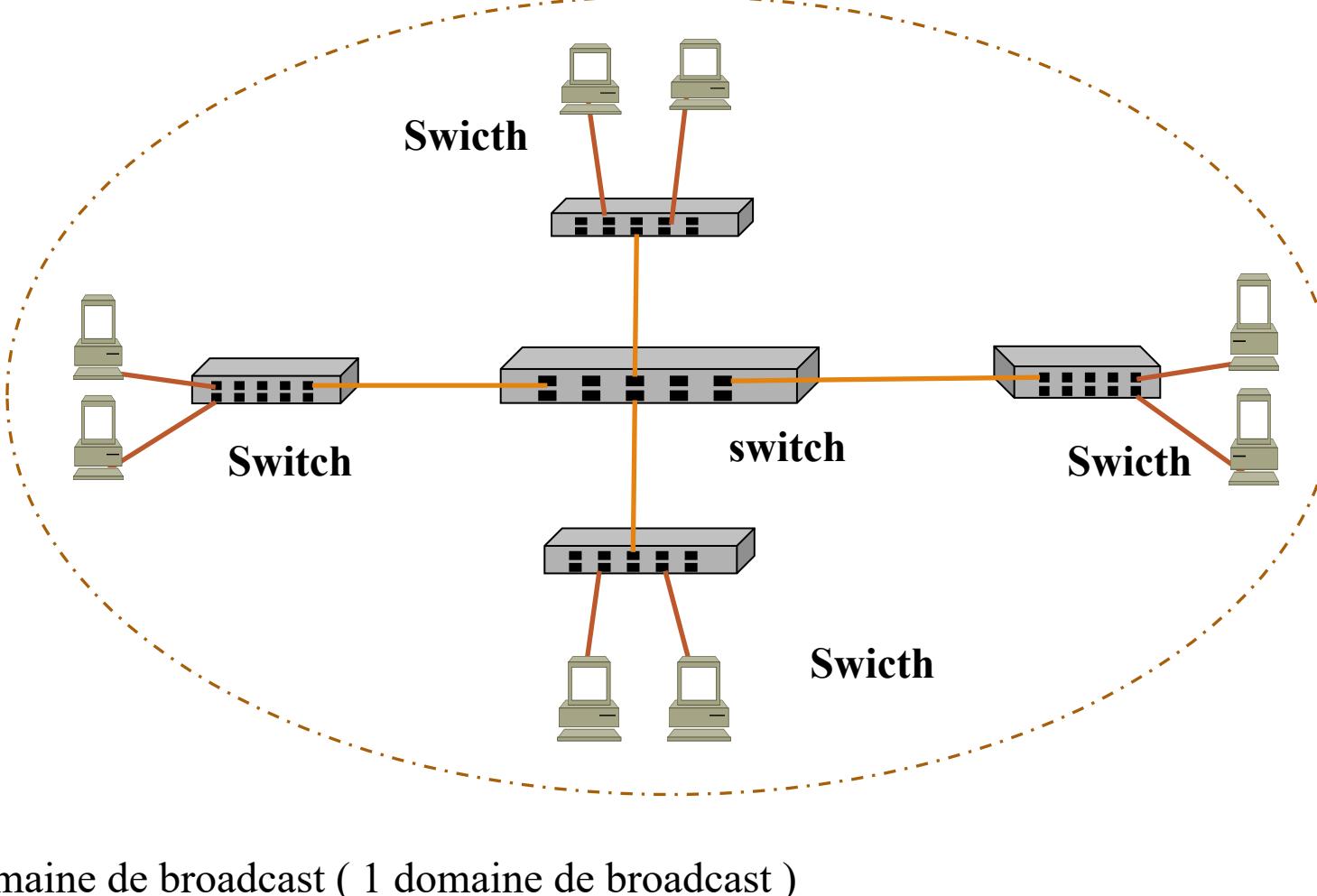


Chapitre3 : Routage - MPLS

Introduction

- Un réseau est l'interconnexion d'un ensemble de machines via des équipements d'interconnexion (commutateurs, hubs).
- Un réseau qui comporte uniquement des commutateurs est appelé réseau commuté.
- Un réseau commuté autorise la propagation des trames de broadcast.
- Un réseau commuté supporte uniquement l'adressage du niveau 2 (adresse MAC).
- Les commutateurs utilisent des tables de correspondance Port/@MAC pour localiser les machines.

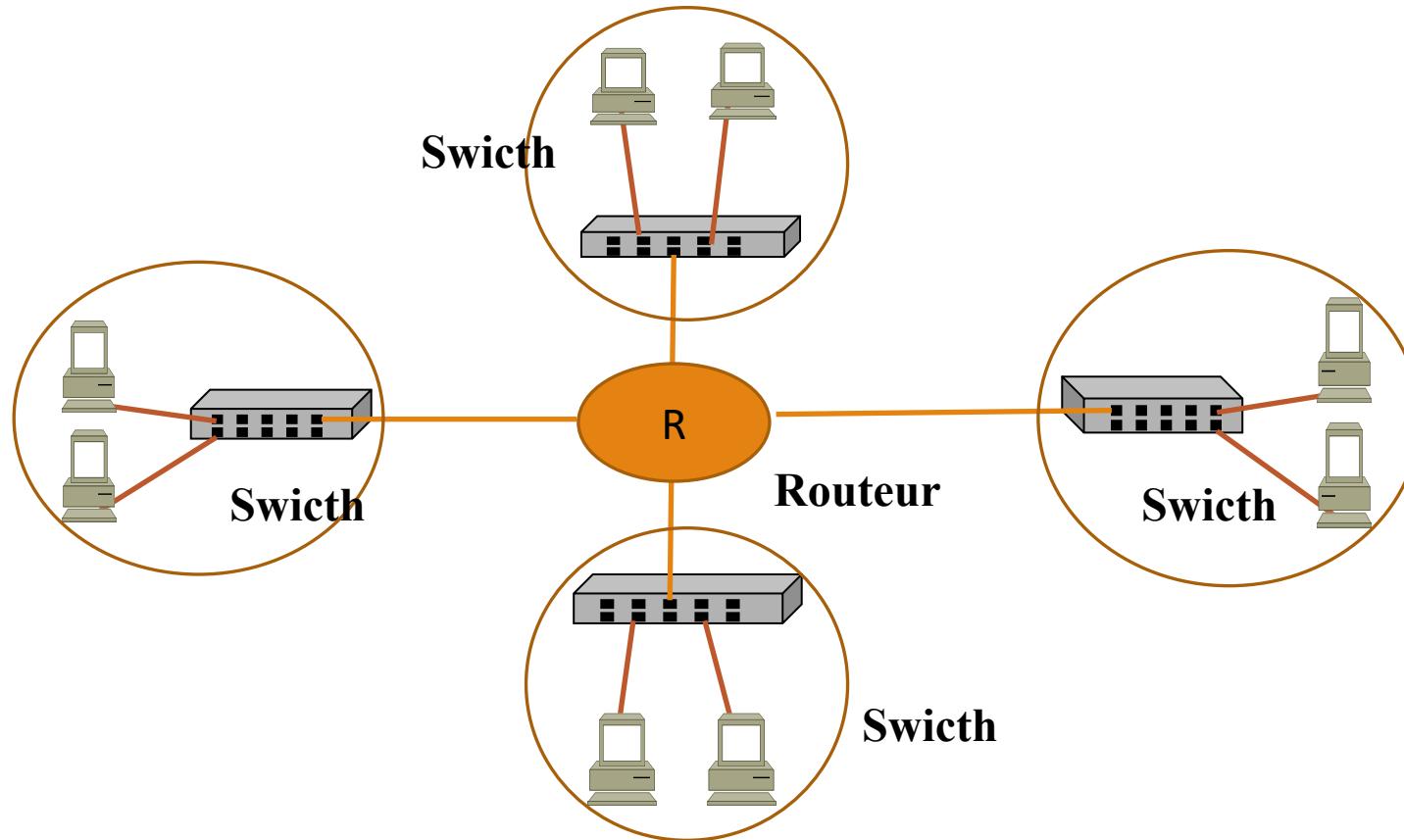
Exemple d'un réseau commuté domaine de broadcast avec des switch



Réseau commuté

- Si le nombre de machines augmente → les performances du réseau diminuent.
- Donc il faut utiliser un équipement qui :
 - supporte l'adressage logique (adresses IP),
 - permet de limiter les domaines du broadcast
 - et qui permet localiser les machines en fonction de leurs adresses IP .
- L'équipement qui offre ces fonctionnalités s'appelle un routeur (c'est un équipement de la couche 3).

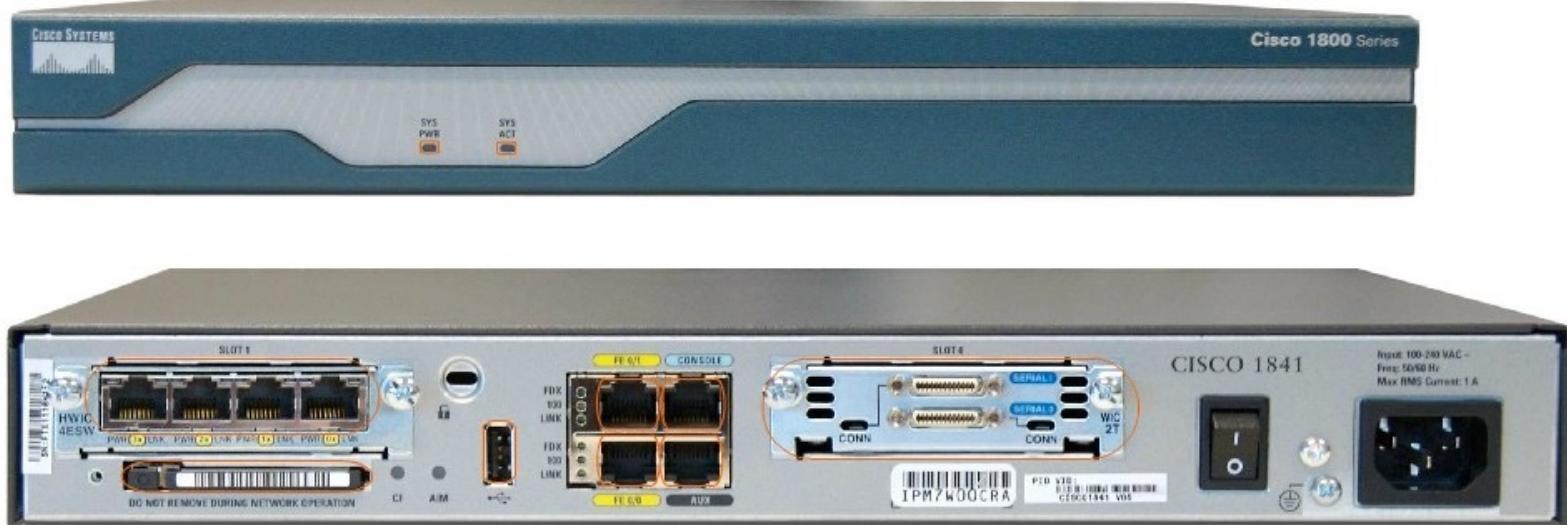
Utilisation d'un routeur



— Domaine de broadcast (4 domaines de broadcast)

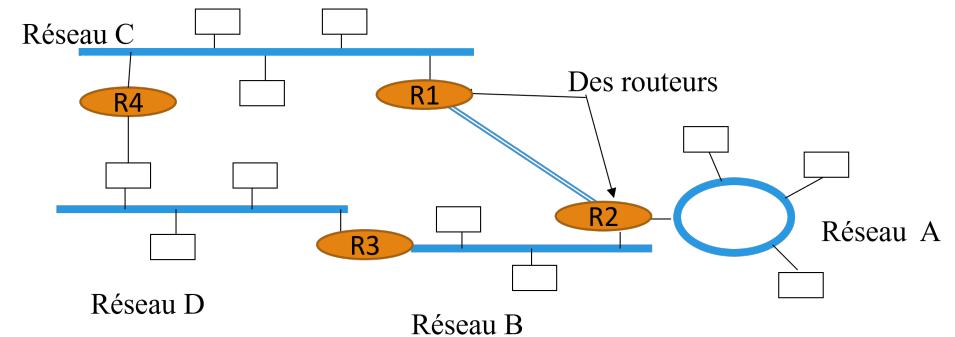
Les routeurs

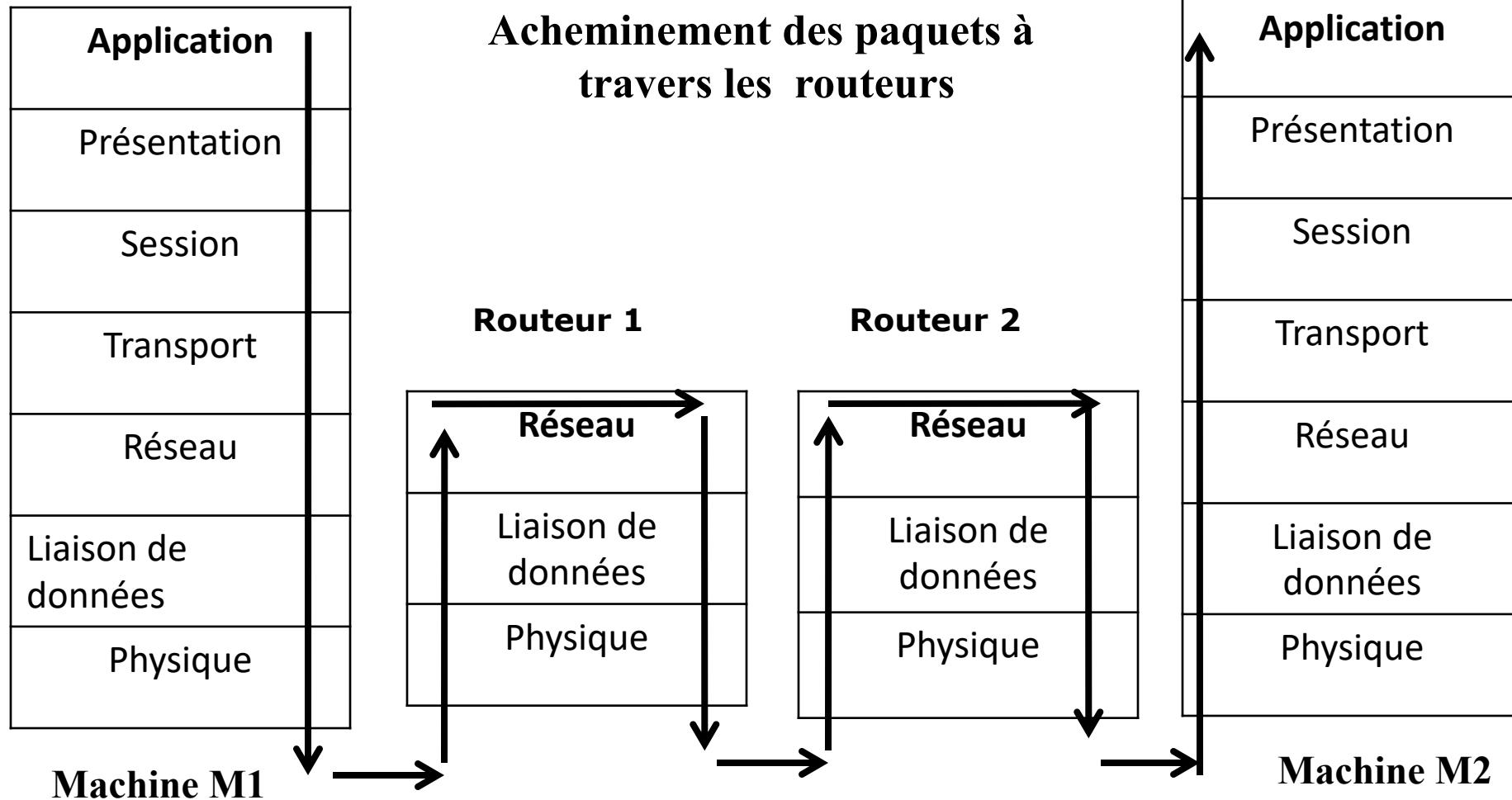
- Un routeur est un équipement qui possède plusieurs interfaces réseaux.
- Chaque interface est connectée à un réseau.
- Il possède la capacité de router (acheminer) les paquets d'un réseau à un autre selon les adresses IP.
- Un réseau qui comporte des routeurs (ou des passerelles) est appelé réseau routé.
- Un réseau routé permet de limiter le domaine du broadcast.



