Maintain Topology: Hello and Broadcast "Link States" at Each Change....

Tables OSPF

Each OSPF router stores routing and topology information in three tables:

- Neighbor table—Stores information about OSPF neighbors
- Topology table—Stores the topology structure of a network
- Routing table—Stores the best routes

1- Discovery of the neighbors

- Routers running OSPF must establish **neighbor relationships** before exchanging routes.
- Neighbors do not exchange routing tables. Instead, they exchange information about the network topology.
- OSPF neighbors are dynamically discovered by sending **hello packets** on each interface:
  - 10 sec for LAN/PPI.
  - 30 sec. NBMA (No Broadcast Multi Acces), Exp: FrameRelay.
- IP Destination, multicast: 224.0.0.5 ....

What's in a Hello?

The fields in hello packets that **must be identical** on both routers in order for the routers to become neighbors, are essentially as follows:

- Subnet
- area ID
- dead-interval timers
- Hello interval timers
- Authentication

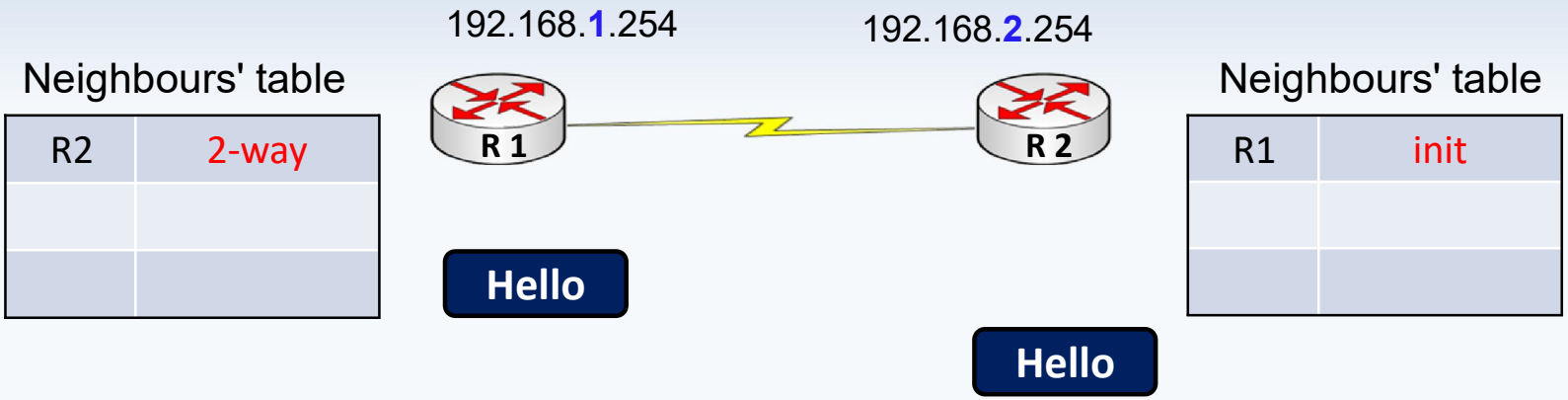
In addition to the configuration parameters, the hello message contains in particular the list of neighbors.

The Hello Packet

Hello packets are used to do the following:

- **Discover** OSPF neighbors and establish neighborhood adjacencies.
- **Advertise** the parameters that two routers must agree on to become neighbors
- **Choose** DR and BDR on multi-access networks such as Ethernet....

Discovery of the neighbours



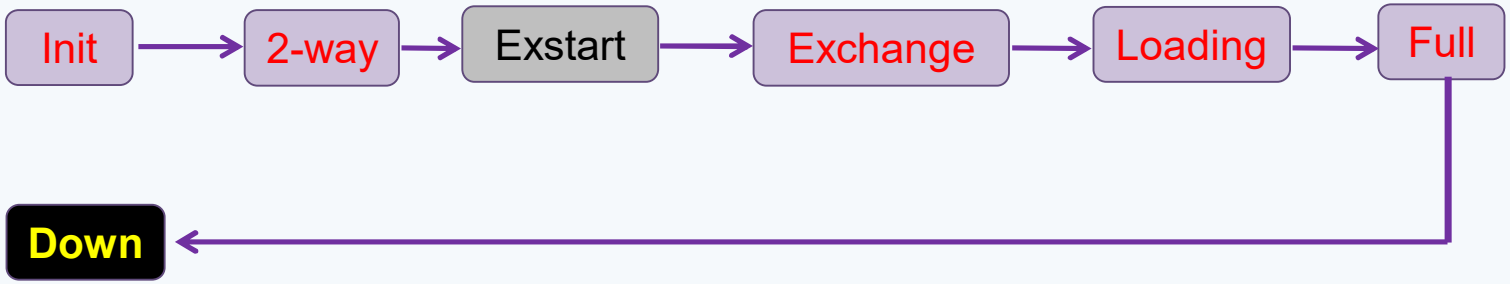

```
Open Shortest Path First
  OSPF Header
      OSPF Version: 2
      Message Type: Hello Packet (1)
      Packet Length: 48
      Source OSPF Router: 192.168.2.254 (192.168.2.254)
      Area ID: 0.0.0.0 (Backbone)
      Packet Checksum: 0x664d [correct]
      Auth Type: Null
      Auth Data (none)
  OSPF Hello Packet
      Network Mask: 255.255.255.0
      Hello Interval: 10 seconds
    ⊞ Options: 0x12 (L, E)
      Router Priority: 1
      Router Dead Interval: 40 seconds
      Designated Router: 0.0.0.0
      Backup Designated Router: 0.0.0.0
      Active Neighbor: 192.168.1.254
  OSPF LLS Data Block
```

Hello Message
de R2 vers R1

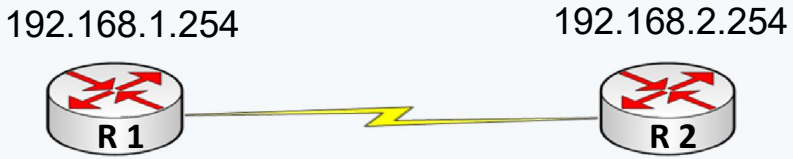
Possible states of OSPF neighbors

- **init**: A Hello message is received from another OSPF router
- **2-way**: The neighbor received the Hello message and replied with a Hello message of his own
- **Exstart**: (case of DR and BDR)
- **Exchange**: Database Descriptor (DBD) packets are exchanged. Routers will use this information to see which **LSAs** (Link-State Advertisments) need to be exchanged.
- **Loading**: loading link states: LSR (Link State Request) and LSU (Link State Update)
- **Full**: End of synchronization, both routers have the database synchronized and are completely adjacent to each other....

Possible states of OSPF neighbors



RouterDeadInterval = 4*HelloInterval



Adjacents

2- Synchronization of databases

Once two neighboring routers are in 2-way mode, they can exchange information on their LSDBs: Link State DataBase. This database contains all the details about the router, especially the interfaces, in terms of:

- IP/mask
- Metric (throughput)
- Type

Link Type	Description
1	Point-to-point connection to another router.
2	Connection to **transit network.**
3	Connection to **stub network**.
4	Virtual Link

2- Database Synchronisation

192.168.1.2	192.168.1.1	OSPF	78 DB Description
192.168.1.2	192.168.1.1	OSPF	78 DB Description
192.168.1.2	224.0.0.5	OSPF	94 Hello Packet
192.168.1.1	224.0.0.5	OSPF	94 Hello Packet
192.168.1.2	192.168.1.1	OSPF	78 DB Description
192.168.1.1	192.168.1.2	OSPF	138 DB Description
192.168.1.2	192.168.1.1	OSPF	118 DB Description
192.168.1.2	192.168.1.1	OSPF	82 LS Request
192.168.1.1	192.168.1.2	OSPF	78 DB Description
192.168.1.1	192.168.1.2	OSPF	70 LS Request
192.168.1.1	192.168.1.2	OSPF	130 LS Update
192.168.1.2	192.168.1.1	OSPF	78 DB Description
192.168.1.2	192.168.1.1	OSPF	98 LS Update
192.168.1.1	192.168.1.2	OSPF	78 DB Description
192.168.1.2	224.0.0.5	OSPF	94 LS Update
192.168.1.1	224.0.0.5	OSPF	98 LS Update
192.168.1.2	224.0.0.5	OSPF	98 LS Update
192.168.1.2	224.0.0.5	OSPF	98 LS Acknowledge
192.168.1.1	224.0.0.5	OSPF	98 LS Acknowledge
192.168.1.2	224.0.0.5	OSPF	94 LS Update
192.168.1.1	192.168.1.2	OSPF	98 LS Update
192.168.1.2	192.168.1.1	OSPF	98 LS Update
192.168.1.1	224.0.0.5	OSPF	98 LS Acknowledge
192.168.1.2	224.0.0.5	OSPF	78 LS Acknowledge

Exchange: Preparing for synchronization (DB)