Les routeurs

- Un routeur peut être :
 - Une machine spécialisée : une machine conçue spécialement pour le routage. Généralement un routeur comporte un système d'exploitation spécialisé (ex : IOS de CISCO)
 - Une machine ordinaire (ordinateur) équipée d'un système d'exploitation (ex : Unix/Linux , Windows) et qui possède plusieurs interfaces. Dans ce cas on parle de passerelle (gateway)





- Avant qu'un paquet arrive sa destination il passe par un ou plusieurs routeurs.
 - Lorsqu'un routeur reçoit le paquet (couche 3), il **va consulter l'entête du paquet** pour extraire des informations qui l'aide à acheminer le paquet.
 - **L'information la plus importante est l'adresse IP de la destination → existe dans l'entête d'un paquet IP.**

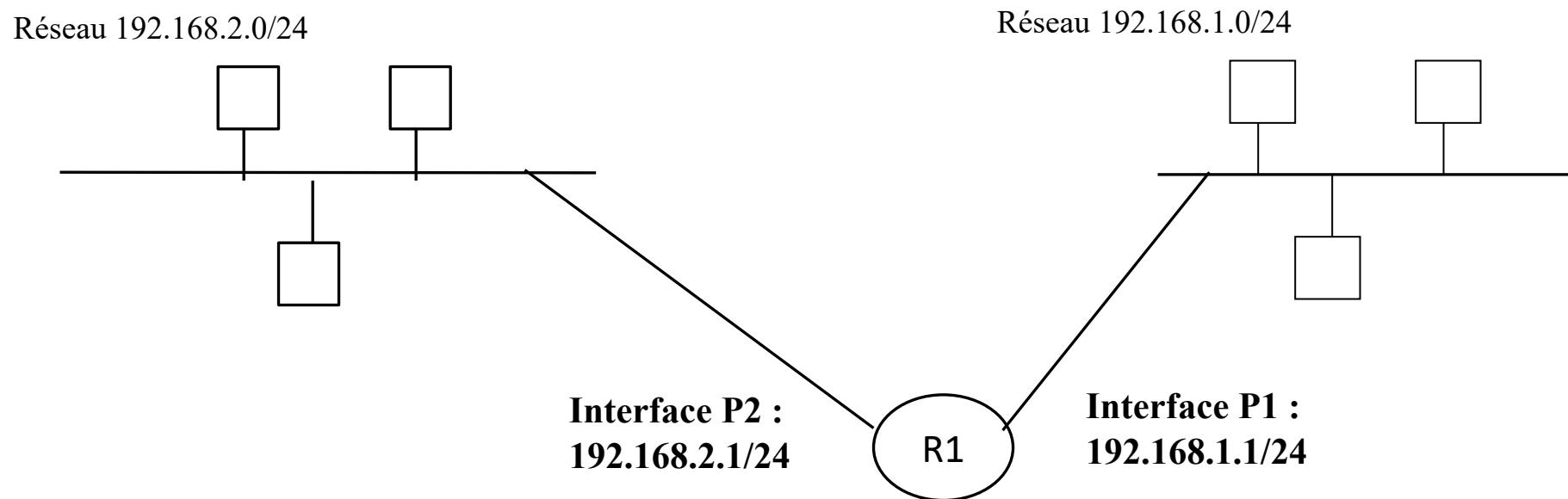
Structure de l'entête datagramme IP

L'**unité de transfert de base dans un réseau internet** est le datagramme qui est constituée d'un en-tête et d'un champ de données:

0	4	8	16	19	24	31					
VERS	HLEN	Type de service	Longueur totale								
Identification		Flags	Offset fragment								
Durée de vie(TTL)	Protocole	Somme de contrôle Header									
Adresse IP Source											
Adresse IP Destination											
Options IP (éventuellement)				Padding							
Données											

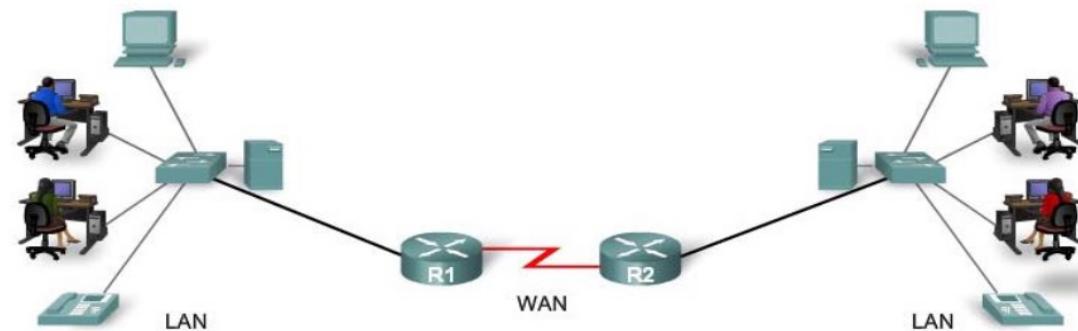
Adresses d'un routeur

- A chaque interface physique (carte réseau) correspond une adresse IP.
- Les routeurs possèdent (par définition) plusieurs interfaces. Une pour chaque connexion.
- Donc à une même machine (un routeur), est associé plusieurs adresses IP.



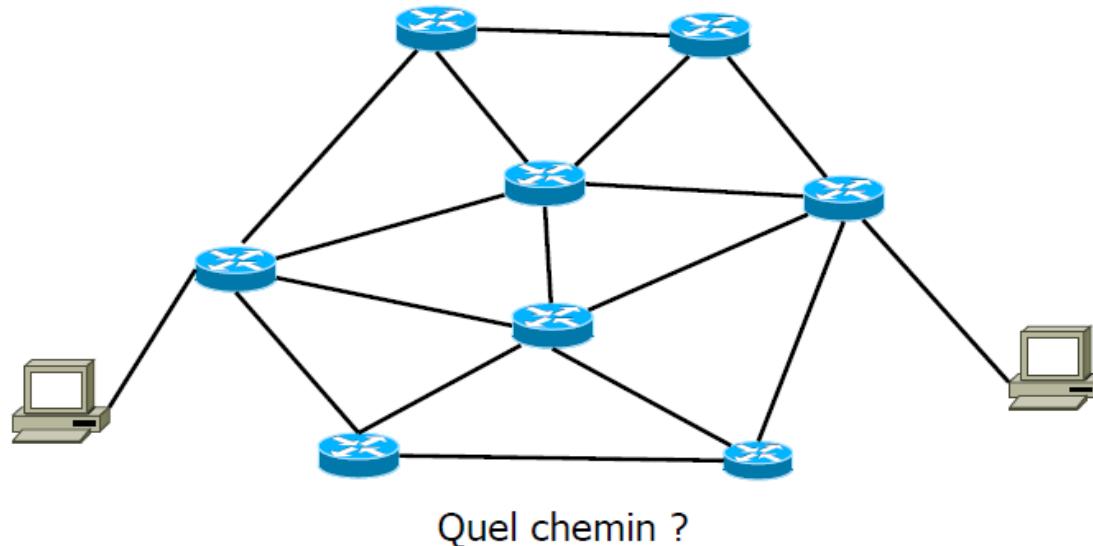
Le routage

- Le routage (routage des paquets) est le processus permettant à un paquet d'être acheminé vers le destinataire lorsque celui-ci n'est pas sur le même réseau physique que l'émetteur.
- La sélection du chemin est le processus que le routeur utilise pour choisir le prochain saut (prochain routeur) du trajet que le paquet empruntera vers sa destination.



Le routage

- Les routeurs coopèrent entre eux de telle manière qu'un datagramme passe d'un routeur à un autre jusqu'à ce que l'une d'entre elles le délivre à son destinataire.
- Le routeur ne connaît pas le chemin complet pour atteindre la destination.
- Un routeur a juste une vision locale du réseau global .



La Table de routage

- Pour assurer la fonction de routage (trouver un chemin) le ROUTEUR CONSULTE une TABLE dite table DE ROUTAGE.
- Les tables de routage contiennent les informations nécessaires à la transmission des paquets sur les autres réseaux.
- Les routeurs emploient des protocoles de routage pour : construire et gérer les tables de routage contenant les informations d'acheminement.

Structure de la table de routage

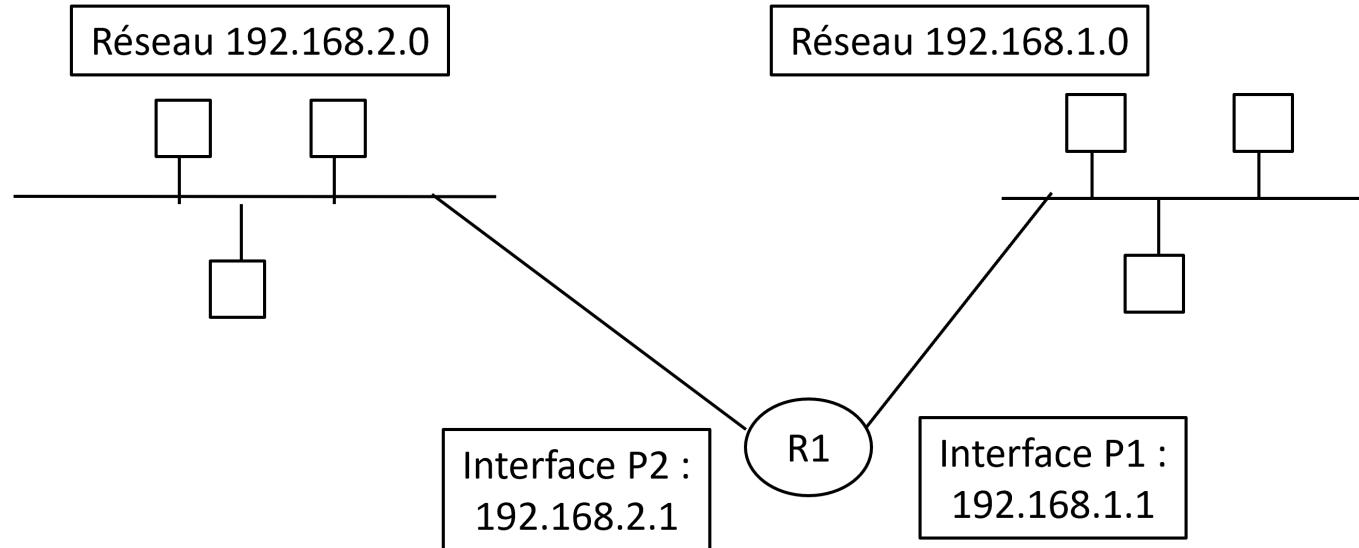
- Une table de routage contient les informations suivantes :
 - **Adresse réseau destination** ainsi que le masque associé à ce réseau.
 - **Adresse IP de la Passerelle** : correspondant à l'adresse du prochain routeur qui va recevoir le paquet destiné au réseau de destination .
 - **Interface de sortie** : puisque le routeur possède plusieurs interfaces , donc il faut indiquer via quelle interface doit transiter le paquet.
- Généralement la table de routage IP, contient seulement les adresses réseaux et non pas les adresses machines afin de ne pas avoir une table de très grande taille.

Structure de la table de routage

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
192.168.1.0/255.255.255.0	Direct (le même réseau)	P1
192.168.2.0/255.255.255.0	192.168.3.1	P2

- Ligne 1 : indique que pour atteindre le réseau 192.168.1.0 passer directement via l'interface P1 . Ce réseau est connecté localement ,**pas besoin de passer par un autre routeur**.
- Ligne 2 : indique que pour atteindre le réseau 192.168.2.0 passer par le routeur d'adresse 192.168.3.1 et via l'interface P2.

Exemple 1 : un routeur et deux réseaux



Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
192.168.1.0/255.255.255.0	Direct	P1
192.168.2.0/255.255.255.0	Direct	P2

Exemple 2 : deux routeurs et deux réseaux

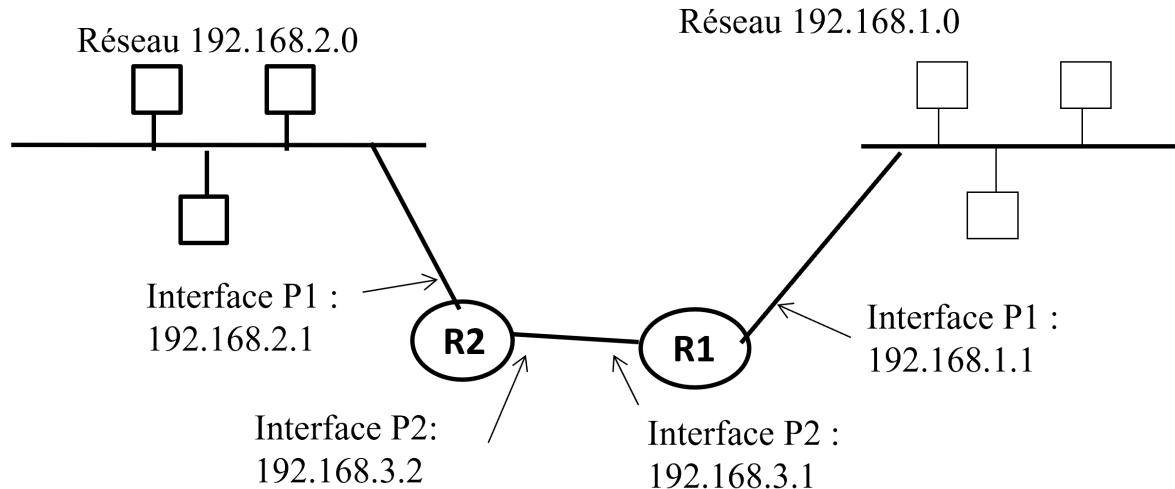


Table de routage de routeur R1

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
192.168.1.0/24	Direct	P1
192.168.3.0/24	Direct	P2
192.168.2.0/24	192.168.3.2	P2

Table de routage de routeur R2

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
192.168.2.0/24	Direct	P1
192.168.3.0/24	Direct	P2
192.168.1.0/24	192.168.3.1	P2

Route par défaut passerelle par défaut

- La route par défaut est utilisée lorsque le chemin de destination n'est pas explicitement précisé dans la table de routage (une destination inconnue) : s'il n'y a pas une route vers le réseau de destination alors par défaut l'envoyer vers cette passerelle
- Grouper la même route pour un ensemble de réseaux pour ne pas encombrer la table de routage.

Exemple: Route par défaut (1/2)

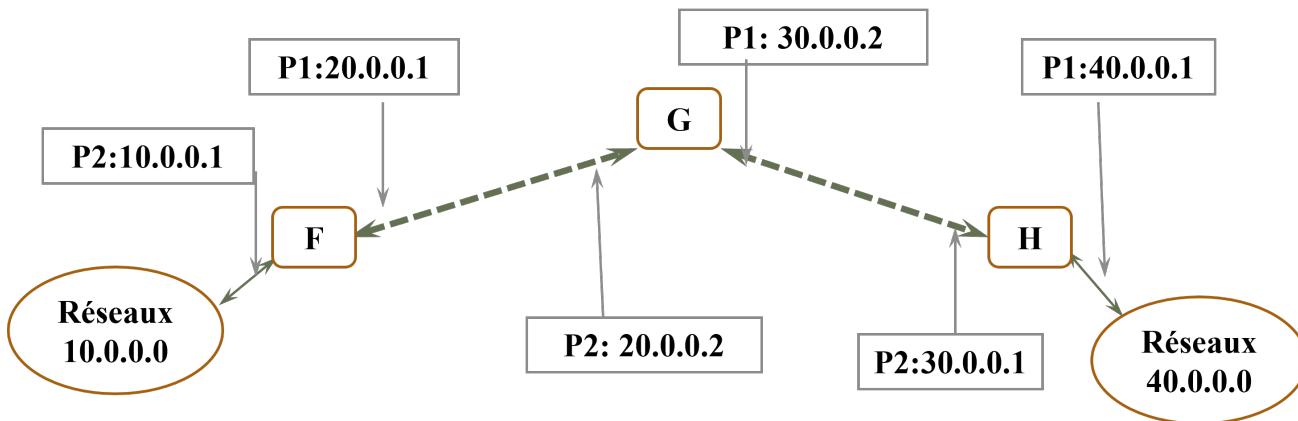


Table de routage de routeur G

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
20.0.0.0/8	Direct	P2
30.0.0.0/8	Direct	P1
40.0.0.0/8	30.0.0.1	P1
10.0.0.0/8	20.0.0.1	P2

Table de routage de routeur F

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
10.0.0.0/8	Direct	P2
20.0.0.0/8	Direct	P1
30.0.0.0/8	20.0.0.2	P1
40.0.0.0/8	20.0.0.2	P1

Table de routage de routeur H

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
40.0.0.0/8	Direct	P1
30.0.0.0/8	Direct	P2
20.0.0.0/8	30.0.0.2	P2
10.0.0.0/8	30.0.0.2	P2

Exemple: Route par défaut (2/2)

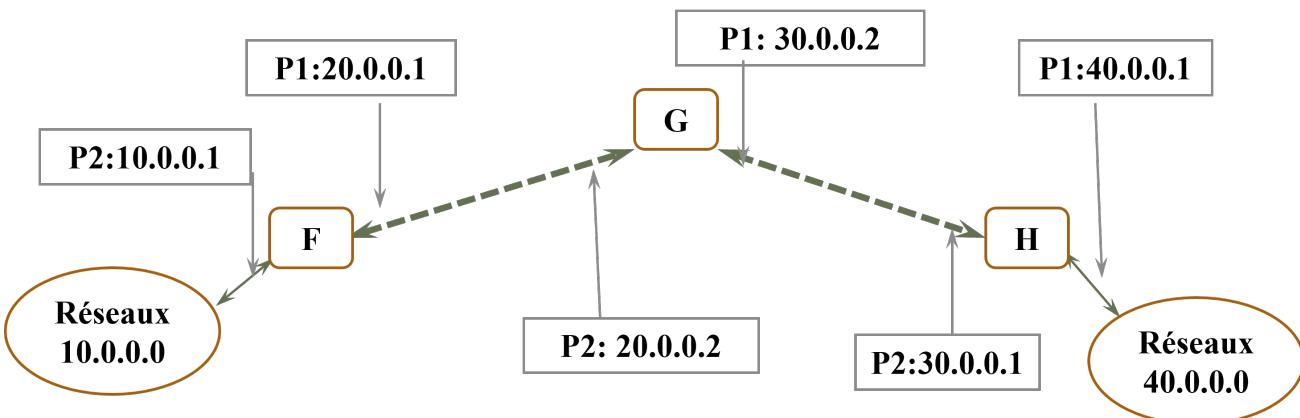


Table de routage de routeur F

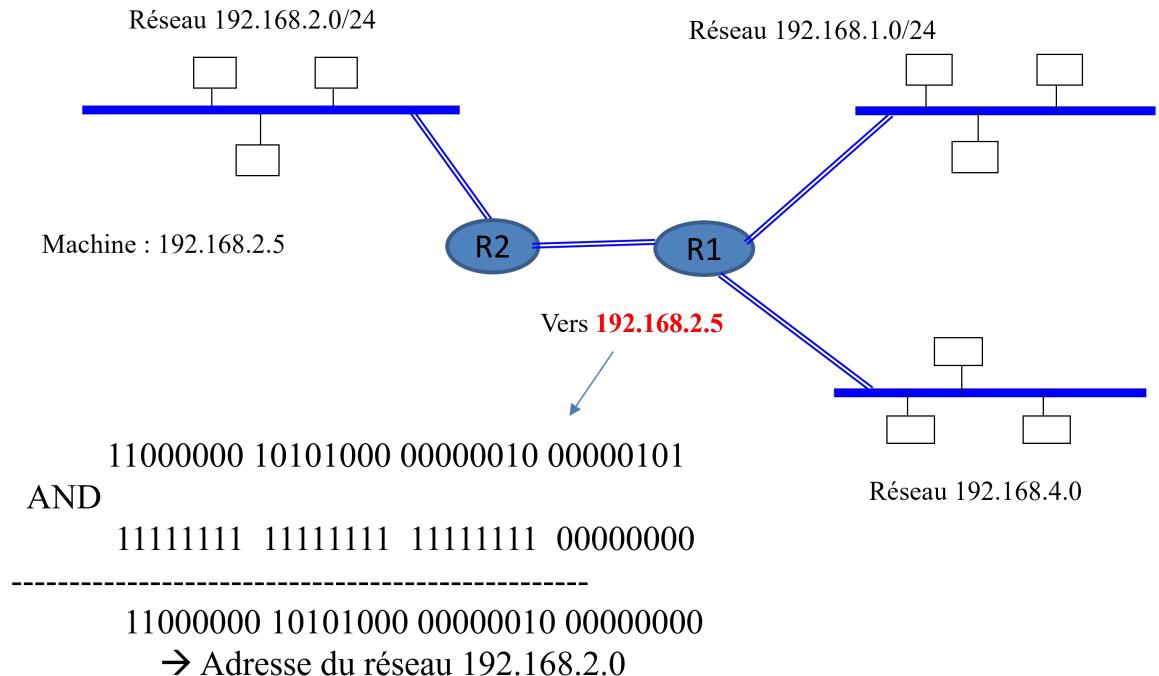
Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
10.0.0.0/8	Direct (le même réseaux)	P2
20.0.0.0/8	Direct (le même réseaux)	P1
0.0.0.0/0.0.0 (tous les autres réseaux) Route par défaut	20.0.0.2	P1

Table de routage de routeur H

Pour atteindre le réseau de destination	Prochain nœud Passerelle	Via l'interface
40.0.0.0/255.0.0.0	Direct (le même réseaux)	P1
30.0.0.0/255.0.0.0	Direct (le même réseaux)	P2
0.0.0.0/0.0.0 (tous les autres réseaux) Route par défaut	30.0.0.2	P2

Calcul de l'adresse réseau de destination

- La première étape du processus de routage est le calcul de l'adresse du réseau de destination.
- Pour cela le routeur ou la machine consulte sa table de routage et effectue un ET LOGIQUE entre l'adresse destination et les masques qui existent dans chaque entrée de la table de routage.



Types de routage

comment construire la table de routage ?

- **Routage statique** : Les informations de routage (tables de routage) sont mises à jour manuellement à chaque modification topologique du réseau.
- **Routage dynamique** : Les informations (tables de routage) relatives à la route sont mises à jour automatiquement entre les routeurs selon un protocole (RIP , OSPF, etc.)

Routage Statique vs Routage Dynamique

	Routage dynamique	Routage statique
Complexité de la configuration	Généralement indépendant de la taille du réseau	Augmente avec la taille du réseau
Connaissances d'administrateur requises	Connaissances avancées requises	Aucune connaissance supplémentaire n'est requise
Modifications apportées à la topologie	S'adapte automatiquement aux modifications apportées à la topologie	Intervention de l'administrateur requise
Évolutivité	Idéal pour les topologies simples et complexes	Idéal pour les topologies simples
Sécurité	Moins sécurisé	Plus sécurisé
Utilisation de la ressource	Utilise l'UC, la mémoire, la bande passante de la liaison	Aucune ressource supplémentaire n'est requise
Prévisibilité	La route dépend de la topologie actuelle	La route menant à la destination est toujours la même

Routage Statique

Routage statique

- La table de routage est mise à jour manuellement.
- Si la topologie change :
 - Ajout d'un réseau : ajouter la ou les règles de routage pour accéder à ce réseau sur l'ensemble des routeurs concernés
 - Suppression d'un réseau : supprimer la ou les règles de routage pour accéder à ce réseau sur l'ensemble des routeurs concernés

Commande Cisco pour configurer une route statique

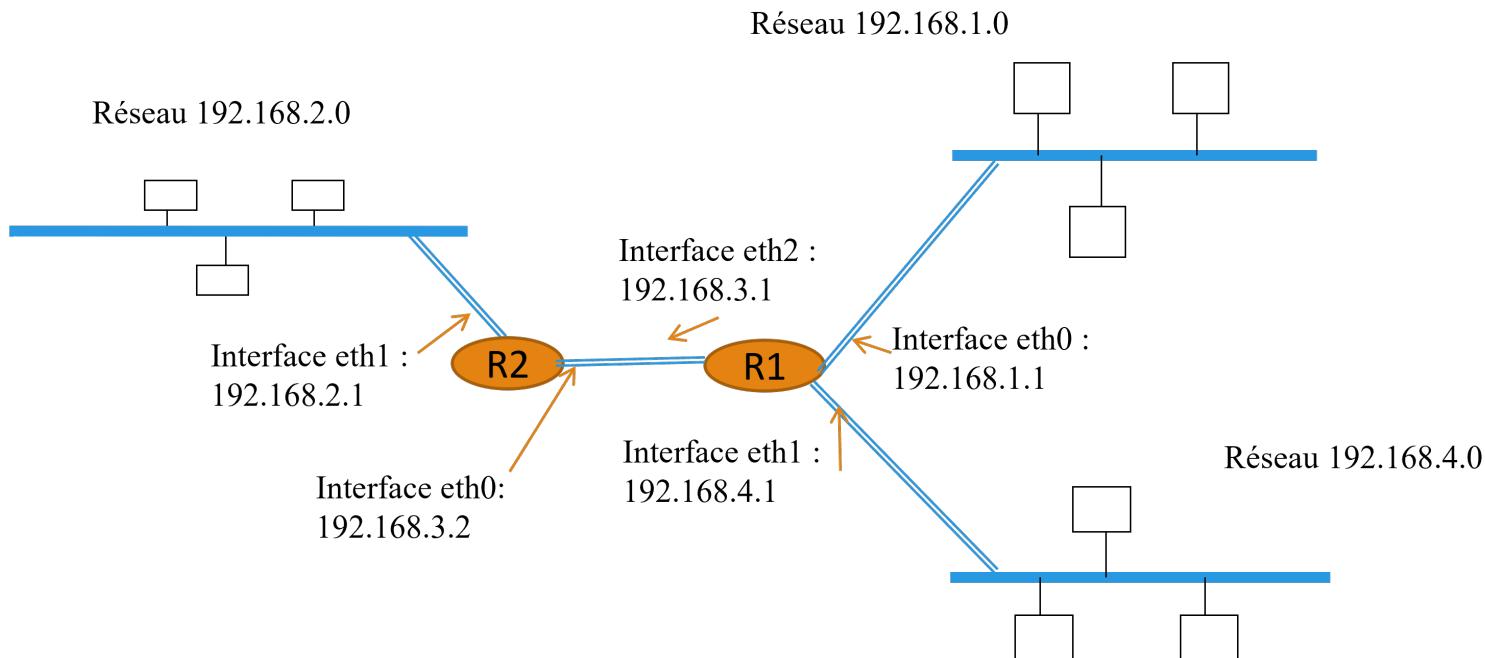
```
Router(config)# ip route network-address subnet-mask  
{ip-address | exit-interface }
```

Paramètre	Description
network-address	Adresse de destination du réseau distant, à ajouter à la table de routage.
subnet-mask	Masque de sous-réseau du réseau distant, à ajouter à la table de routage. Vous pouvez modifier le masque de sous-réseau pour résumer un groupe de réseaux.
ip-address	Généralement appelé adresse IP du routeur de tronçon suivant.
exit-interface	Interface sortante utilisée pour transférer des paquets au réseau de destination.

Commande Linux pour configurer une route statique

Sur le routeur R1 : Ajouter une route vers le réseau 192.168.2.0/24 en passant par la passerelle 192.168.3.2 via l'interface eth2:

```
route add -net 192.168.2.0 netmask 255.255.255.0 gw 192.168.3.2 dev eth2
```



Configuration d'un Routeur CISCO (1/3)

Syntaxe des commandes de configuration des paramètres de base d'un routeur

Attribution d'un nom au routeur	Router(config)#hostname name
Définition des mots de passe	Router(config)#enable secret password Router(config)#line console 0 Router(config-line)#password password Router(config-line)#login Router(config)#line vty 0 4 Router(config-line)#password password Router(config-line)#login
Configuration d'une bannière de message du jour	Router(config)#banner motd # message #

Configuration d'un Routeur CISCO (2/3)

Syntaxe des commandes de configuration des paramètres de base d'un routeur	
Configuration d'une interface	<pre>Router(config)#interface type number Router(config-if)#ip address address mask Router(config-if)#description description Router(config-if)#no shutdown</pre>
Enregistrement des modifications apportées à un routeur	<pre>Router#copy running-config startup-config</pre>
Vérification des informations renvoyées par les commandes show	<pre>Router#show running-config Router#show ip route Router#show ip interface brief Router#show interfaces</pre>

Configuration d'un Routeur CISCO (3/3)

Vérification de la configuration de routeur de base

show running-config

```
R1#show running-config
!
version 12.3
!
hostname R1
!
interface FastEthernet0/0
description R1 LAN
ip address 192.168.1.1 255.255.255.0
!
interface serial0/0/0
description Link to R2
ip address 192.168.2.1 255.255.255.0
clock rate 64000
!
banner motd ^C
*****
WARNING!! Unauthorized Access Prohibited!!
*****
^C
!
line con 0
password cisco
login
line vty 0 4
password cisco
login
!
end
```

Routage Dynamique

Routage dynamique

- Types d'algorithmes de routage
 - Le routage statique (non adaptatif)
 - Décisions de routage précalculées et chargées initialement dans le routeur
 - Les routes sont fixes. Il faut les changer “à la main” en cas de modifications du réseau
 - algorithme du plus court chemin
 - routage par inondation «flooding»
 - Le routage dynamique (adaptatif)
 - décisions de routage évoluant selon les modifications de topologie et de trafic
 - Dialogue entre routeurs
 - routage à vecteur de distance «distance vector»
 - routage à état de lien «link state»