Loading: Loading of Link states

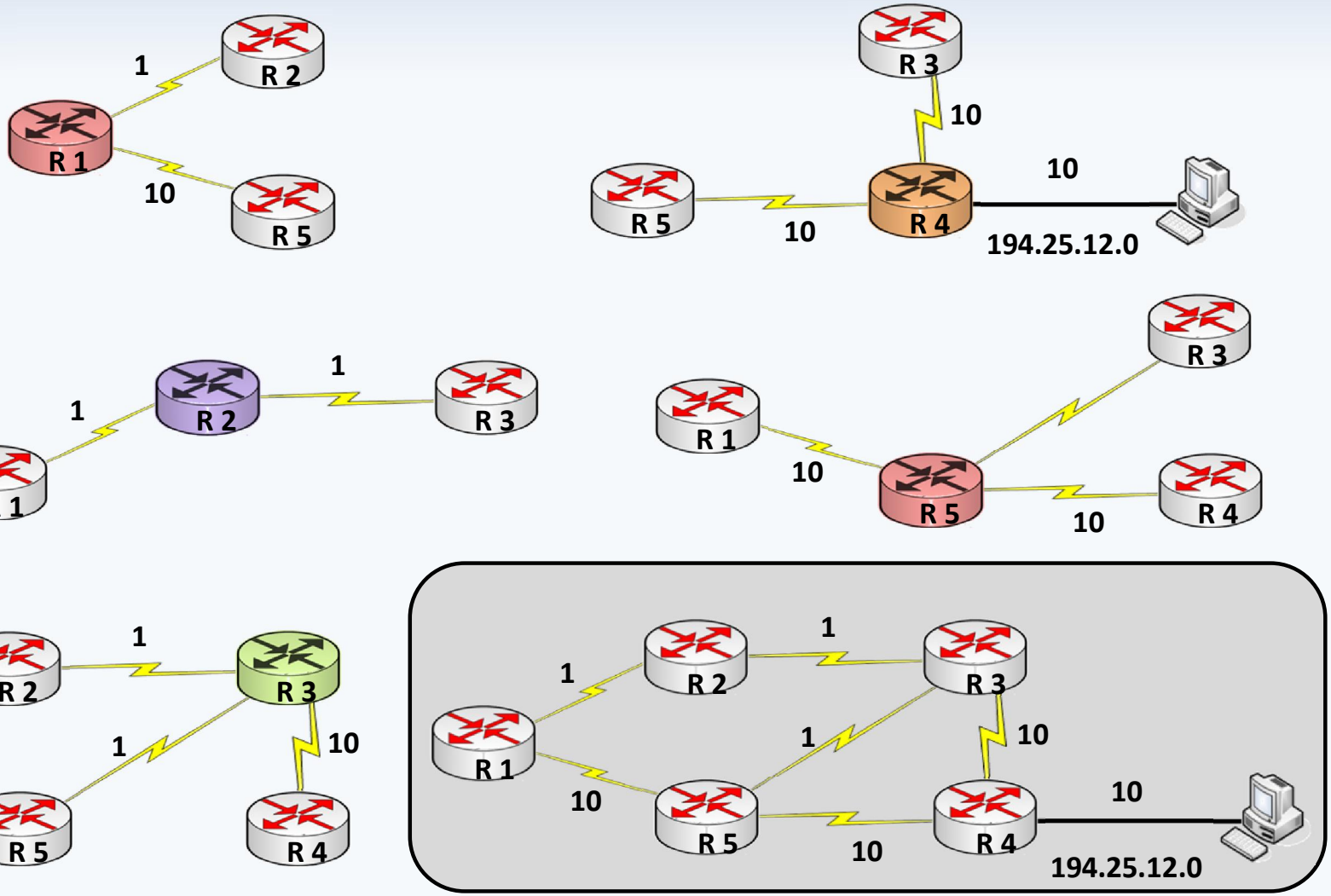
Full : End of thesynch

LSA, LSR, LSU and LSAck

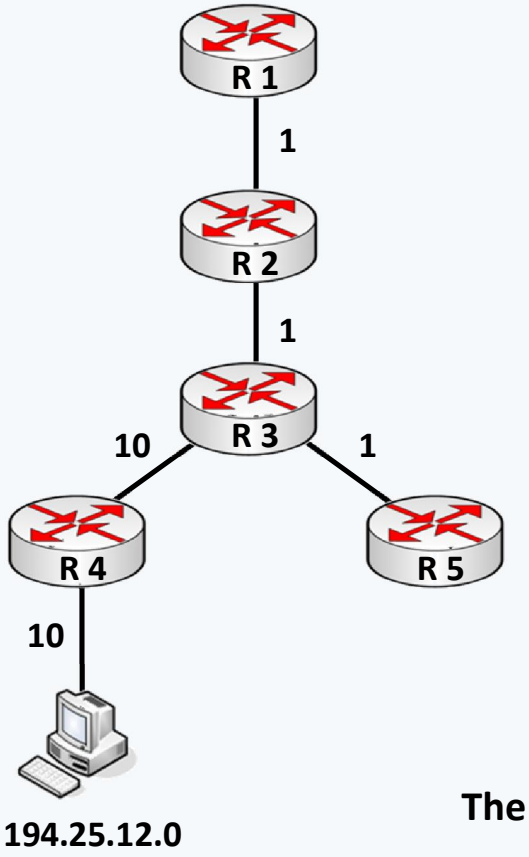
- Database Description (DBD): Routers use DBD packets to describe their own Link State Data Base (LSDB) databases for synchronization.
- Each router checks its database and sends an **LSR** (Link State Request) requesting all **LSAs** (Link-State Advertisments) not found in its table,
- The other router **responds with the LSU** (Link State Update) which contains all the requested LSAs,
- LSAck (Acknowledgment): When an LSU is received, the router sends a link state **acknowledgment** to confirm the receipt....

RIP : sending a large amount of information to only your neighbours,
OSPF : send a small piece of information (the neighbors) to all routers,

Example of transforming LSDBs into a directed graph

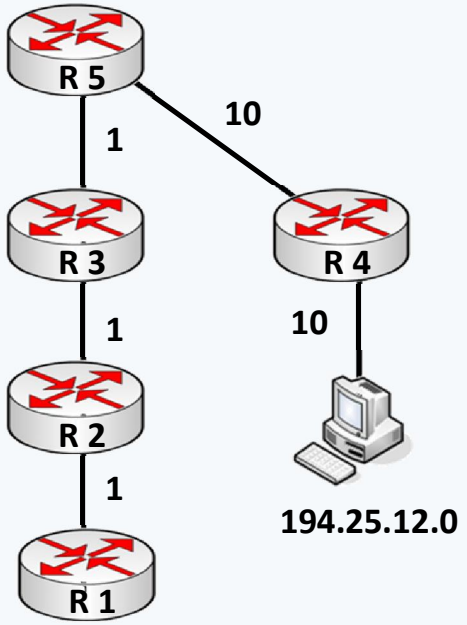


Example - The Election of the Best routes



From - to	Cost
R1, R2	1
R1, R5	10
R2, R3	1
R2, R1	1
R3, R2	10
R3, R4	10
R3, R5	1
R4, R5	10
R4, R3	10
R4, 192.168.1.0	10
R5, R4	10
R5, R3	1
R5, R1	10

The topology database(Network Map)



Example - Determining a Route Table

The topology database describes the network, but is not used directly for routing. The routing table is determined by applying the SPF algorithm on a topological basis.

- **On R1**, here is an excerpt from the routing table calculated by the OSPF about the 194.95.12.0 network:

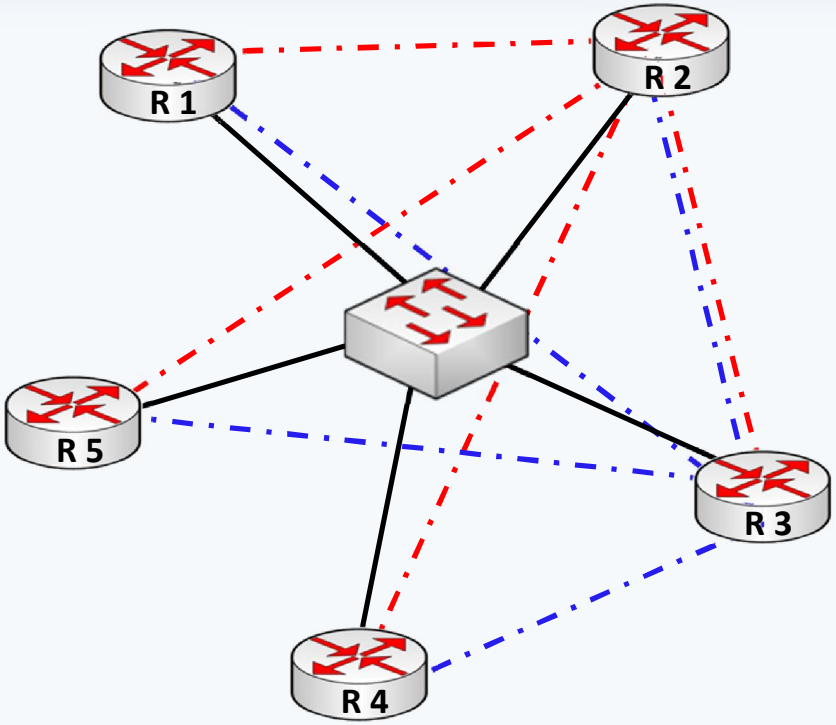
Destination Network	Means of attaining it	Cost
194.95.12.0	R2	22

- **On R5,** we will have the following excerpt:

Destination Network	Means of attaining it	Cost
194.95.12.0	R4	20

In OSPF, a router **calculates** the shortest paths to destinations **instead of** obtaining routing information through route advertisements.

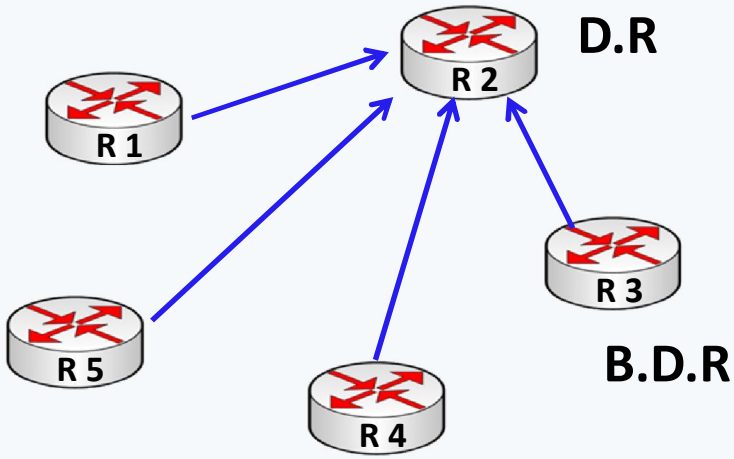
Diffusion problem



One of the fundamental characteristics of the OSPF is the limitation of dissemination to the strict minimum !!