Routage dynamique

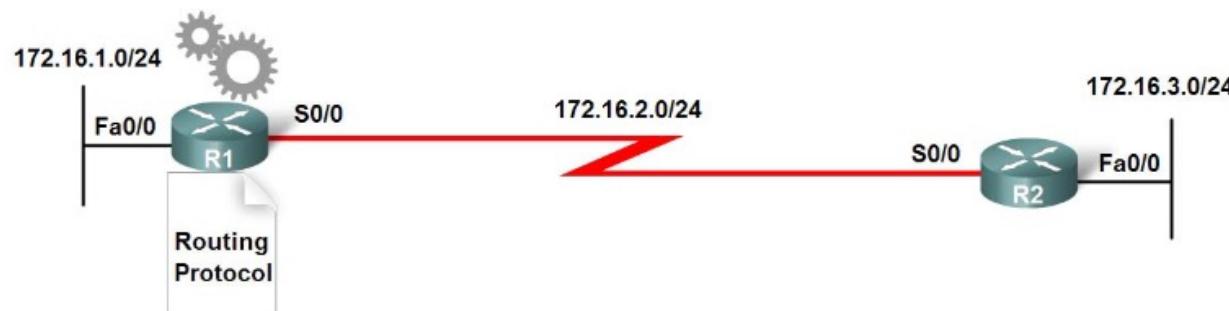
1. **Algorithme:** pour déterminer le meilleur chemin

Ex. RIP → Bellman, OSPF → Djikistra, EIGRP → Dual

2. **Structures de données/bases de données:** pour conserver les informations de routage

3. **Messages:**

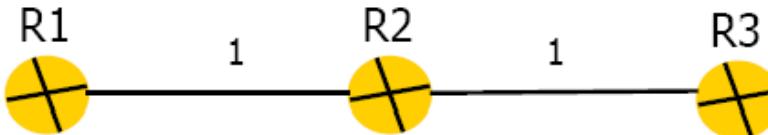
- messages pour découvrir les routeurs voisins
- informations de routage (Mise-à-jour)



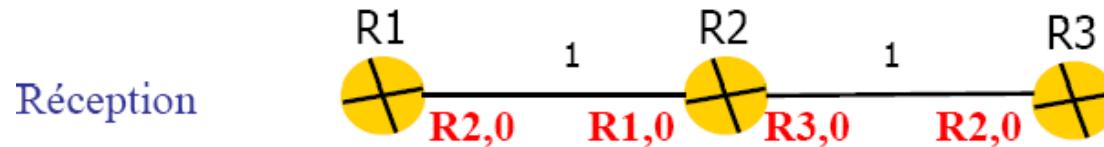
Routage dynamique

- Méthode *Distance Vector*: Méthode **itérative** par transmission des tables de routage. Chaque routeur
 - Transmet sa table de routage initiale à ses voisins immédiats.
 - Fusionne sa table avec les tables reçues de ses voisins immédiats.
 - En cas de modification, transmet la nouvelle table à ses voisins immédiatsConvergence non garantie.
- Méthode *Link State*: **non itérative** par transmission d'état des liaisons. Chaque routeur
 - Transmet l'état de ses liaisons à ses voisins immédiats.
 - Retransmet les informations reçues de ses voisins à ses autres voisins (propagation par inondation).
 - Calcule sa table de routage.Pas de problème de convergence.

Exemple de Routage à vecteur de distance (1/3)

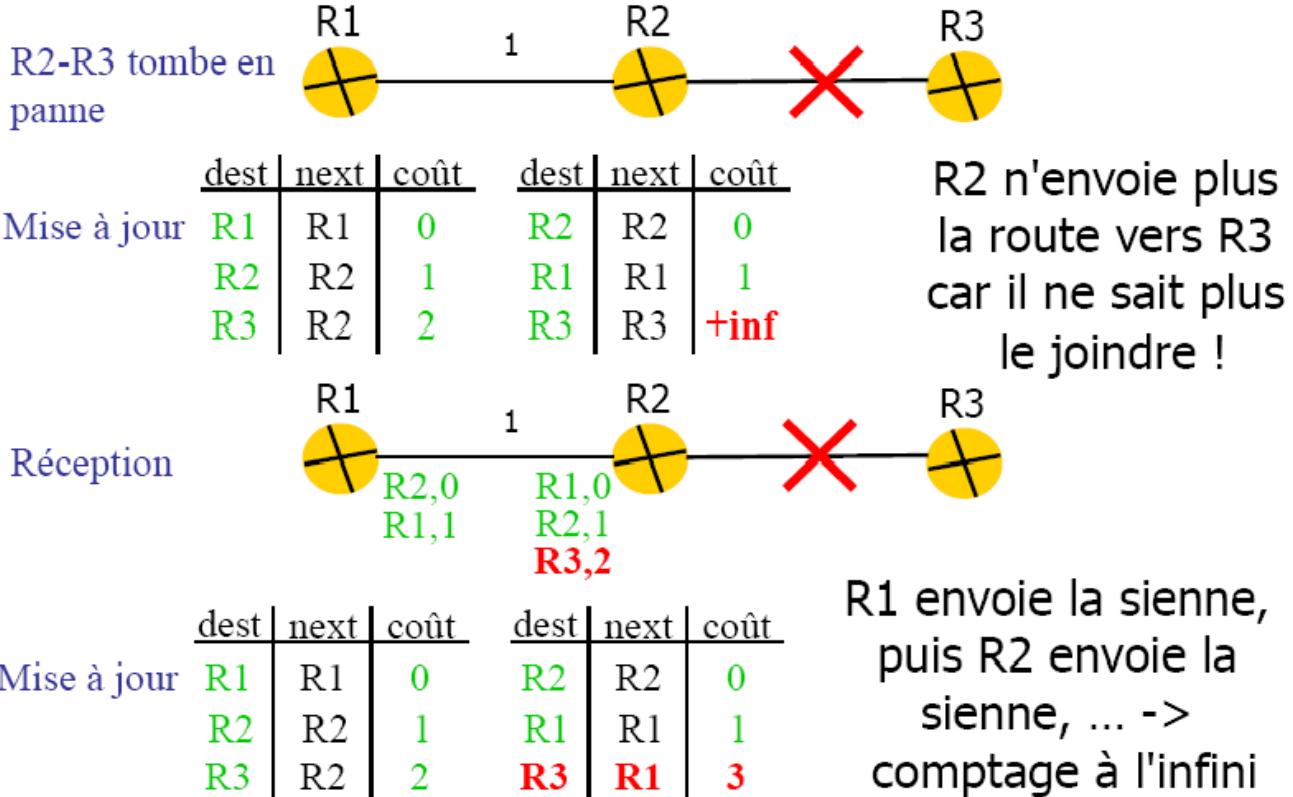


	dest	next	coût
Etat initial	R1	R1	0
	R2	R2	0
	R3	R3	0

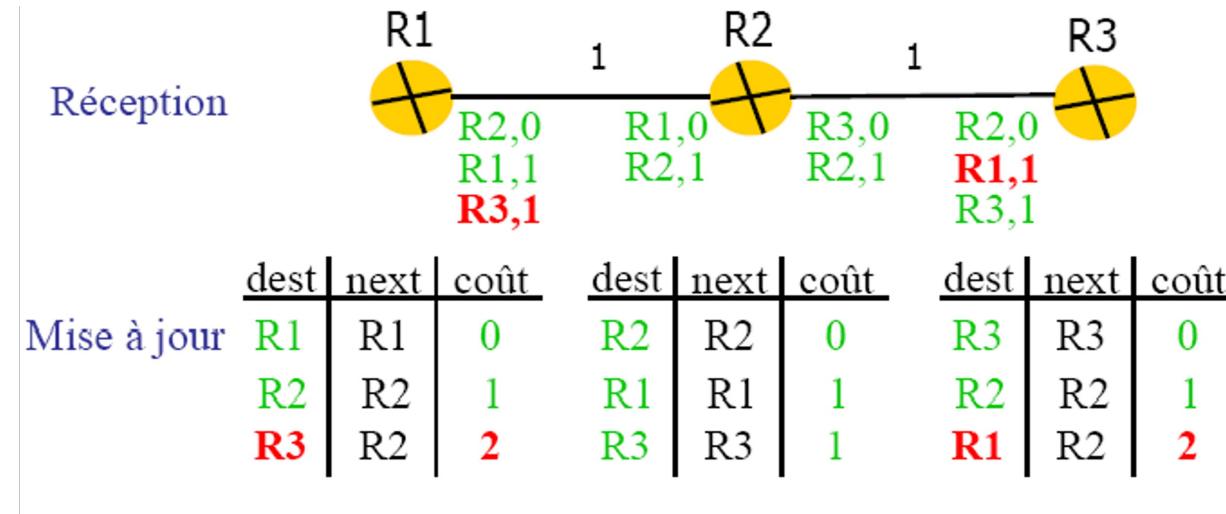


	dest	next	coût
Mise à jour	R1	R1	0
	R2	R2	1
	R3	R3	1
	R1	R2	0
	R2	R1	1
	R3	R3	0
	R2	R2	1

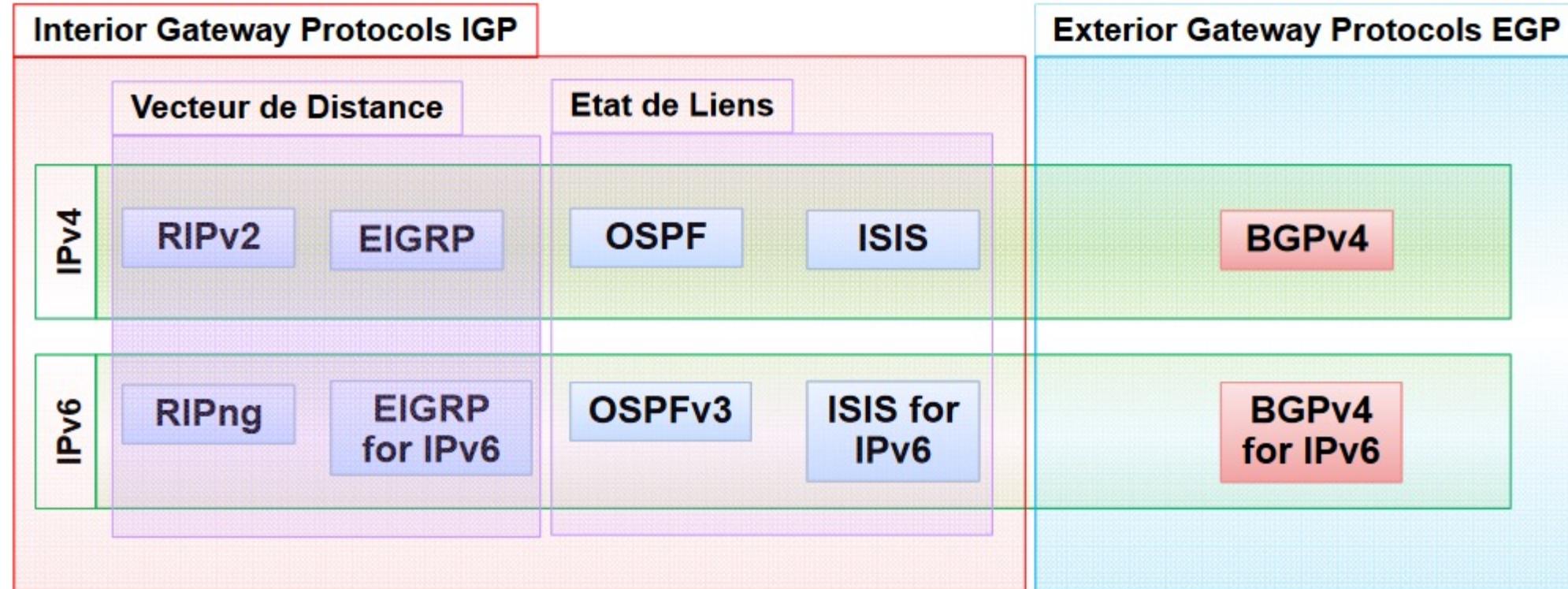
Exemple de Routage à vecteur de distance (2/3)



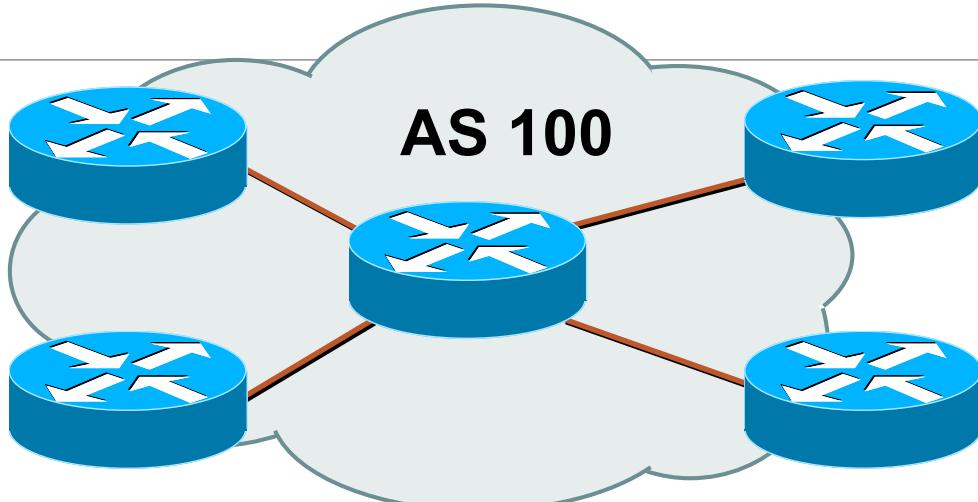
Exemple de Routage à vecteur de distance (3/3)



Routage dynamique: IPv4 & IPv6



Système autonome (AS)



Ensemble de réseaux partageant la même politique de routage

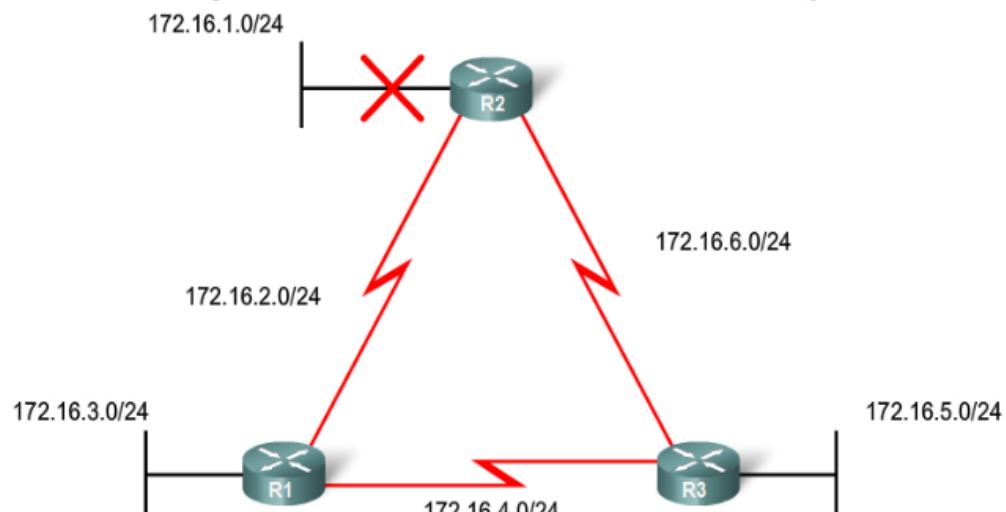
Utilisation d'un même protocole de routage

Généralement sous une gestion administration unique

Utilisation d'un IGP au sein d'un même AS

Routage dynamique

- On parle de convergence lorsque les tables de routage de tous les routeurs sont **parfaitement cohérentes**
- Un réseau n'est pas complètement **opérationnel** tant qu'il n'est pas convergé.
- Le temps de convergence doit être **bref**
- RIP met du temps à converger, EIGRP et OSPF sont **plus rapides**



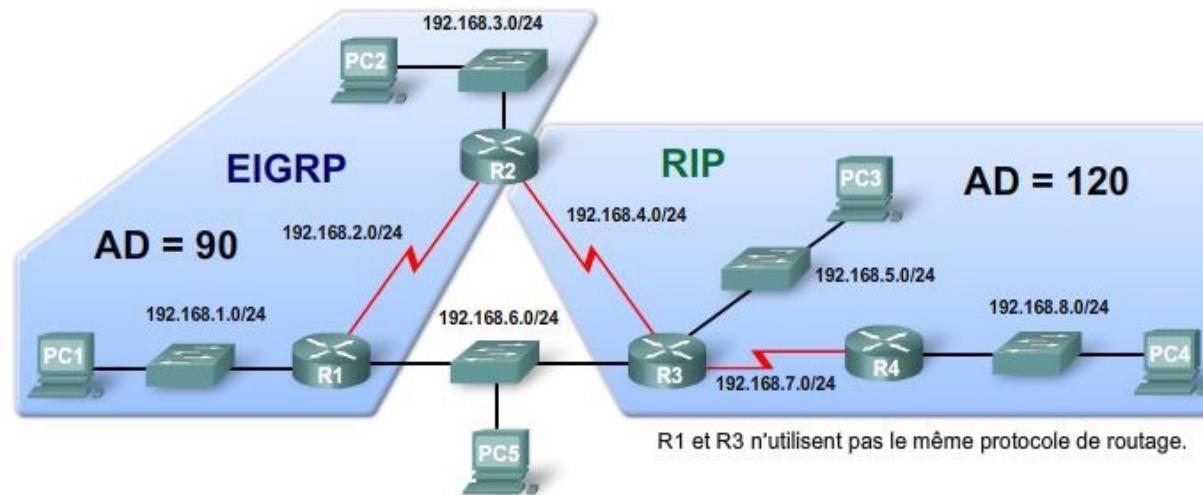
Distance Administrative (AD)

- La distance administrative (AD) définit la préférence d'une source de routage.
- Chaque source de routage est classée par ordre de priorité, du plus préférable au moins préférable, à l'aide d'une valeur de distance administrative.
- Les routeurs Cisco utilisent la distance administrative (AD) pour sélectionner le meilleur chemin lors de la découverte du même réseau de destination à partir d'au moins deux sources de routage différentes.
- La distance administrative est une valeur entière comprise entre 0 et 255. Plus la valeur est faible, plus la source de la route est privilégiée. Une distance administrative de 0 est idéale.
- Seul un réseau directement connecté a une distance administrative égale à 0, laquelle ne peut pas être modifiée.
- Une distance administrative de 255 signifie que le routeur ne se fiera pas à la source de cette route et qu'elle ne sera pas installée dans la table de routage.

Distances administratives par défaut

Origine de la route	Distance administrative
Connecté	0
Statique	1
Récapitulatif de routage du protocole EIGRP	5
Protocole BGP externe	20
EIGRP interne	90
Protocole IGRP	100
Protocole OSPF	110
Protocole de routage IS-IS	115
Protocole RIP	120
Protocole EIGRP externe	170
Protocole BGP interne	200

Exemple de distance administrative



```
R2#show ip route
(**résultat omis**)

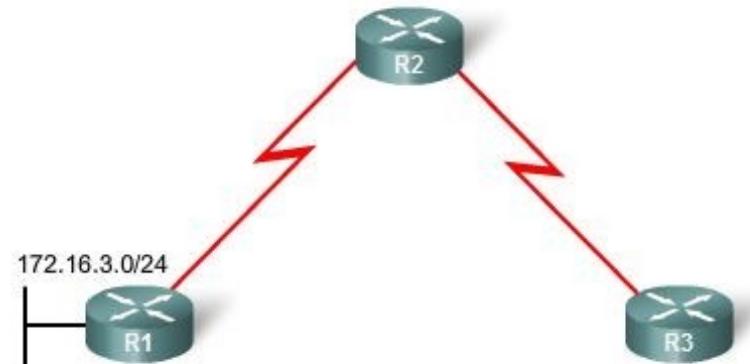
Gateway of last resort is not set

D  192.168.1.0/24 [90/2172416] via 192.168.2.1, 00:00:24, serial0/0/0
C  192.168.2.0/24 is directly connected, Serial0/0/0
C  192.168.3.0/24 is directly connected, FastEthernet0/0
C  192.168.4.0/24 is directly connected, Serial0/0/1
R  192.168.5.0/24 [120/1] via 192.168.4.1, 00:00:08, serial0/0/1
D  192.168.6.0/24 [90/2172416] via 192.168.2.1, 00:00:24, serial0/0/0
R  192.168.7.0/24 [120/1] via 192.168.4.1, 00:00:08, Serial0/0/1
R  192.168.8.0/24 [120/2] via 192.168.4.1, 00:00:08, Serial0/0/1
```

Mesures et Métriques

- La mesure est utilisée pour déterminer quel chemin est préférable en présence de plusieurs chemins vers le même réseau distant.
- RIP utilise le nombre de sauts.
- EIGRP utilise une combinaison de bande passante et de délai.
- OSPF fait appel à la bande passante.

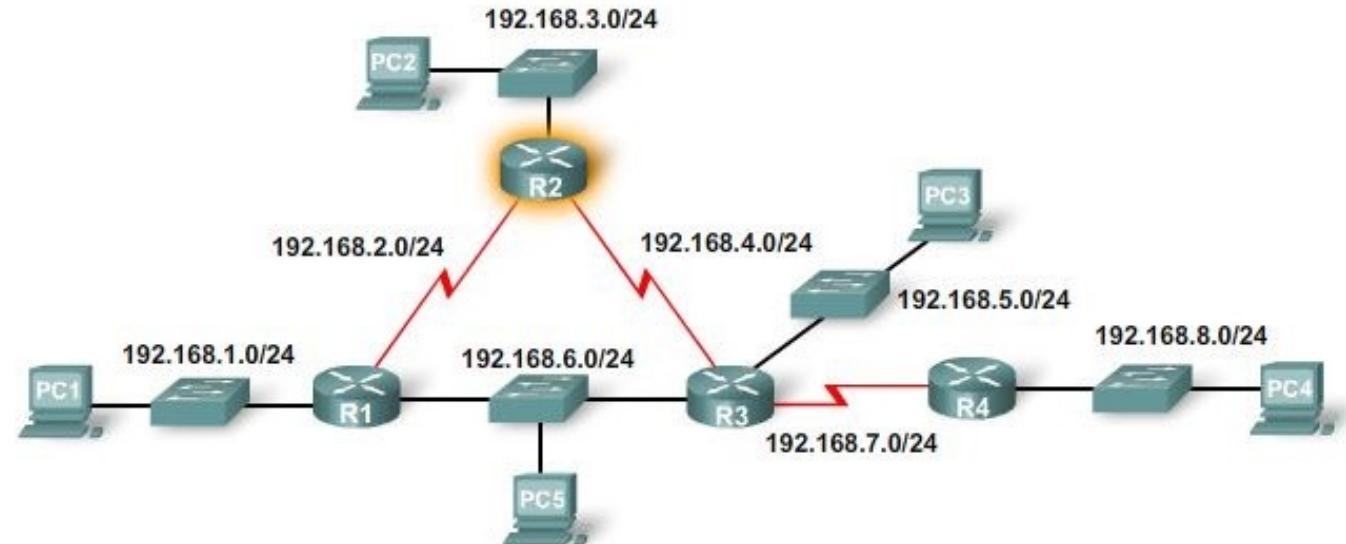
Réseau	Sauts
172.16.3.0	1



Réseau	Sauts
172.16.3.0	0

Réseau	Sauts
172.16.3.0	2

Exemple de mesure dans RIP



```
R2#show ip route
(**résultat omis**)

Gateway of last resort is not set

R    192.168.1.0/24 [120/1] via 192.168.2.1, 00:00:24, Serial0/0
C    192.168.2.0/24 is directly connected, serial0/0
C    192.168.3.0/24 is directly connected, FastEthernet0/0
C    192.168.4.0/24 is directly connected, Serial0/1
R    192.168.5.0/24 [120/1] via 192.168.4.1, 00:00:26, Serial0/1
R    192.168.6.0/24 [120/1] via 192.168.2.1, 00:00:24, Serial0/0
                           [120/1] via 192.168.4.1, 00:00:26, serial0/1
R    192.168.7.0/24 [120/1] via 192.168.4.1, 00:00:26, Serial0/1
R    192.168.8.0/24 [120/2] via 192.168.4.1, 00:00:26, Serial0/1
```

Elle se trouve à 2 sauts de R2 dans 192.168.8.0/24

Protocole RIP

- RIP «Routing Information Protocol», RFC 1058
- Routage IGP «Interior Gateway Protocol» interne à un système autonome (AS)
- Les routeurs n'ont pas une vue globale du réseau. Ils s'échangent les distances qu'ils connaissent
- Messages UDP sur le port 520 échangés toutes les 30s
- D'autres temporiseurs sont utilisés
- d'indisponibilité «invalid timer» : si le message d'un voisin n'est pas reçu au bout de 180s, le routeur annonce l'indisponibilité du voisin aux autres
- d'effacement «flush timer» : si rien ne vient informer l'indisponibilité au bout de 270s, l'entrée est supprimée de la table de routage

Protocole RIP

- A servi de base pour AppleTalk, IPX, Banyan Vines ...
- RIP v 2, RFC 1387 et RFC 1388
 - permet le routage des sous-réseaux (véhicule le netmask)
 - identifie les routes externes utilisées par un EGP
 - interopère "raisonnablement" avec RIP v 1
 - diffusion multicast (224.0.0.9) plutôt que broadcast
 - authentification

Algorithmes de RIP

- Chaque **routeur** envoie à tous ses voisins périodiquement (30s) une liste (**@réseau, coût**) qu'il connaît d'après sa table de routage
- A la réception d'un paquet RIP arrivant d'un routeur voisin d'adresse **@routeur**, pour chaque (**@destination, cout**) contenue dans le paquet reçu faire
 - Si **@destination** inconnue
 - Rajouter dans la table de routage l'entrée (**@destination, @routeur, coût+1**)
 - Armer timer
 - Si **@destination** connue
 - Si **@routeur_voisin= @routeur** alors
 - Changer table (**@destination, @routeur, coût+1**)
 - Relancer timer
 - Si **@routeur_voisin ≠ @routeur** et **coût+1<coût_présent** alors
 - Changer table (**@destination, @routeur, coût+1**)
 - Relancer timer

Routage à état de liens

- Une liaison LAN est assimilée à un nœud dans le graphe représentant la carte du réseau
- Avantages : chaque routeur calcule indépendamment la carte du réseau, ainsi l'algorithme converge plus rapidement que le routage à vecteur de distance
- Inconvénients : nécessite un espace mémoire plus important pour le stockage de la carte du réseau, les calculs sont assez complexes
- Exemple : OSPF (Internet), IS-IS (DECnet, CLNP, Internet, Netware)
- Nous allons décrire l'OSPF

Routage à état de liens

- Chaque routeur :

- surveille activement l'état de tous ses liens
 - détermine ses voisins (HELLO)
 - mesure les coûts vers chacun de ses voisins (ECHO). La prise en compte de la charge des liens risque de créer des oscillations. Les coûts peuvent aussi être assignés statiquement.
- construit un message traduisant l'état des liens (LSA : «Link State Advertisements»)
 - adresse du routeur source
 - numéro de séquence et durée de vie (âge)
 - liste des voisins et les coûts d'acheminement associés
- diffuse ce message LSA à tous les autres routeurs via ses voisins.
- construit une carte topologique et calcule les plus courts chemins (algorithme de Dijkstra).

OSPF «Open Shortest Path First»

Le protocole OSPF est un protocole de routage dynamique IGP (Interior Gateway Protocol), à état de liens qui est ouvert. On peut donc l'utiliser avec des constructeurs hétérogènes (Cisco, HP, Alcatel...).

OSPF a été créé par le groupe de travail IETF en 1988, il utilise l'algorithme de Dijkstra (algorithme du plus court chemin).

Il existe 3 versions OSPF :

OSPF V1 : création du protocole OSPF

OSPF V2 : ajout de l'authentification

OSPF V3 : pour l'IP version 6

OSPF «Open Shortest Path First»

- RFC 1247, 1583
- Protocole ouvert non lié à un propriétaire
- Routage IGP «Interior Gateway Protocol» interne à un système autonome (AS)
- Accepte une variété de métriques : distances, délais, débits,...,
- Algorithme dynamique, capable de s'adapter aux changements topologiques,
- Acceptation du routage par « type de service » (particulier au traitement du champ service du datagramme IP)
- Gestion d'un niveau de sécurité destiné à éviter l'attaque des tables de routage

Que signifie état de liens ?

Il y a une notion d'état de liens dans le protocole OSPF, il est indispensable de bien la comprendre.

Un état de liens dans OSPF est une description de l'interface d'un routeur avec les éléments suivants :

- Son adresse IP
- Son masque
- Le type de réseau
- Son voisin (un routeur)
- L'ensemble des liens OSPF est enregistré dans une base de données appelée link-state database, qui est identique sur tous les routeurs d'une même aire.

Pour mettre à jour sa base de données on utilise des LSA (Link State Advertisements) nous y reviendrons plus tard, car il existe différents types de LSA.

Les caractéristiques OSPF

Les caractéristiques notables du protocole OSPF sont les suivantes :

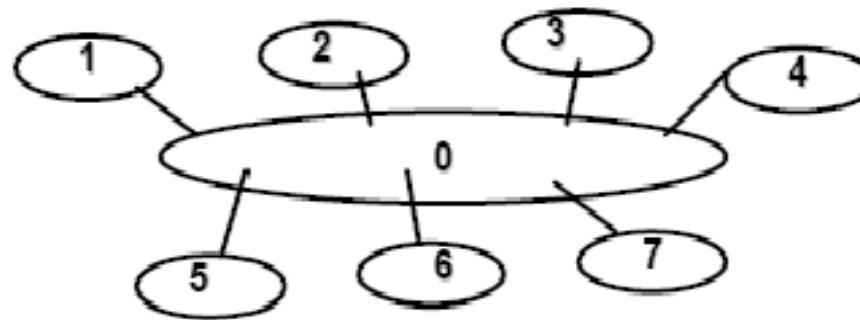
- ❖ Convergence rapide les mises à jour sont incrémentielles
- ❖ Utilisé dans les grands réseaux (supérieur à 15 routeurs contrairement au protocole RIP)
- ❖ OSPF utilise une notion d'aire qui permet de réduire le trafic réseau (nous y reviendrons plus tard dans ce cours)
- ❖ Les aires OSPF sont hiérarchisées, l'aire 0 est obligatoire (aire Backbone)
- ❖ Supporte le VLSM (netmask à longueur variable)
- ❖ Adresses multicast réservées à OSPF pour les MAJ
- ❖ Authentification possible sous OSPF
- ❖ Prends en compte les routes venant d'autres protocoles
- ❖ La métrique utilise la bande passante
- ❖ C'est un protocole ouvert
- ❖ La distance administrative est égale à 110.