Designated Router and Backup Designated Router

- To reduce the amount of OSPF traffic on transit networks, OSPF chooses a D.R and a backup D.R (B.D.R).
- The DR is responsible for updating all other OSPF routers when a change occurs in the network.
- The RLO monitors the RD and takes over if the current RD fails.
- All other routers become DROthers.
- Multicast address 224.0.0.6



DROthers routers remain in the **2way state** with each other. On the other hand, in the **full state** with the DR and the BDR.

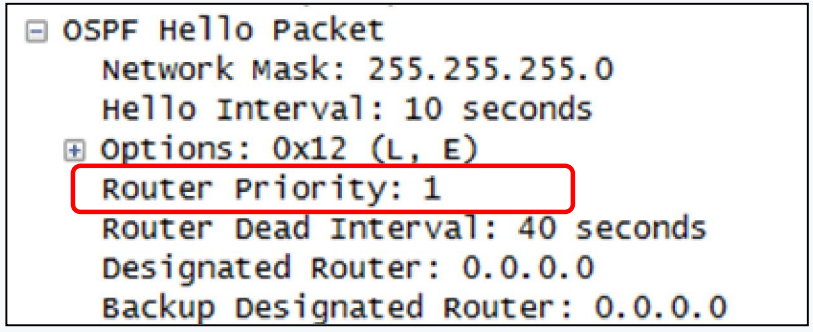
Election of the RD/BDR

- The DRs / BDRs are elected on the basis of the information contained in the OSPF Hello package.
- The Hello packet lists each router ID and a **priority value** (from 0 to 255).
- The one with a high priority is elected DR, with the BDR being the second priority.
- If the priority is the same, the highest RID (router ID) is used to elect a DR....

Election of the RD/BDR

Criterion 1: Priority (interface)

- The default value is 1
- The router with the highest priority is elected as the DR
- This value can be changed by the Admin
- By setting the priority to "0", the router is forced to be Drother...

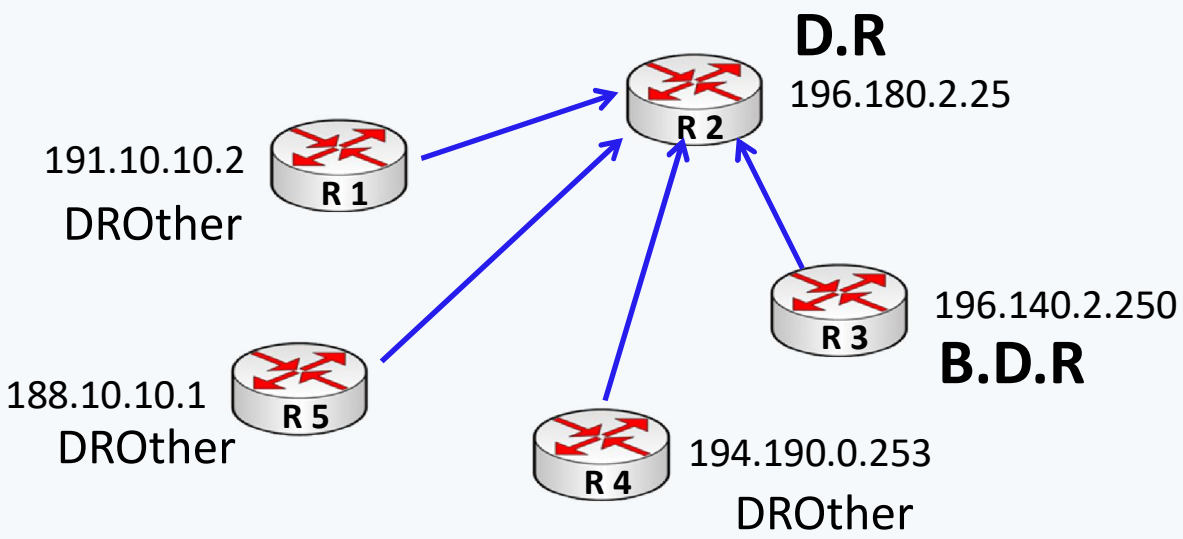


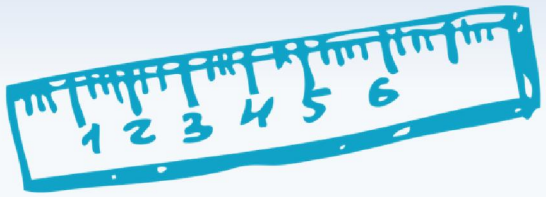
The identifier of a router

- Each OSPF router has an **identifier**
- The OSPF ID must be unique in an OSPF topology.
- A router identifier is determined using one of the following, **per priority**:
  - Manually, using the router-id command under the OSPF process (an IPv4 address)
  - The highest IP address of the router's loopback logical interfaces
  - The highest IP address of the router's physical interfaces....

Election of the RD/BDR

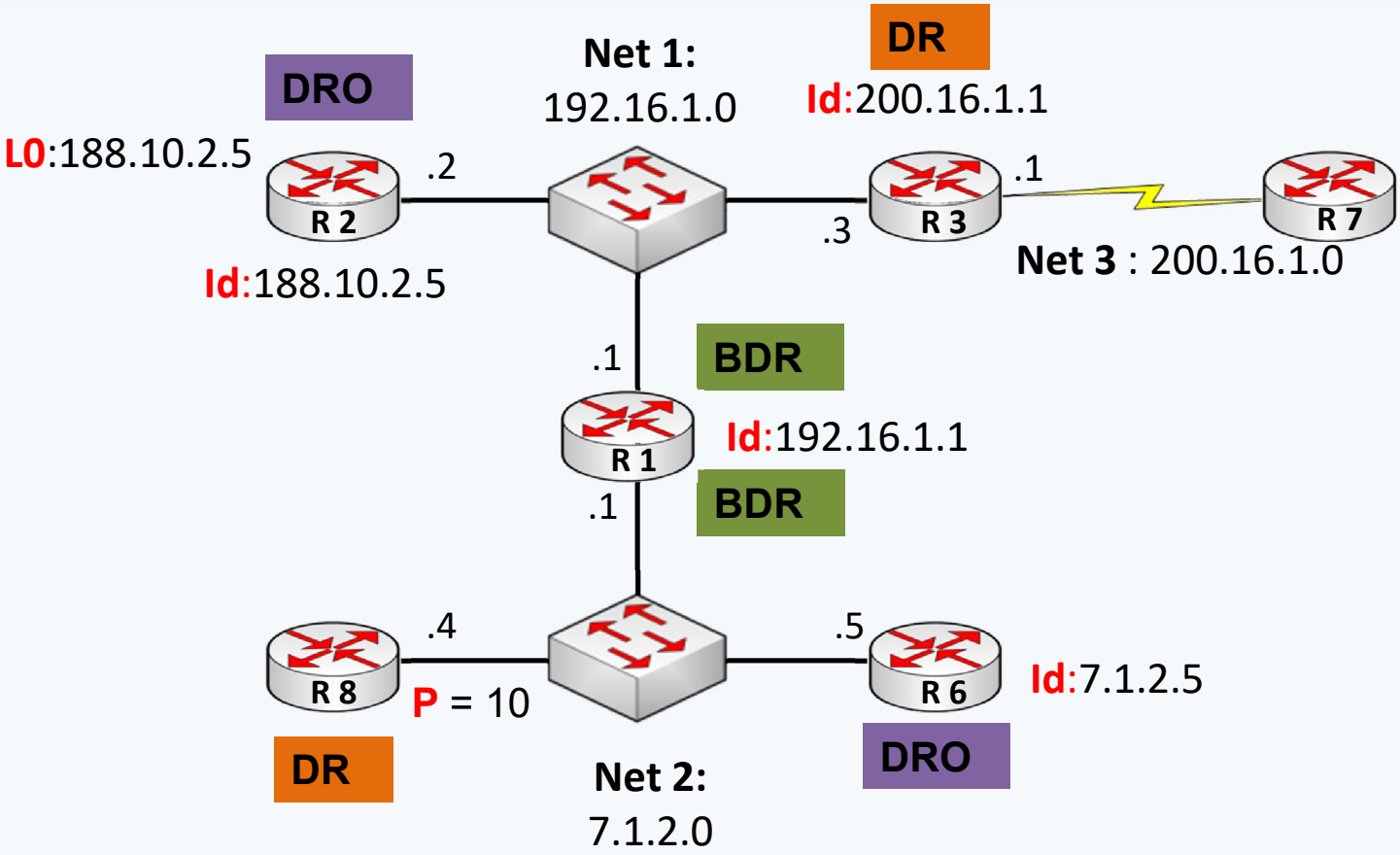
Criterion 2: Router ID





The logical address is preferred because it is **always UP**

Example:

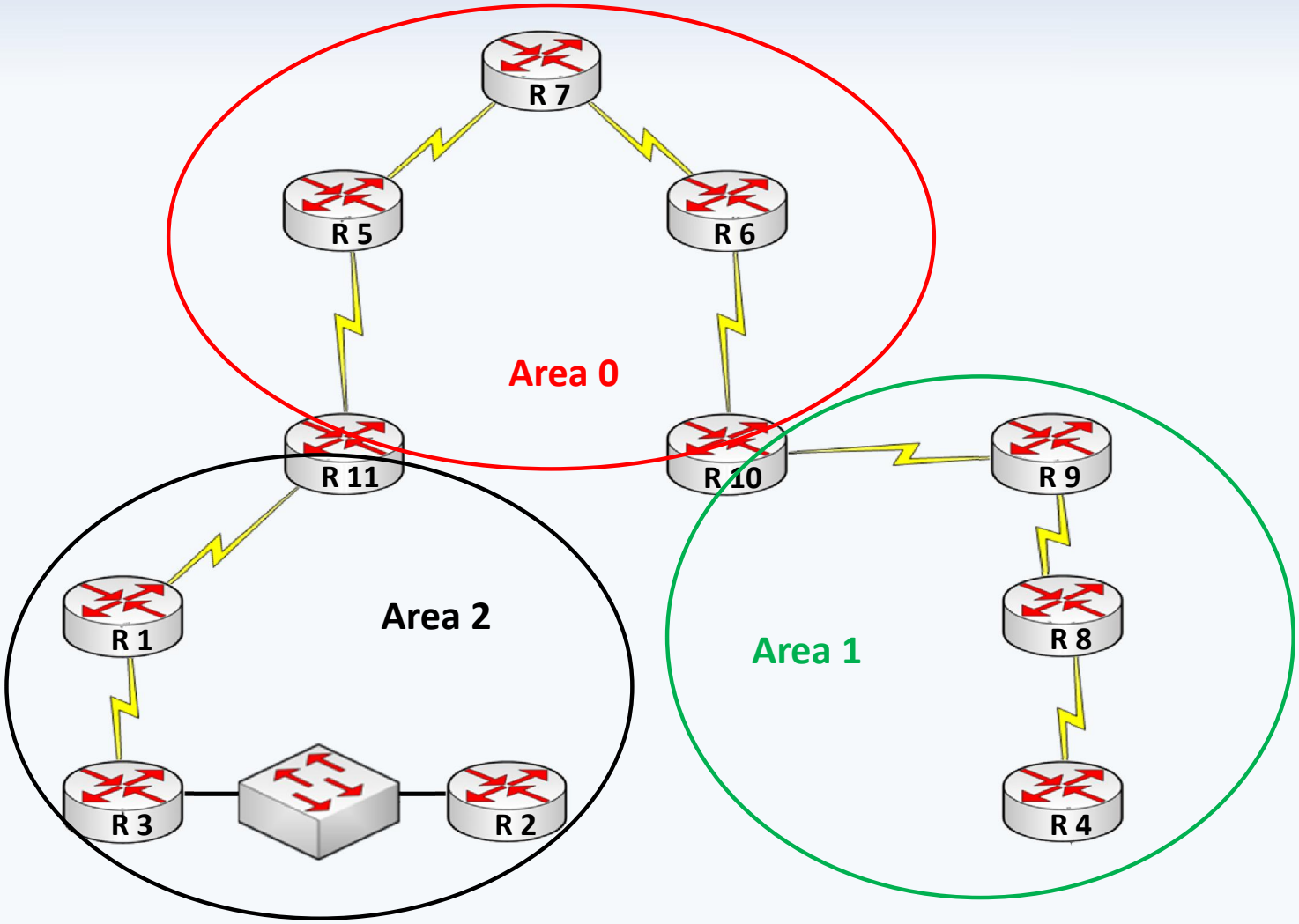

DRO
L0:188.10.2.5
Net 1:
192.16.1.0
DR
Id:200.16.1.1
.2
.3
.1
R 2
R 3
R 7
Id:188.10.2.5
Net 3 : 200.16.1.0
.1
BDR
R 1
Id:192.16.1.1
.1
BDR
.4
.5
R 8
R 6
P = 10
Id:7.1.2.5
DR
DRO
Net 2:
7.1.2.0

The Priority is by interface,
ID is per equipment (Global)

Design

- OSPF uses the concept of **areas**.
- A zone is a **logical grouping** of contiguous networks and routers.
- All routers in the same zone have **the same zone ID**, and also the same topology table, but they don't know the routers in the other zones.
- The main advantage is that the size of the topology and route table are **reduced**, it takes less time to run the OSPF, and **updates are reduced**.
- Each zone of the network must connect to the **backbone** area (**area 0**).

Design

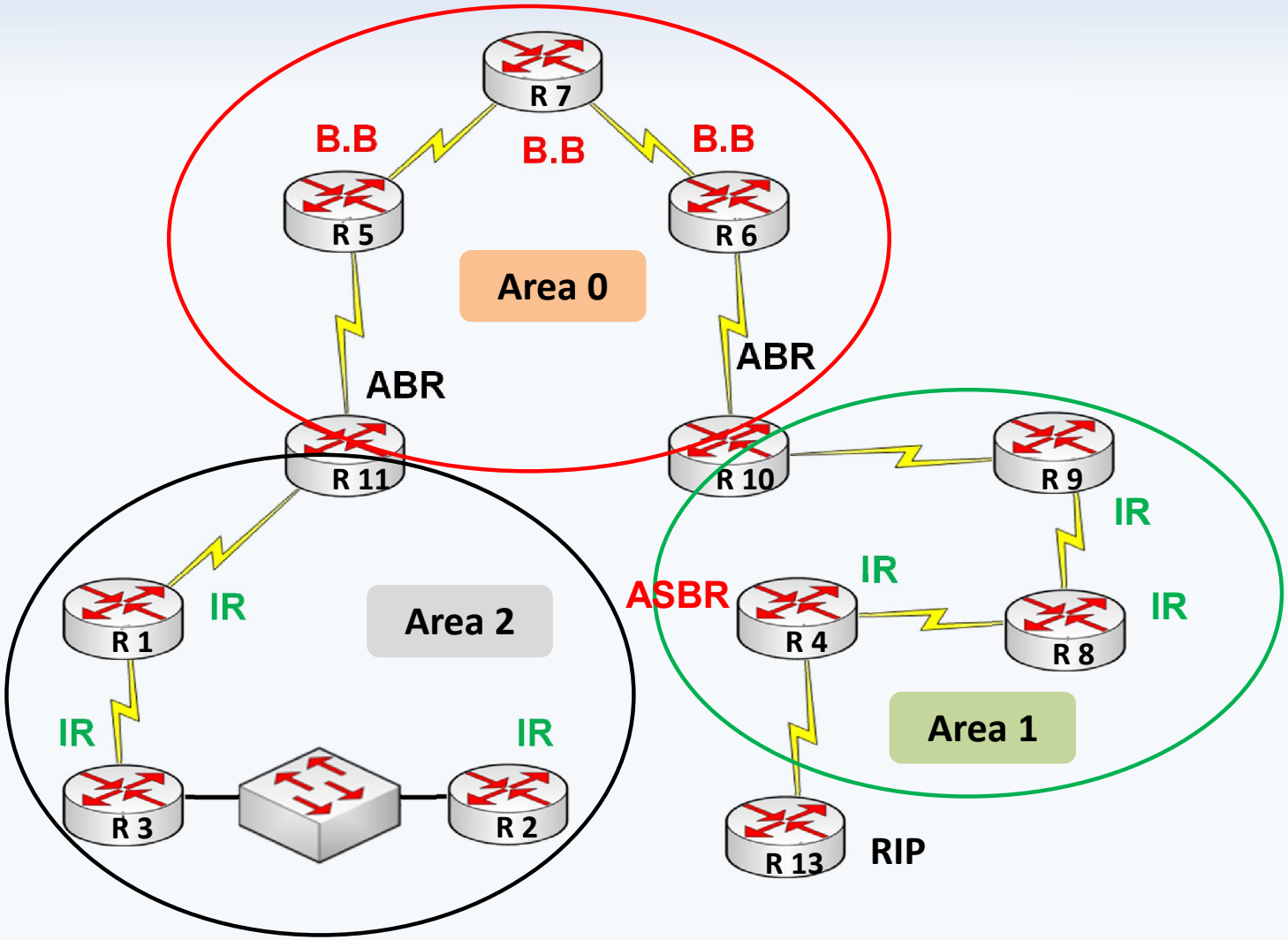


Definitions

- **BackBone (B.B)**: Each zone in the OSPF network must connect to the backbone area (area 0).
- **Area Border Router (A.B.R):** A router that has interfaces in more than one zone (for example, zone 0 and zone 1).

  Its role is to announce route summaries to neighbouring areas
- **Internal Router (I.R)**: Routers that share the same zone ID, and are not ABRs
- **Autonomous System Boundary Router (A.S.B.R)**: A router that connects an OSPF network to other routing domains (such as a RIP network)....

Definitions (2)