En contrepartie OSPF a les trois inconvénients suivants :

- ❖ OSPF consomme de la mémoire, car il maintient une base de données pour chaque routeur (état des liens)
- ❖ Consomme de la CPU au démarrage du processus OSPF, car il doit construire sa base de données
- ❖ Protocole complexe à mettre en place

OSPF «Open Shortest Path First»

- Le routage est hiérarchisé : un système autonome (AS) est découpé en zones ou aires «area»
- Une zone est un ensemble de réseaux contigus. Chaque zone se comporte comme un réseau indépendant, elle ne connaît que l'état des liaisons internes à la zone
- La zone 0 est l'épine dorsale d'un système autonome. **Toute autre zone est connectée à cette épine dorsale**, soit directement, soit par un tunnel



- Deux niveaux de routage : intra et inter zone
- Gérer une topologie hiérarchique (les « bords » du réseau sont organisés en arbre alors que le centre est en graphe)

Zone OSPF

Une caractéristique principale d'OSPF est de supporter des inter réseaux très larges grâce au regroupement des routeurs dans des entités logiques appelées areas ou zones.

La communication inter-areas ne laisse passer que les échanges d'information minimale de routage dans le seul objectif de connecter les zones entre elles.

Il en résulte que les efforts de calcul de routes ne s'opèrent qu'au sein d'une même zone.

Les routeurs d'une zone ne sont pas affectés (en calcul) par les changements intervenus dans une autre zone.

Dans un contexte où OSPF demande beaucoup de ressources en CPU et en mémoire, cette notion de conception est très importante.

OSPF «Open Shortest Path First»

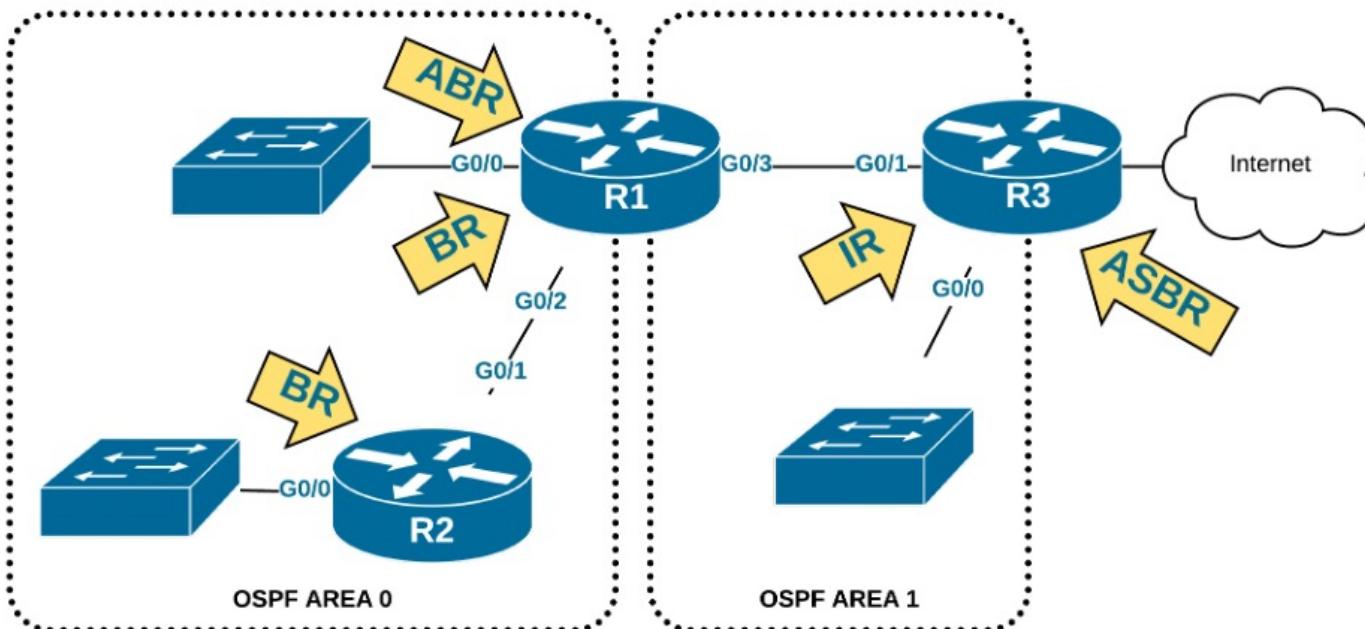
- A l'intérieur d'une zone

- Chaque routeur dispose d'une base de données topologique (informations d'état des liens)
- Le même algorithme du plus court chemin
- Un routeur au moins connecté à l'épine dorsale
- Si un routeur est connecté à deux zones, il doit exécuter l'algorithme du plus court chemin pour les deux zones

Les routeurs dans les zones OSPF

Dans une zone OSPF, les routeurs jouent plusieurs rôles en fonction de leur positionnement dans l'infrastructure, ce qui donne :

- **Routeur ABR (Area Border Router)** : il fait le lien entre 2 zones OSPF
- **Routeur Interne (IR)** : routeur dans une zone OSPF
- **Routeur Backbone** : routeur situé dans l'aire 0 (aire Backbone)
- **Routeur ASBR (Autonomous System Area Border Router)** : routeur faisant le lien entre le protocole OSPF et un autre protocole de routage (redistribution de routes entre OSPF et RIP dans les deux sens, par exemple)

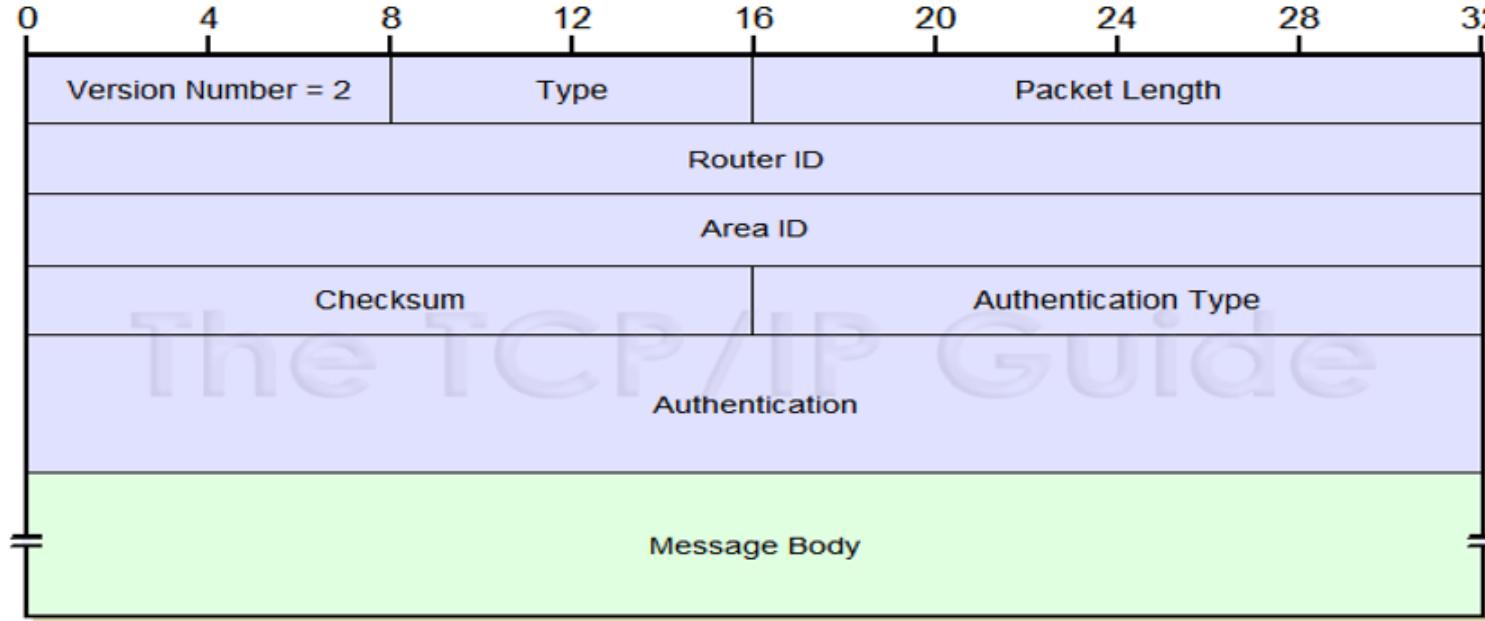


OSPF «Open Shortest Path First»

- Trois types de chemins :
 - **Chemin intra-zone**: le plus simple, puisque chaque routeur d'une zone connaît la topologie de la zone
 - **Chemin inter-zone**, il nécessite 3 étapes :
 - Aller de la source vers l'épine dorsale (dans la zone source)
 - Transiter à travers l'épine dorsale jusqu'à la zone de destination
 - Transiter dans la zone destination jusqu'à la destination
 - **Chemin inter-systèmes autonomes**
 - Nécessite un protocole particulier (BGP)
- Bases de données maintenues sur un routeur :
 - « Adjacencies Database » => les voisins
 - « Topology Database » => toutes les routes
 - « Forwarding Database » => les meilleures routes

Les paquets OSPF

L'échange d'informations entre les routeurs s'effectuent au travers différents paquets, ceux-ci ont chacun un rôle bien spécifique.



Type: Indicates the type of OSPF message:

Type Value	OSPF Message Type
1	Hello
2	Database Description
3	Link State Request
4	Link State Update
5	Link State Acknowledgment

■ Messages OSPF

- Annonce (*Hello Packet*, type 1) : utilisé pour la détection des routeurs, le choix des routeurs désignés. Transmis à l'adresse 224.0.0.5 (*AllSPFRouters*).
- Description de base de données (*Database Description Packet* type 2): permet lors de l'apparition d'un nouveau routeur de lui transférer le contenu de la base de données.
- Message de demande de données (*Link State Request Packet*, type 3) : permet d'obtenir d'un routeur voisin une mise à jour de la base de données
- Message de mise à jour de l'état des liaisons (*Link State Update Packet*, type 4) : transmet les modifications de topologie :
 - sur un réseau local, aux routeurs désignés (adresse 224.0.0.6, *AllDRouters*), le routeur désigné transmet l'information aux autres (224.0.0.5, *AllSPFRouters*).
 - sur une liaison point à point, à un routeur distant.
- Message d'acquittement d'état des liaisons (*Link State Acknowledgment Packet*, type 5) : acquitte la mise à jour de la base de données.

■ Relations avec les voisins

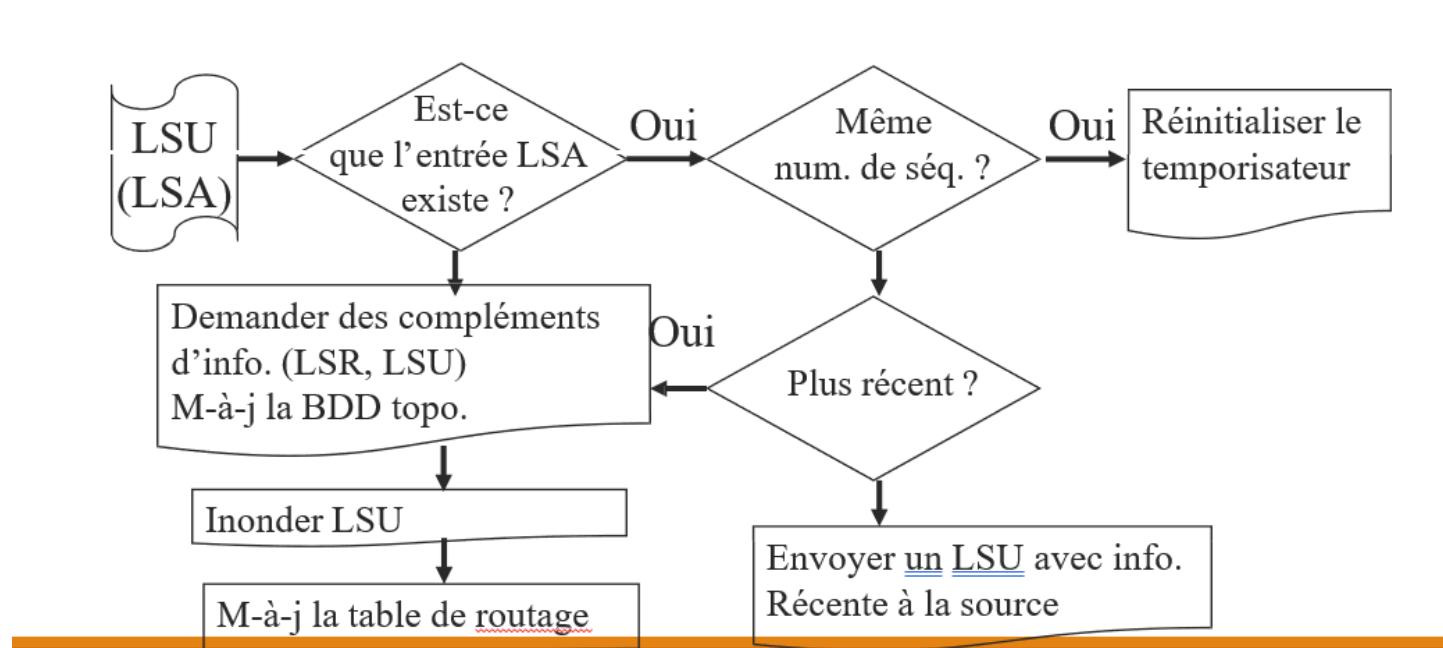
- Découverte des routeurs voisins (adjacents), du DR et du BDR
 - **Message** de type **Hello** envoyé par un routeur à ses voisins donnant une vision sur le voisinage du routeur :
 - identité (adresse IP la plus élevée parmi les adresses de ses interfaces),
 - priorité,
 - période du hello « HelloInterval »
 - temporisation de panne « RouterDeadInterval »,
 - DR, BDR,
 - identités des voisins qui ont envoyé récemment des paquets hello
 - identité de la zone ...

■ Initialisation d'un lien-

- Établissement d'une relation de type « Maître/esclave » (le maître a l'identité la plus élevée)
- Le maître envoie un **message DBD** (« **D**atabase **D**escription ») qui représente un résumé de l'état des liens, sous forme d'un ou plusieurs paquets, chaque paquet possède un numéro de série
- L'esclave acquitte les paquets DBD (« **L**ink **S**tate **A**cknowledgment ») et envoie un message DBD
- Le maître acquitte les paquets DBD
- Une fois les DBD échangés, si un routeur constate que certaines entrées sont moins récentes, il envoie une requête **LSR** « **L**ink **S**tate **R**equest » à laquelle répond le routeur adjacent (ayant l'info. La plus récente) par un message **LSU** « **L**ink **S**tate **U**pdate »

Maintient de l'information de routage

- Lors du changement de l'état d'un lien un LSU est envoyé (inondé)
- Un LSA (« Link state Announcement ») est envoyé quand l'état d'une ligne change ou toutes les 30 minutes



Le paquet LSA

- Il existe 5 types de messages LSA

- Type 1 « Router-LSA » : information concernant l'états des liens envoyée par tout routeur interne, inondée dans la zone
- Type 2 « Network-LSA » : description de la liste des routeurs sur LAN envoyée par le DR, inondée dans la zone
- Type 3 « Summary-LSA type 3 » : description des destinations en dehors de la zone (mais dans le même AS) envoyée par tout routeur ABR (inter-zone), vers tous les routeurs de la zone
- Type 4 « Summary-LSA type 4 » : description des routeurs ASBR (AS Boundary Router) envoyée par les ABR, vers tous les ABRs et ensuite inondée dans les zones respectives (à ces ABR)
- Type 5 « AS-External-Link » : description des destinations en dehors du AS envoyée par les ASBR, inondée dans le AS

Activation du protocole de routage OSPF et des réseaux participant aux annonces
Le numéro du processus ospf est 100.