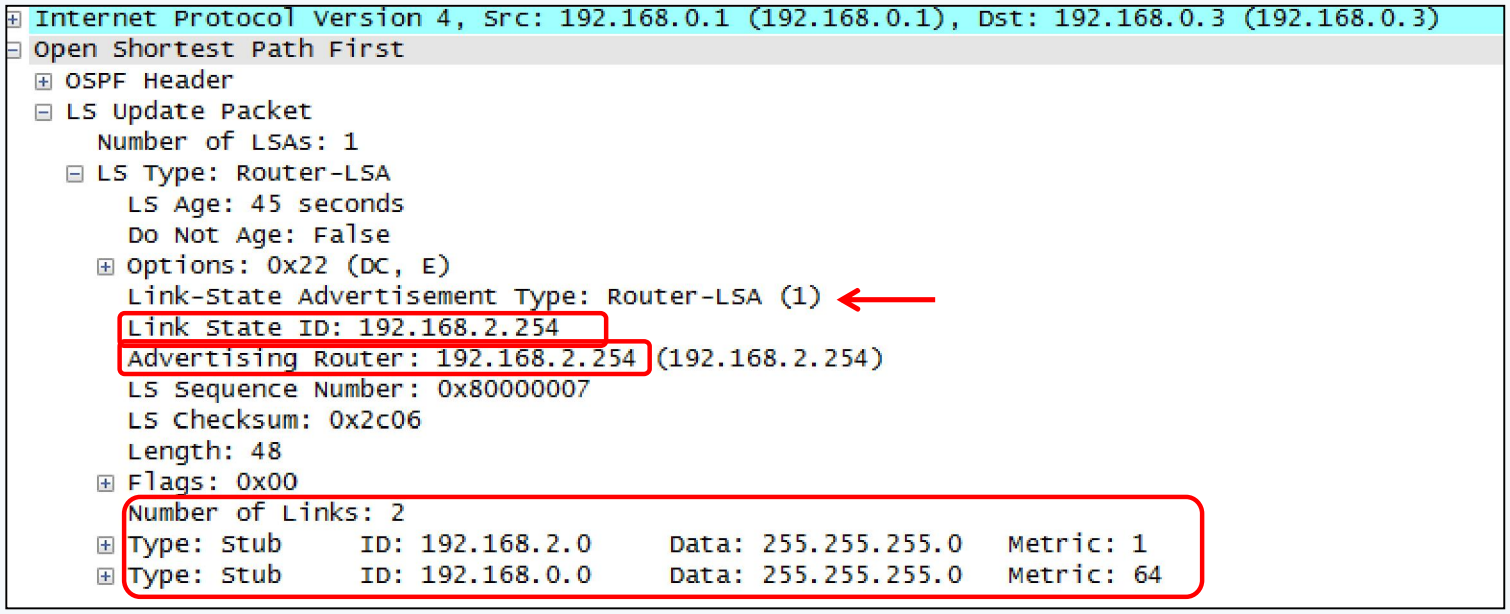
OSPF Topology, LSAs

LSAs (Link-State Advertisments) are found in the form of several types (LSA1, LSA2,... LSA5):

- LSA1 and LSA2: used to describe the states of router interfaces,
- LSA3, LSA4 and LSA5: used to exchange IP routes...
- Broadcast to all routers to have the same vision of the topology,

LSA type 1: (Router-LSA)

- Generated by all routers, to present to the network
- Describes the link states of a router (type, IP, metric, etc.)
- Broadcast within the **same area**



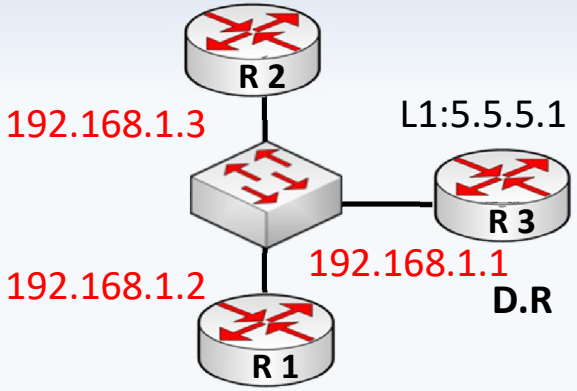
LSA type 1: (Router-LSA)

Receiving the LSA1 packet.

- The router consults the existing database.
- If the input (the link and its cost) is not present, add this input and broadcast the information to all neighbours except the LSA transmitter.
- If the input is present and the sequence number of the LSA is larger than the sequence number corresponding to the entry, modify the entry and broadcast the information to all neighbors except the re-transmitter of the LSA.
- If the entry is present and the sequence number of the LSA is less than or equal to the sequence number corresponding to the entry: do nothing.
- All of these link states form the link-state database. The topology table, or topology, is the same on all routers in an area.

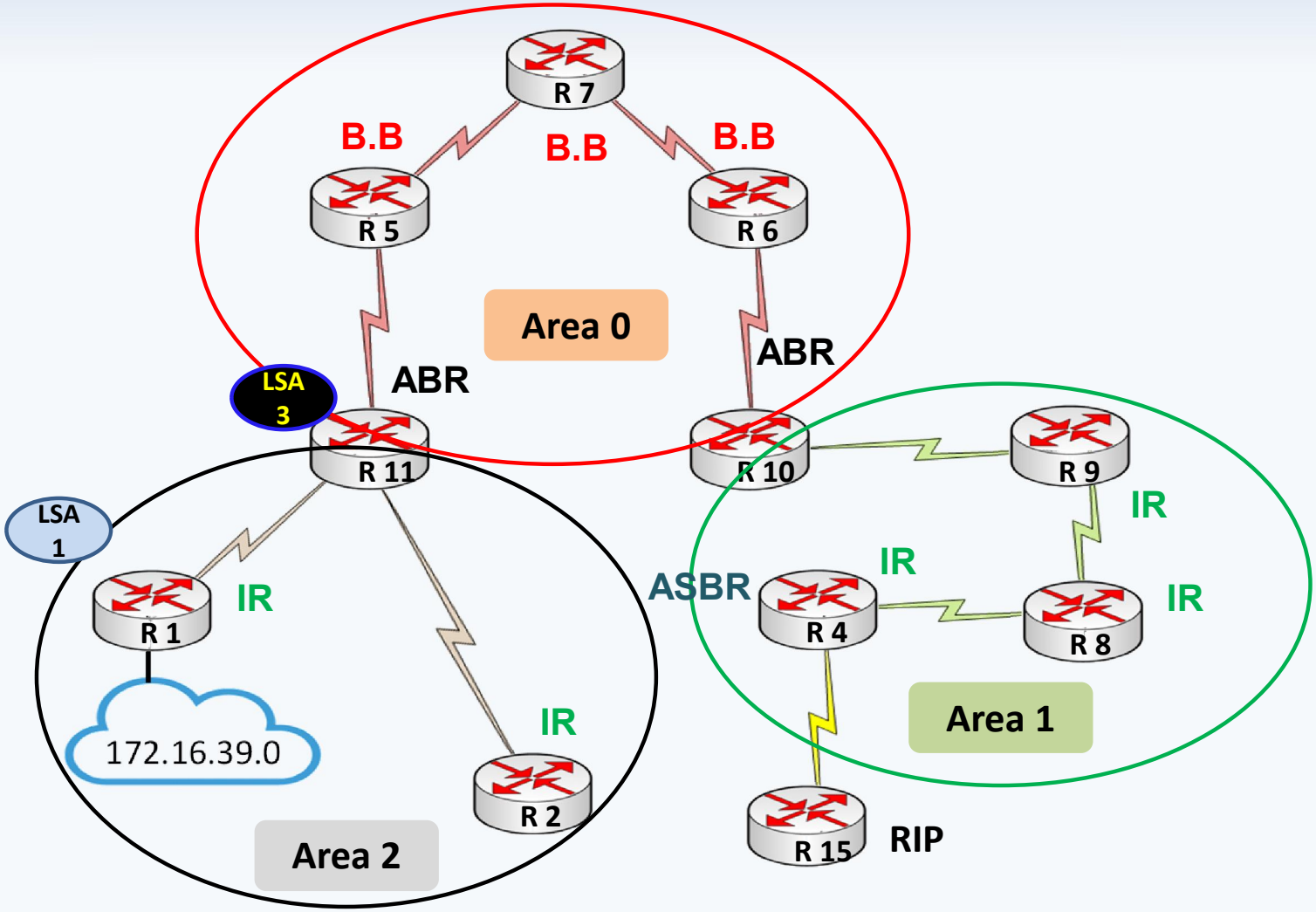
LSA type 2: (Network-LSA)

- ✓ Generated by the DR
- ✓ To describe OSPF routers adjacent to the DR
- ✓ Broadcast within the same area



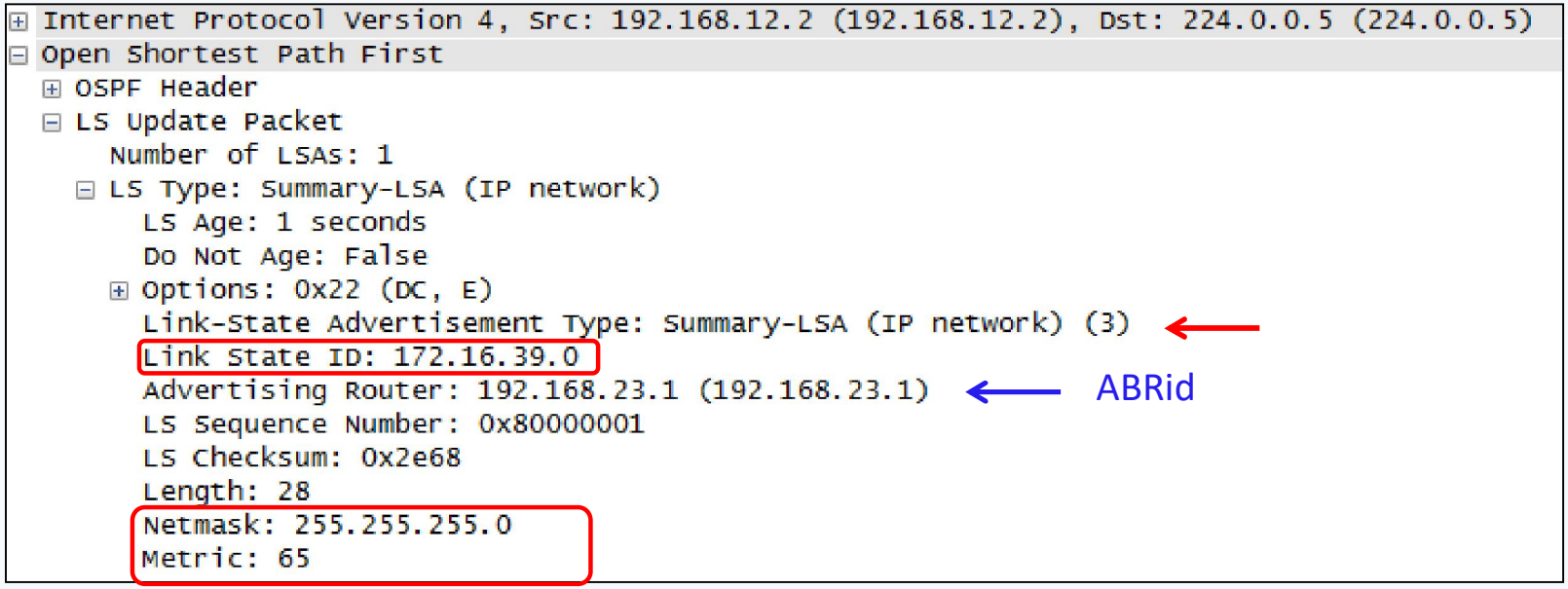

```
⊞ Internet Protocol Version 4, Src: 192.168.1.1 (192.168.1.1), Dst: 224.0.0.5 (224.0.0.5)
⊟ Open Shortest Path First
  ⊞ OSPF Header
  ⊟ LS Update Packet
      Number of LSAs: 1
    ⊟ LS Type: Network-LSA
        LS Age: 1 seconds
        Do Not Age: False
      ⊞ Options: 0x22 (DC, E)
        Link-State Advertisement Type: Network-LSA (2)
        Link State ID: 192.168.1.1
        Advertising Router: 5.5.5.1 (5.5.5.1)
        LS Sequence Number: 0x8000002e
        LS Checksum: 0x01a2
        Length: 36
        Netmask: 255.255.255.0
        Attached Router: 5.5.5.1
        Attached Router: 192.168.1.2
        Attached Router: 192.168.1.3
```

LSA type 3: (Summary-LSA)

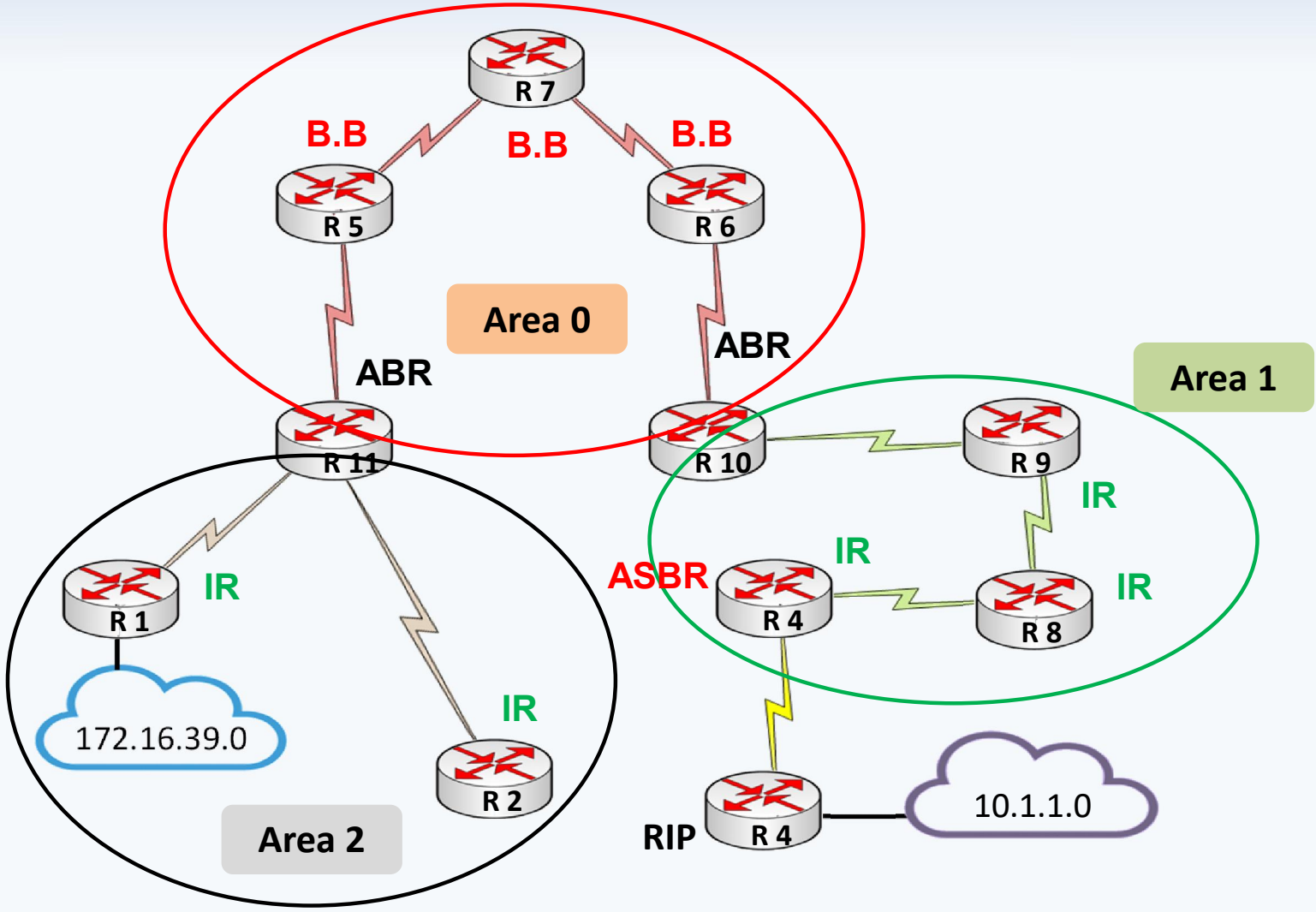


LSA type 3: (Summary-LSA)

- ✓ Advertises an IP network from another zone (summary).
- ✓ Broadcast in all areas
- ✓ Generated by ABR

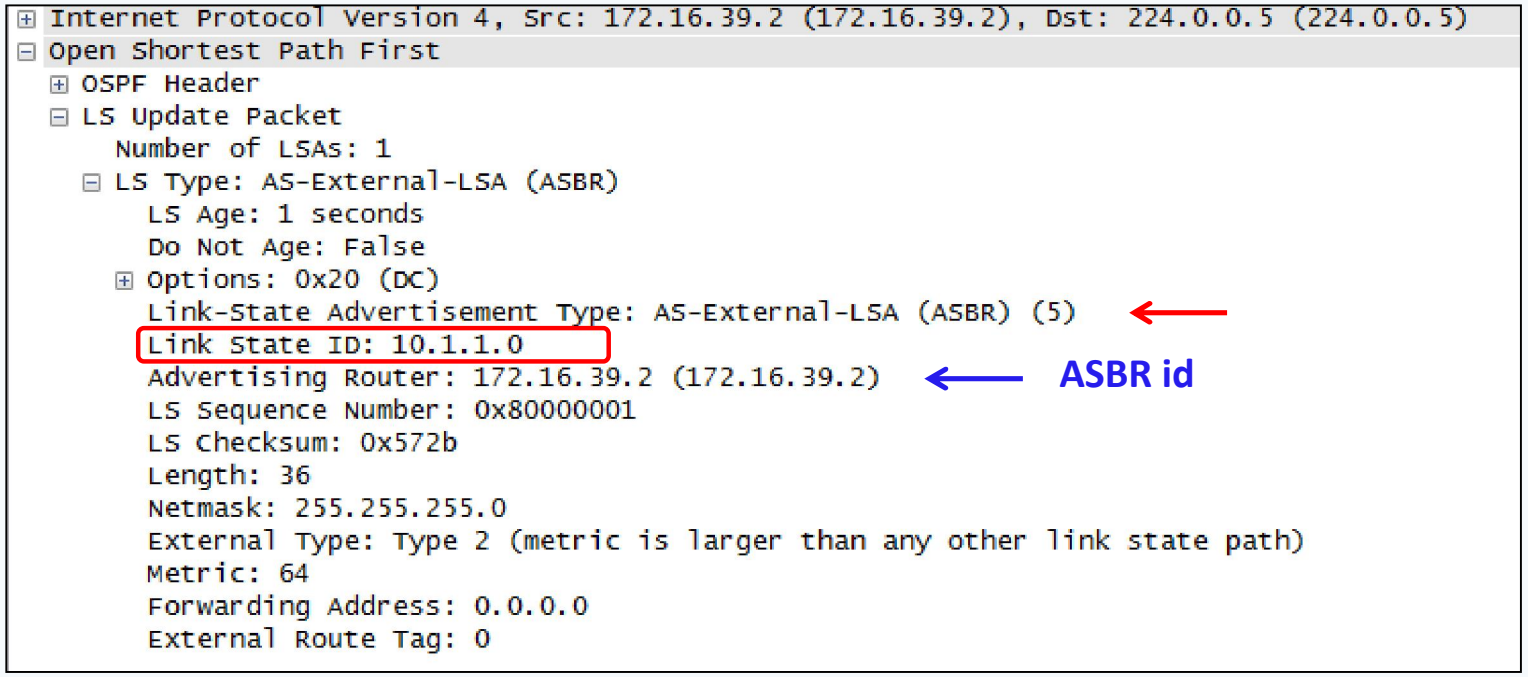


LSA Type 5 – (AS External LSA)



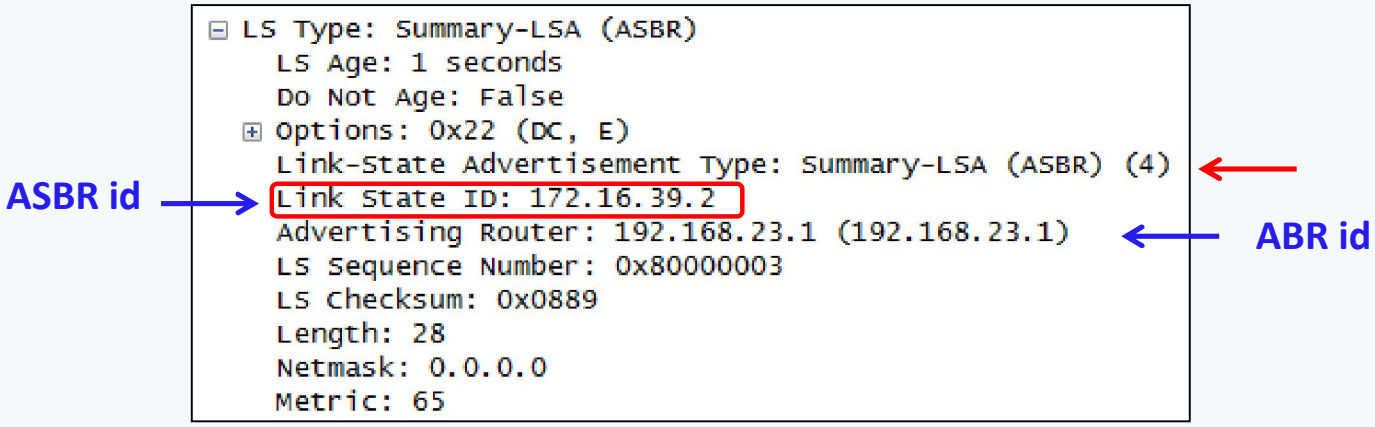
LSA Type 5 – (AS External LSA)