L'adresse IP de l'interface connectée à un autre routeur est 10.0.3.1/24 et le
numéro de l'aire est 0.

```
R1(config)#router ospf 100
R1(config-router)#network 10.0.3.0 0.0.0.255 area 0
R1(config-router)#network 10.0.0.0 0.0.0.255 area 0
```

Affichage des informations concernant ospf

```
R1#show ip ospf
Routing Process "ospf 100" with ID 172.16.0.1
Start time: 00:18:40.612, Time elapsed: 00:08:28.352
Supports only single TOS(TOS0) routes
Supports opaque LSA
Supports Link-local Signaling (LLS)
Supports area transit capability
Router is not originating router-LSAs with maximum metric
Initial SPF schedule delay 5000 msec
Minimum hold time between two consecutive SPFs 10000 msec
Maximum wait time between two consecutive SPFs 10000 msec
```

```
R1#show ip ospf neighbor

Neighbor ID Pri State Dead Time Address Interface
172.16.0.3 1 FULL/BDR 00:00:39 10.0.3.2 FastEthernet2/0
172.16.0.2 1 FULL/DR 00:00:39 10.0.0.2 FastEthernet1/0
R1#
R1#
R1#show ip ospf neighbor detail
Neighbor 172.16.0.3, interface address 10.0.3.2
In the area 0 via interface FastEthernet2/0
Neighbor priority is 1, State is FULL, 34 state changes
DR is 10.0.3.1 BDR is 10.0.3.2
Options is 0x12 in Hello (E-bit, L-bit)
Options is 0x52 in DBD (E-bit, L-bit, O-bit)
LLS Options is 0x1 (LR)
Dead timer due in 00:00:35
Neighbor is up for 00:11:57
```

Afficher les voisins ospf

BGP «Border Gateway Protocol»

- Rôle du BGP
 - Utilisé pour échanger des informations de routage entre systèmes autonomes
- Méthode path vector
 - Les routes sont créées progressivement par addition des domaines : chaque route reçue est complétée par l'identification du domaine local et transmise aux voisins
 - Détection des boucles : une route reçue contenant le domaine local est terminée

BGP «Border Gateway Protocol»

- | BGP-1 RFC 1105, BGP-2 RFC 1163, BGP-3 RFC 1267, BGP-4 RFC 1654 1655
 - ni vraiment distance vector ni vraiment link state
 - ➔ Distance vecteur avec chemins (**Path-vector**)
 - entre routeurs configurés pour s'accepter mutuellement d'être voisins
 - transmet le "chemin d'AS" en entier ce qui permet la **détection** simple et efficace des **boucles** (adapté à des topologies complexes maillées)
 - échange des informations de routage par une connexion TCP afin d'assurer la **fiabilité** des échanges
 - Mise à jour **incrémentale** (**partielle**)
 - Mise à jour **seulement en cas de modification**
 - Utilisation de **métrique**
 - inclut un système **d'authentification** des messages échangés

Vocabulaire

Voisin (Neighbor)

- Routeur avec qui on a une session BGP

NLRI/Préfixe

- NLRI - network layer reachability information
- Informations concernant l'accessibilité (ou pas) d'une route (réseau + masque)

Router-ID (identifiant de routeur)

- Adresse IP la plus grande du routeur

Route/Path (chemin)

- Préfixe (NLRI) annoncé par un voisin

Vocabulaire (2)

Transit - transport de vos données par un réseau tiers, en général moyennant paiement

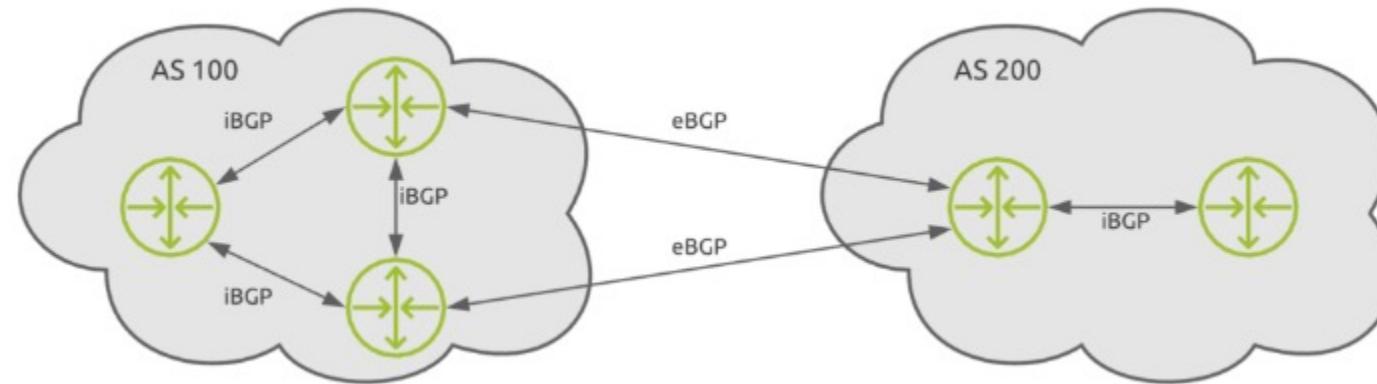
Peering - accord bi-latéral d'échange de trafic

- chacun annonce uniquement ses propres réseaux et ceux de ses clients à son voisin

Default - route par défaut, où envoyer un paquet si la table de routage de donne aucune information plus précise

BGP «Border Gateway Protocol»

BGP est le protocole utilisé dans l'épine dorsale de l'Internet. Il permet aux organisations qui ont leur propre AS (généralement les fournisseurs de services Internet et les grandes organisations) de s'interconnecter avec d'autres. Ce type d'interconnexion entre AS s'appelle un peering.

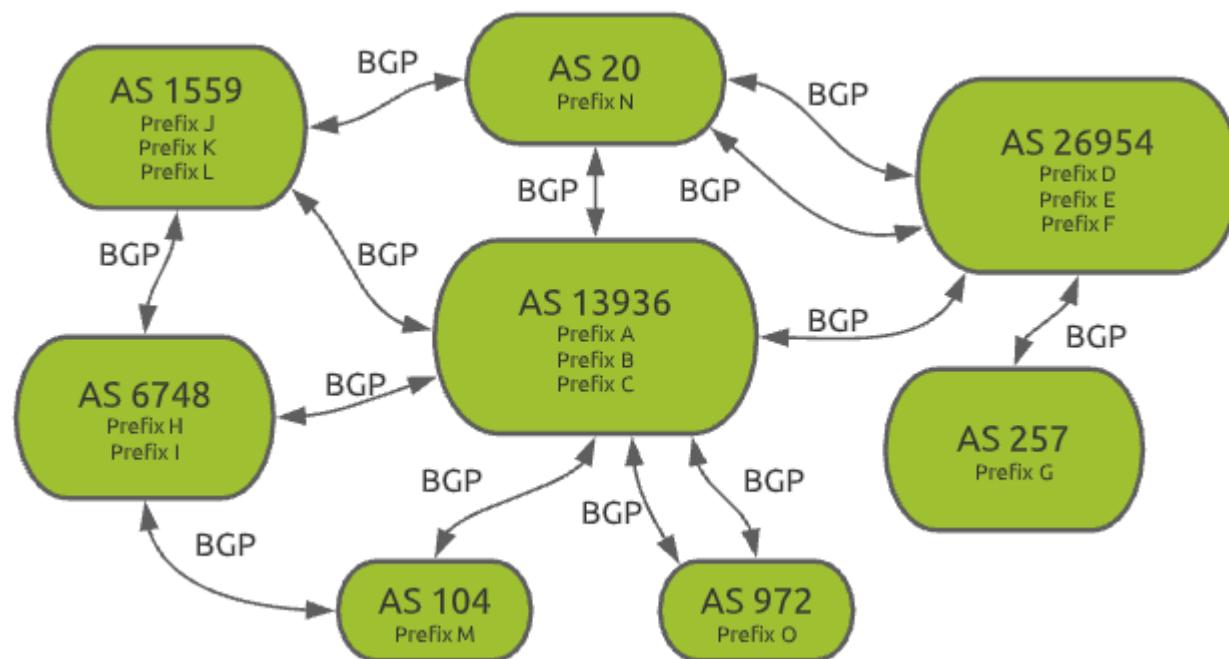


Le routage BGP

Contrairement aux autres protocoles de routage, il n'y a pas de processus de découverte des pairs.

Chaque intervenant BGP, appelé « pair », échange des informations de routage avec ses pairs voisins sous la forme d'annonces de préfixes réseau.

Avec les annonces de préfixes, l'information est suffisante pour construire un graphique de la connectivité des AS, comme illustré ci-dessous



Comme vous pouvez le constater, la communication entre deux préfixes peut souvent se faire par des chemins différents.

Le préfixe J de l'AS 1559 peut par exemple atteindre le préfixe G de l'AS 257 via l'AS 20 ou l'AS 13936.

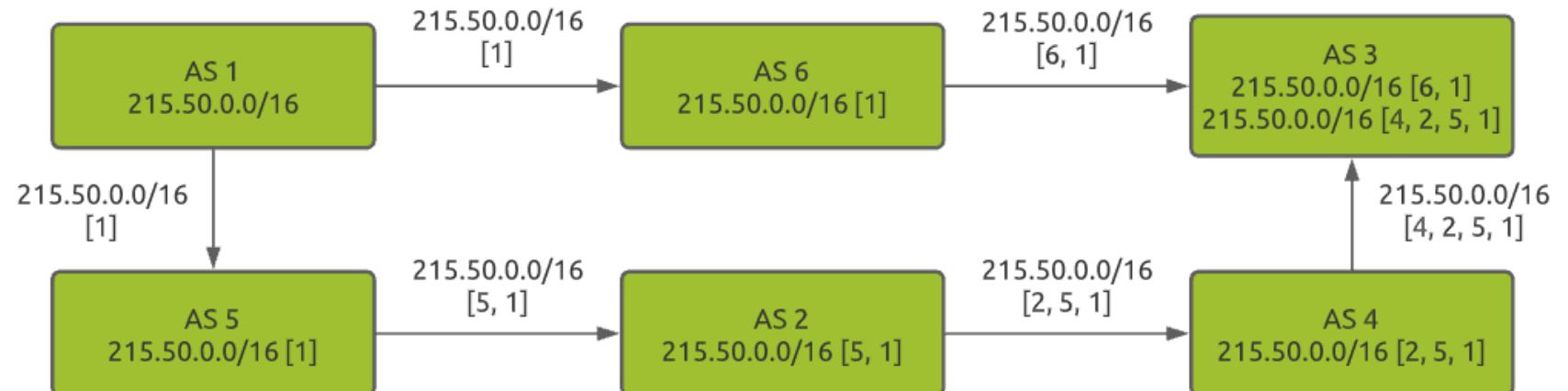
Alors comment les routeurs choisissent-ils entre les différentes routes possibles ?

La décision de routage

- Le chemin de l'AS BGP

BGP ne fonctionne pas comme les autres protocoles de routage traditionnels qui utilisent des mesures comme la distance ou les coûts (par exemple celui de la bande passante) pour prendre des décisions de routage. Au lieu de cela, BGP utilise divers attributs pour acheminer le trafic.

L'attribut principal de BGP est appelé « chemin AS ». Il s'agit d'une liste de numéros d'AS décrivant le chemin inter-AS vers une destination. Le chemin AS est si important pour la fonction de BGP que le protocole est souvent appelé protocole de routage à vecteur de chemin.



La figure ci-dessus montre comment le chemin AS est propagé.

le meilleur chemin est choisi en fonction des politiques, qui sont configurées via divers filtres de préfixes, en annonçant des routes spécifiques ou en manipulant les attributs BGP.
Lorsqu'une destination est accessible par deux chemins différents, BGP sélectionne le meilleur chemin en évaluant séquentiellement les attributs du chemin :

- Poids
- Préférence locale
- Origine
- Longueur du chemin AS
- Code d'origine
- MED (Multi Exit Discriminator)
- Chemin eBGP sur chemin iBGP
- Chemin IGP le plus court vers le prochain saut BGP
- Chemin le plus ancien
- ID du routeur
- Adresse IP du voisin.

Comment et quand les données du protocole de routage BGP sont-elles échangées ?

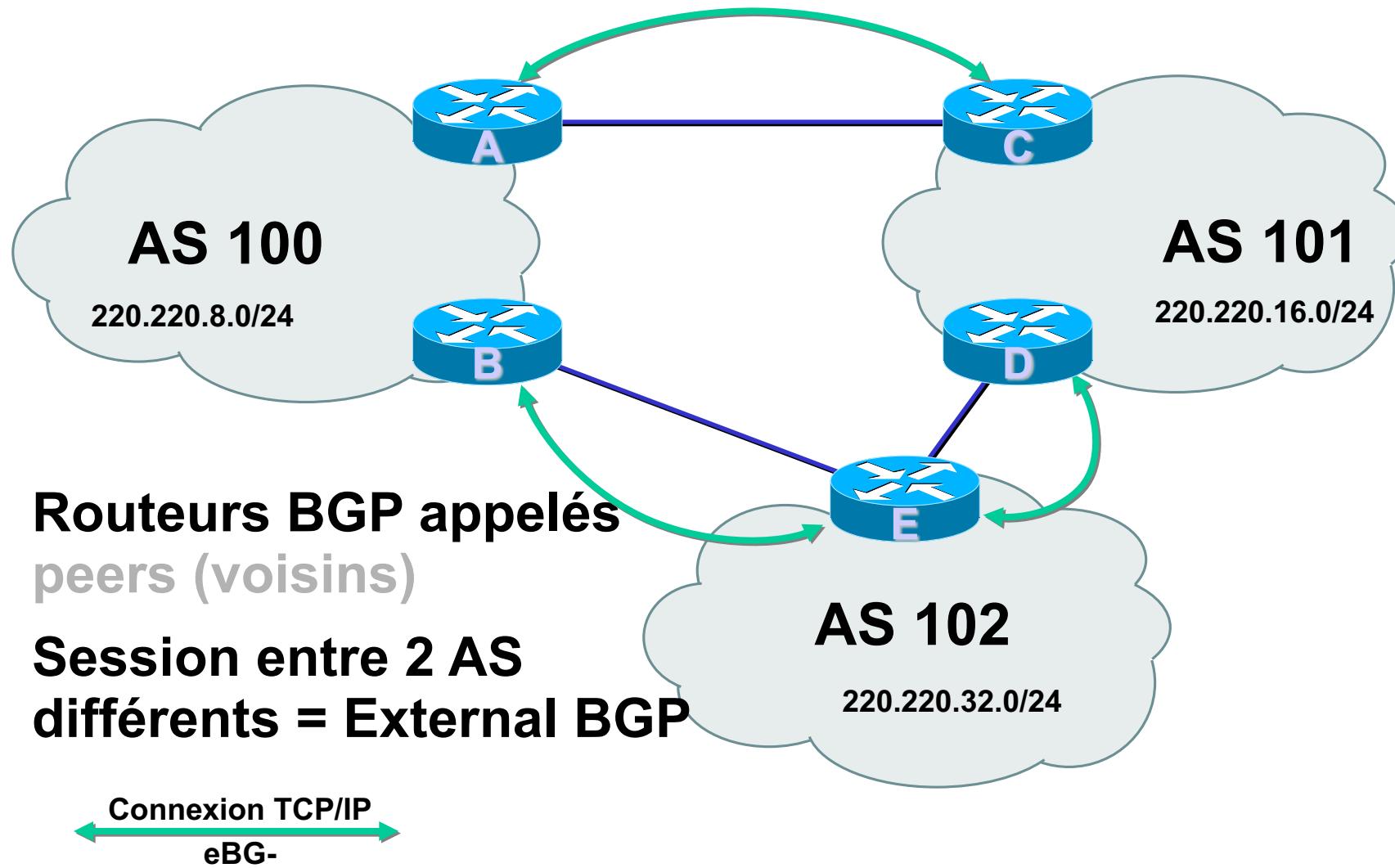
BGP utilise le protocole de transport TCP pour transférer les données. Cela permet une livraison fiable des mises à jour BGP. BGP utilise le port TCP 179 pour cela.