- ✓ To advertise an IP network from another routing protocol (AS)
- ✓ Generated by an ASBR
- ✓ Broadcast (unchanged) to all zones

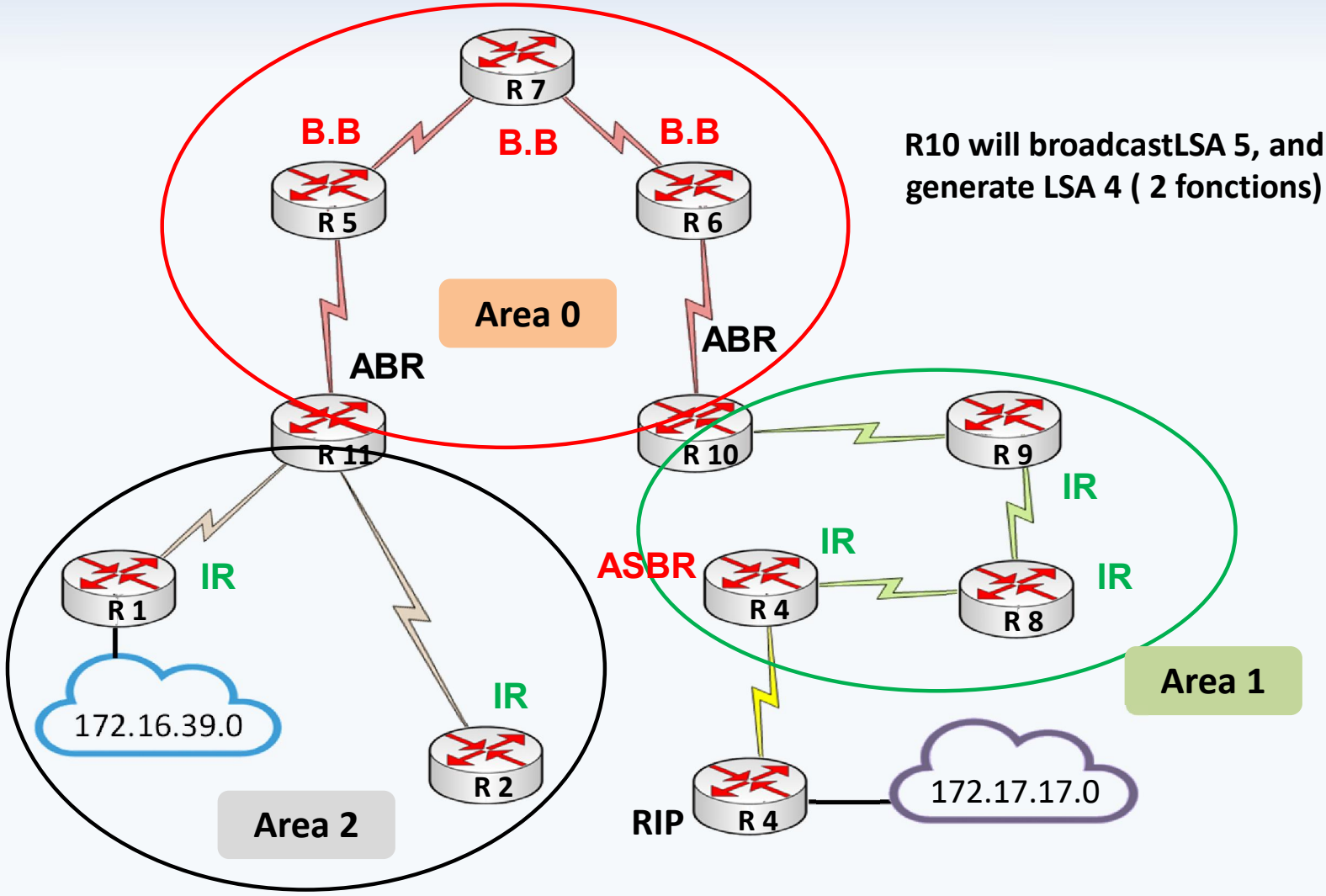


LSA type 4: ASBR summary-LSA

- ✓ Generated by ABR
- ✓ Defines a route to an ASBR
- ✓ Broadcast in all areas



LSA type 4: ASBR summary-LSA



R10 will broadcastLSA 5, and generate LSA 4 (2 fonctions)

LSAs Summary

LSA Type	Name	Description
1	Router	Describes the link states of a router (type, IP, metric, etc.)
2	Network	Lists OSPF routers adjacent to the DR
3	Summary	Advertises an IP network from another zone
4	ASBR summary	Defines a route to an ASBR
5	AS external	Advertises an IP network from another routing protocol (AS)

OSPF Algorithm

- Each OSPF router maintains a link state database containing LSAs received from all other routers.
- When a router has received all the LSAs and built its local database, OSPF uses the shortest path algorithm to create an SPF tree.
- The algorithm was designed by E. W. Dijkstra (1930-2002), a Dutch mathematician.
- The best path is the one that has the least "expensive" cost, i.e. the lowest.
- This algorithm accumulates costs along each route, from source to destination....

An OSPF database

```
R1#sho ip ospf database

            OSPF Router with ID (172.16.11.1) (Process ID 1)

                Router Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum Link count
172.16.11.1     172.16.11.1     12          0x800000B2 0x006EA7 4
192.168.23.1    192.168.23.1    77          0x800000AC 0x008CB2 2
192.168.50.1    192.168.50.1    12          0x8000000B 0x00F557 2

                Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
192.168.15.2    192.168.50.1    12          0x80000001 0x00E4D8

                Summary Net Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
172.16.39.0     192.168.23.1    77          0x80000006 0x00246D
192.168.23.0    192.168.23.1    77          0x800000AA 0x0055FB

                Summary ASB Link States (Area 0)

Link ID         ADV Router      Age         Seq#       Checksum
172.16.39.2     192.168.23.1    77          0x80000006 0x00028C

                Type-5 AS External Link States

Link ID         ADV Router      Age         Seq#       Checksum Tag
10.1.1.0        172.16.39.2     376         0x80000004 0x00512E 0
```

OSPF Algorithm

- The SPF tree is then used to populate the IP route table, with the best paths to each network.
- Each link has a cost,
- We can force the cost of an interface
- Each node has a name and has a complete database of all the links and thus has a complete knowledge of the physical topology...

OSPF metric

- OSPF uses cost as a metric to determine the best path to a destination.
- The default value for the cost depends on the bandwidth value of a link. In general, the lower the bandwidth, the higher the cost.
- The formula for calculating the cost of a link is:

$$Co\hat{u}t = \frac{10p8}{\text{Bande passante}(Bit/s)}$$

- For a 100 Mbps link: The cost = 1

$$Co\hat{u}t = \frac{100}{\text{Bande passante}(M.Bit/s)}$$

OSPF metric

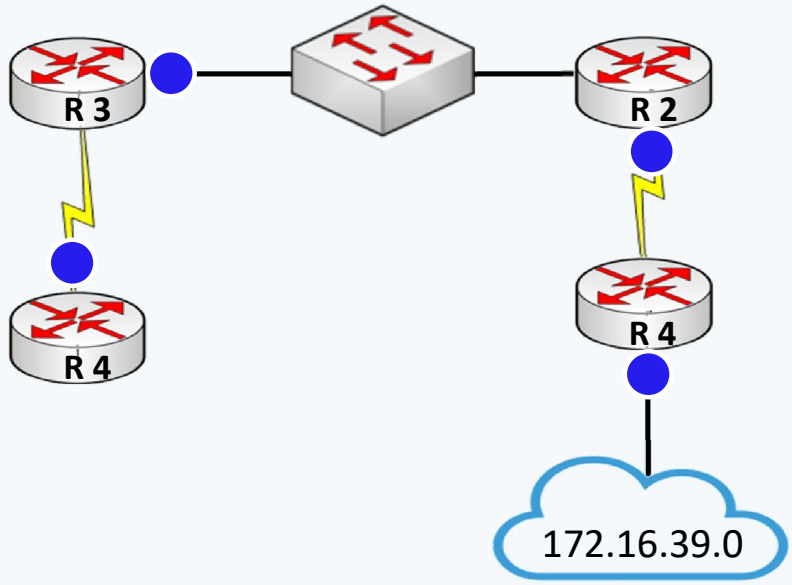
Without further configuration, OSPF considers any link equal to or greater than 100 Mbps to be represented with the best cost of 1.

Support	Cost
64 Kb/s	1562
E1 (ligne série 2048kbps)	48
Ethernet (10 Mb/s)	10
Fast Ethernet 100Mbps	1
1 Gb/s	1

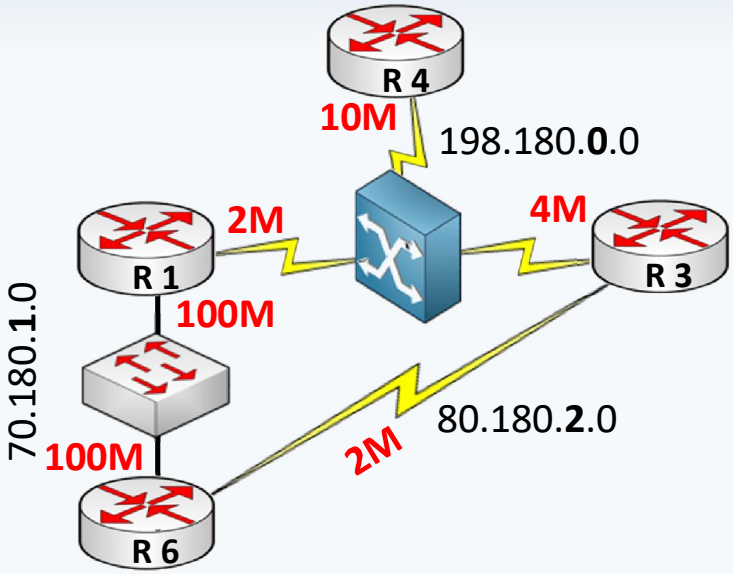
OSPF metric

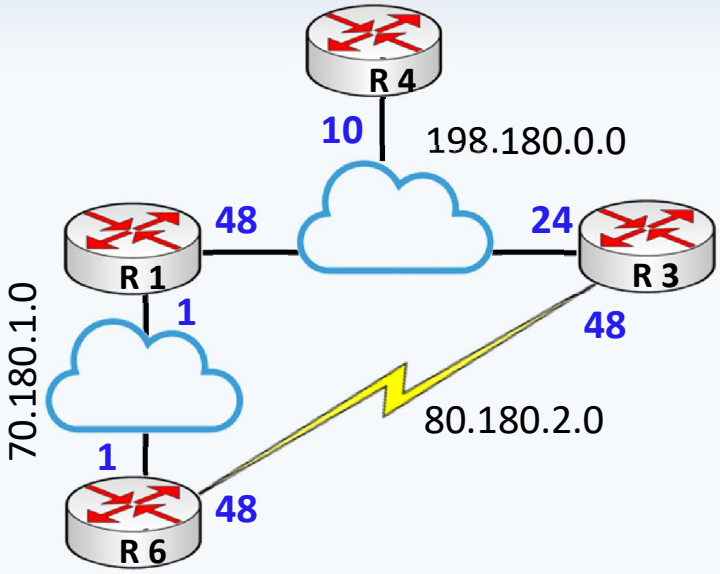
- We can change the calculation reference (on all routers), 100 (Mbps) by default. For example, to adapt to 10G:

  (config-router)#auto-cost reference-bandwidth 1000
- The cost of a path is: the sum of the cost of the interfaces



Example

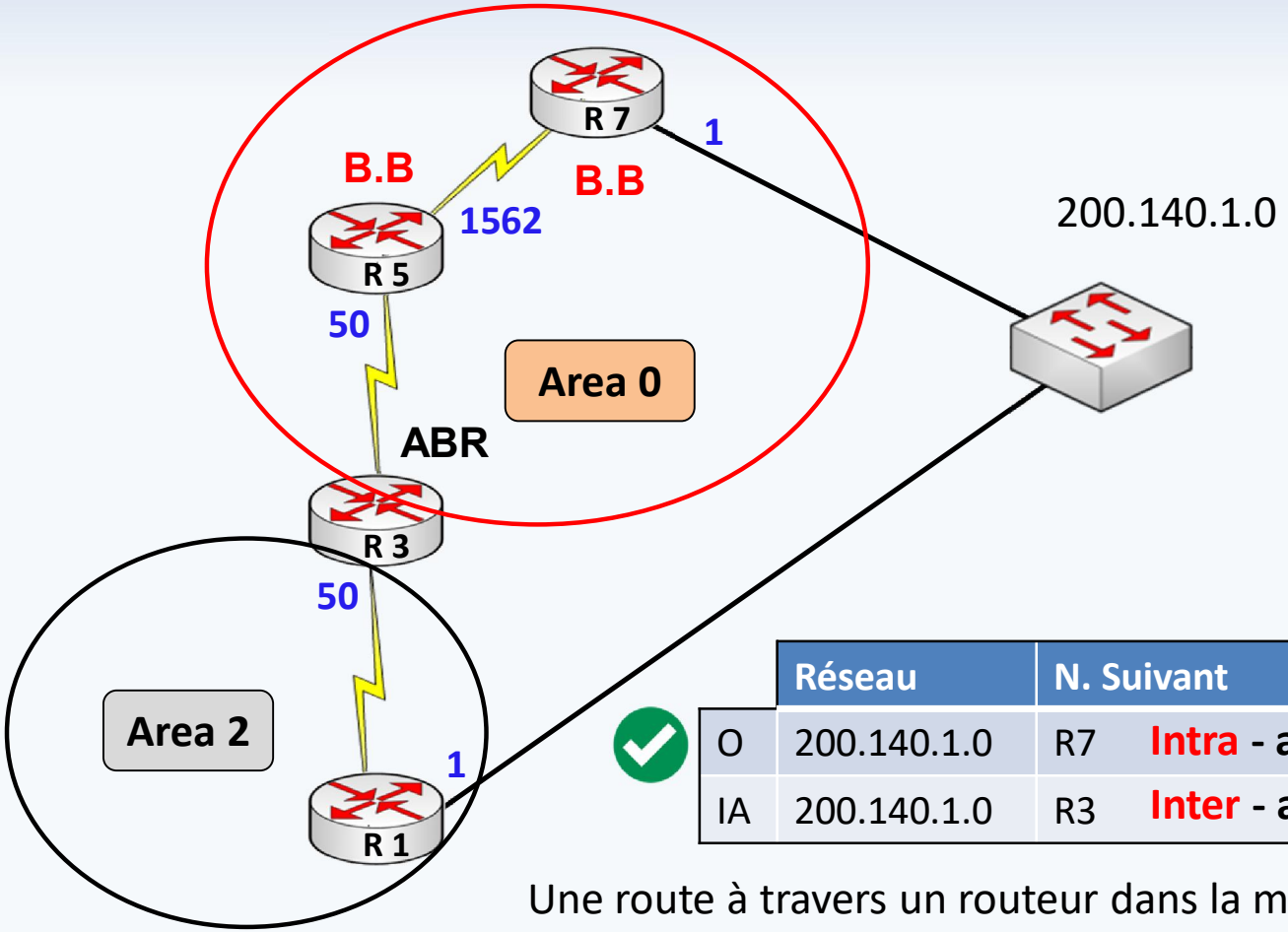




Réseau	N. Suivant	cout
198.180.0.0	.	0
70.180.1.0	.	0
80.180.2.0	R6	49

Réseau	N. Suivant	cout
198.180.0.0	.	0
70.180.1.0	R1	11
80.180.2.0	R3	58

Cas spécial

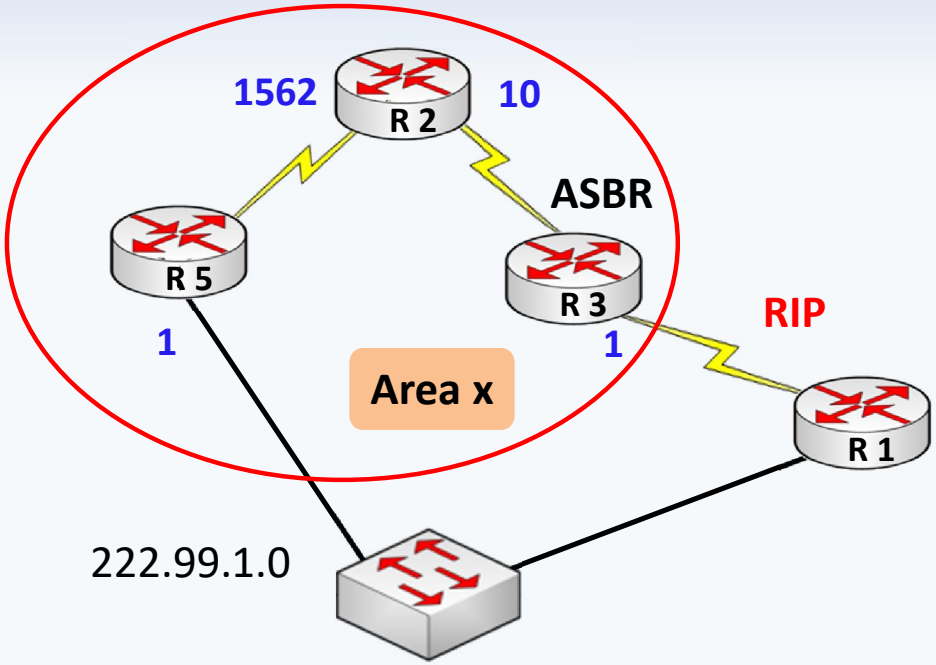


R5

	Réseau	N. Suivant		cout
O	200.140.1.0	R7	Intra - area	1563
IA	200.140.1.0	R3	Inter - area	101

Une route à travers un routeur dans la même zone est préférée à une route à travers un ABR

Special case



R2

	Network	Next hop	cost
O	222.99.1.0	R5 **Intra - area**	1563
E	222.99.1.0	R3 **external**	11

A route through a router in the same area is preferred to a route through an ASBR

References

1. *A. Tanenbaum, "Computer Network".*
2. *Keshav, "An Engineering Approach to Computer Networking".*
3. *L. Toutain, "Réseaux Locaux et Internet".*
4. https://datascientest.com/
5. https://networkcorp.fr/
6. https://www.networklab.fr/
7. https://cisco.goffinet.org/
8. http://anthonyreault.free.fr/
9. http://gponsolution.com/
10. https://nsrc.org/
11. https://loopedback.com/
12. https://www.pynetlabs.com/