Il utilise le modèle FSM (Finite State Machine) pour maintenir une table de tous les pairs BGP et leur état opérationnel.

Par rapport aux autres protocoles de routage, BGP n'envoie pas de mises à jour périodiques des données de routage. Au lieu de cela, il envoie des mises à jour uniquement lorsque des changements se produisent sur le réseau. Par exemple, ces changements peuvent être dus à des réinitialisations de session, des défaillances de lien et des changements de politique.

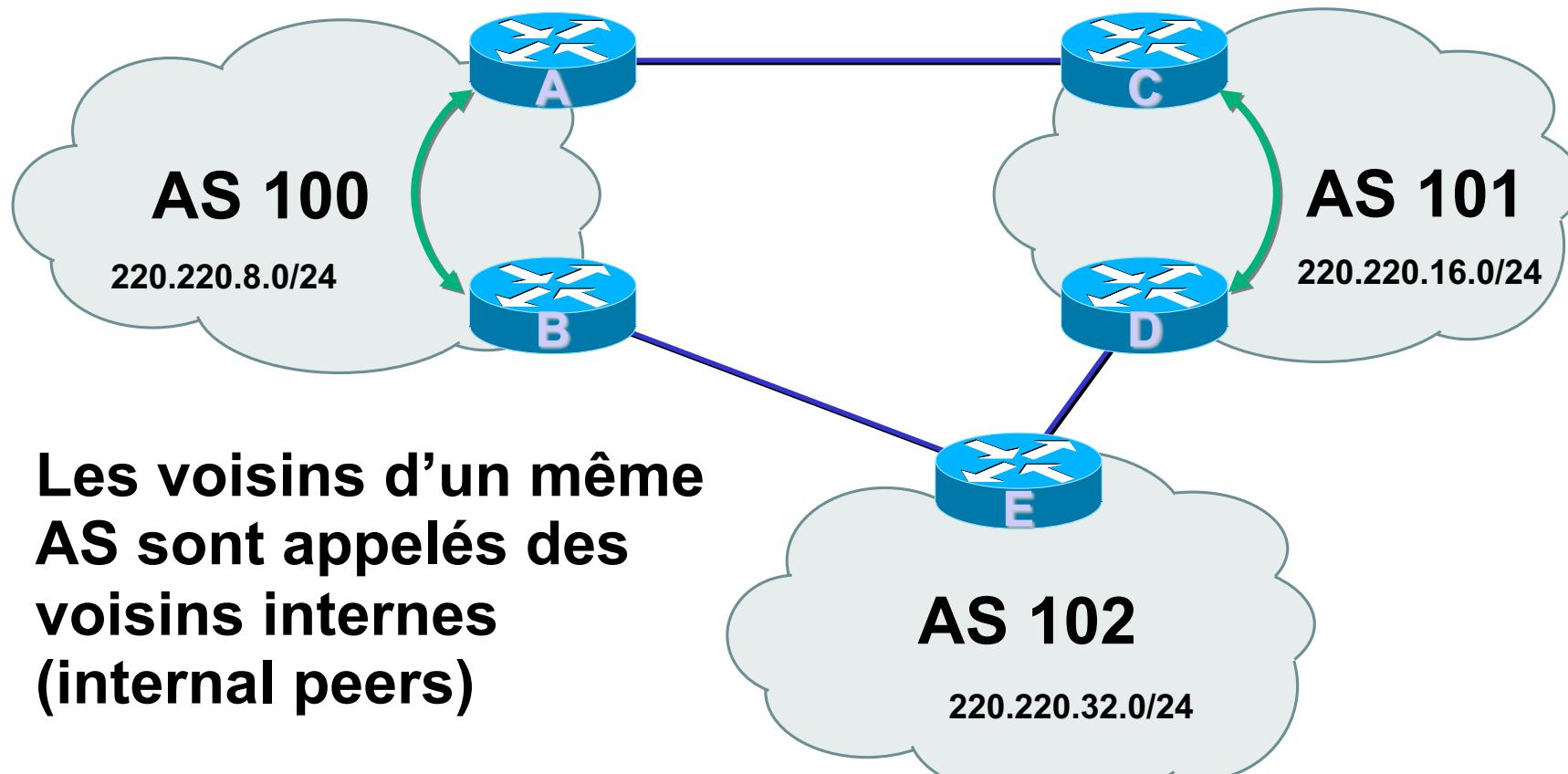
Enfin, BGP envoie périodiquement des messages keep-alive pour vérifier la connexion TCP.

Sessions BGP



Note: les voisins eBGP doivent être directements raccordés.

Sessions BGP

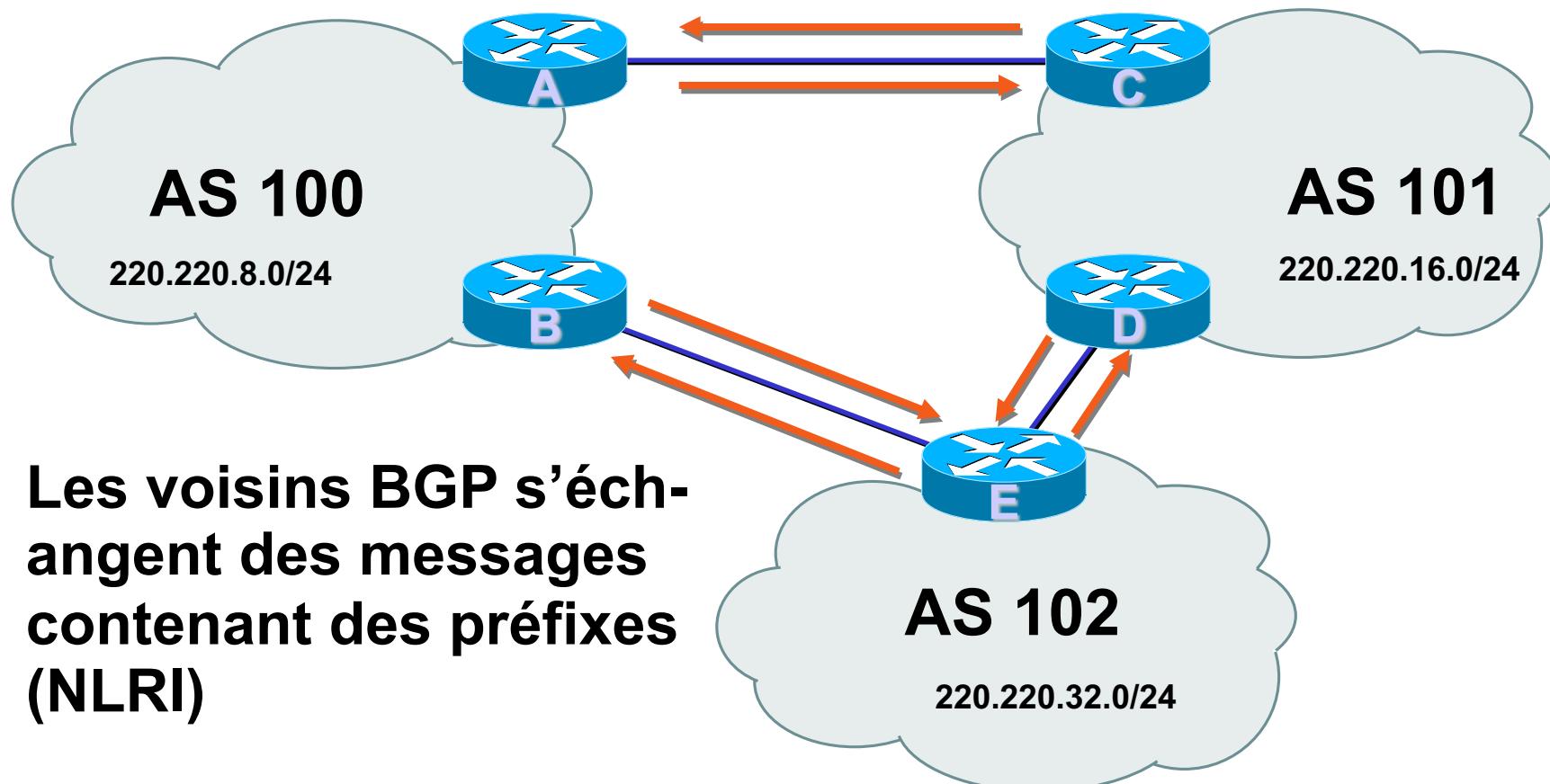


Les voisins d'un même AS sont appelés des voisins internes (internal peers)

Connexion TCP/IP
iBGP

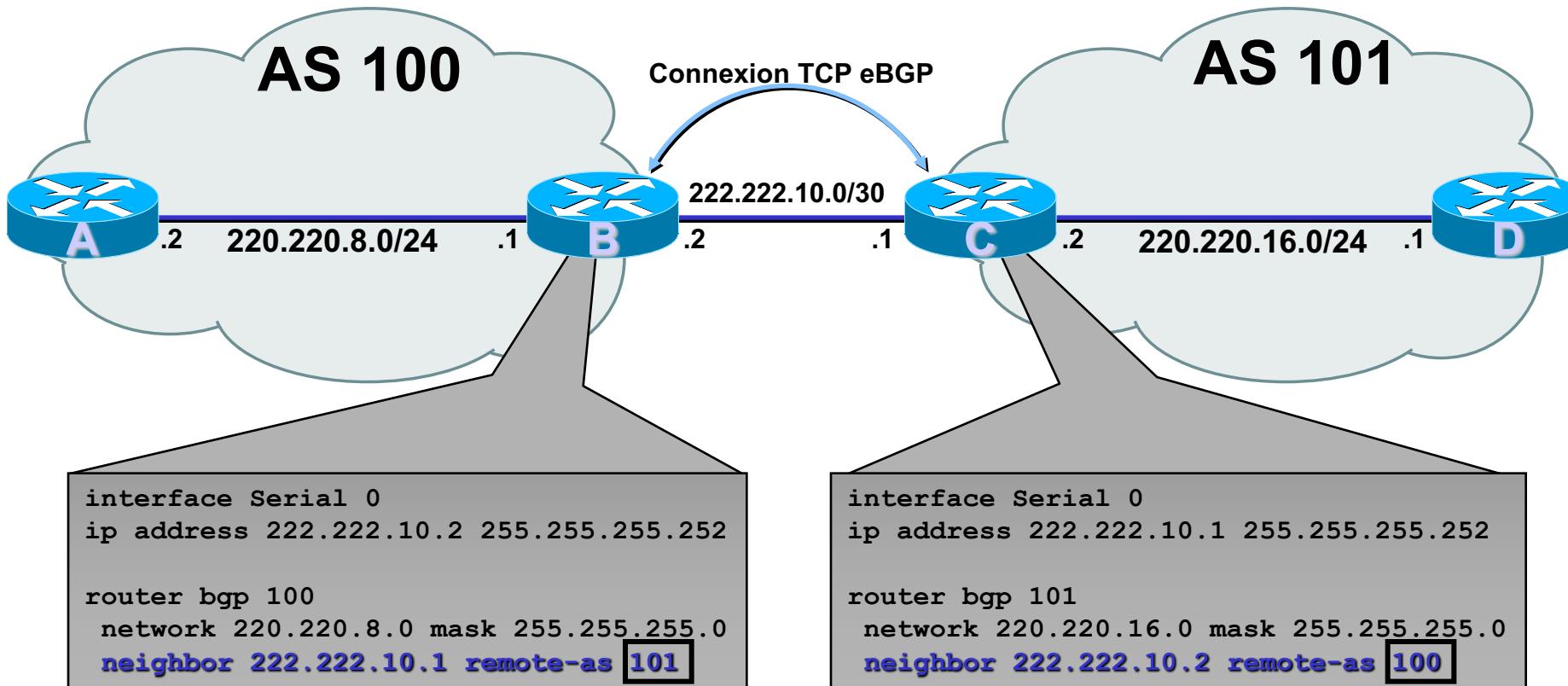
Note: les voisins iBGP peuvent ne pas être directement connectés.

Sessions BGP



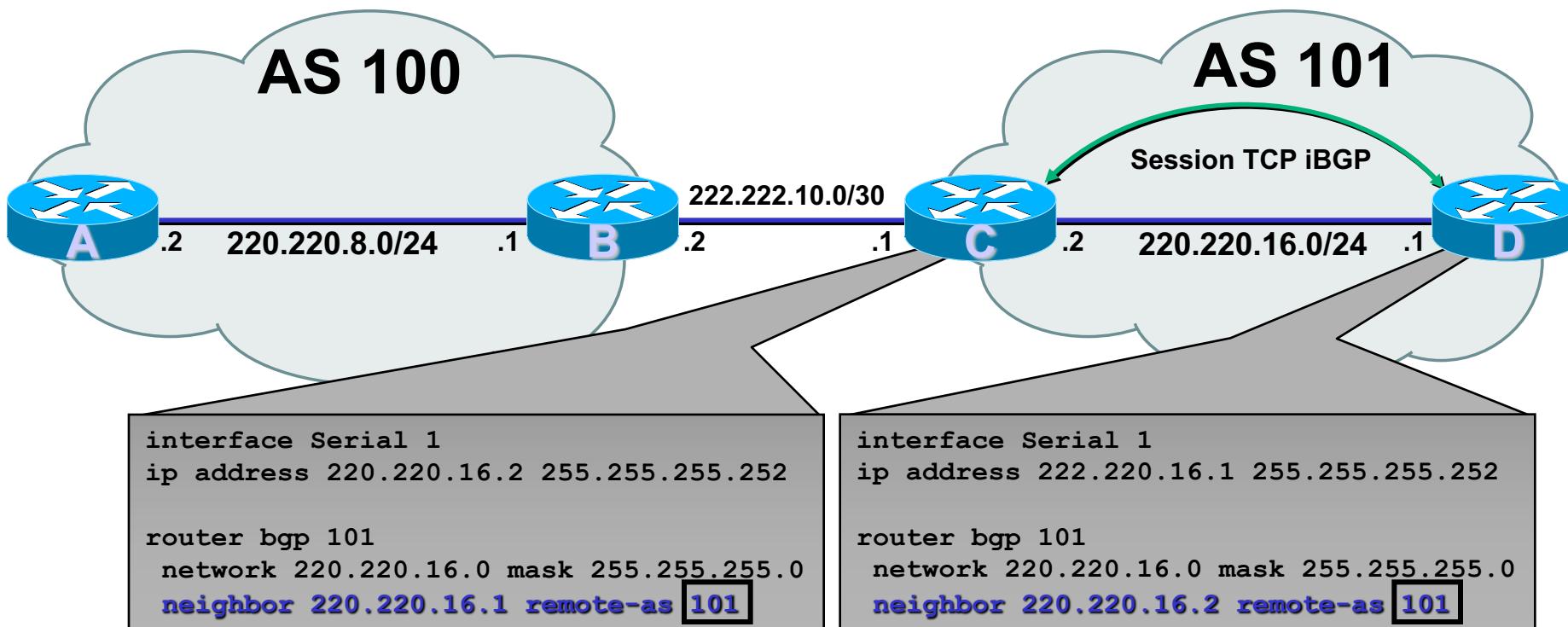
Message de mise
à jour BGP

Configuration de sessions BGP



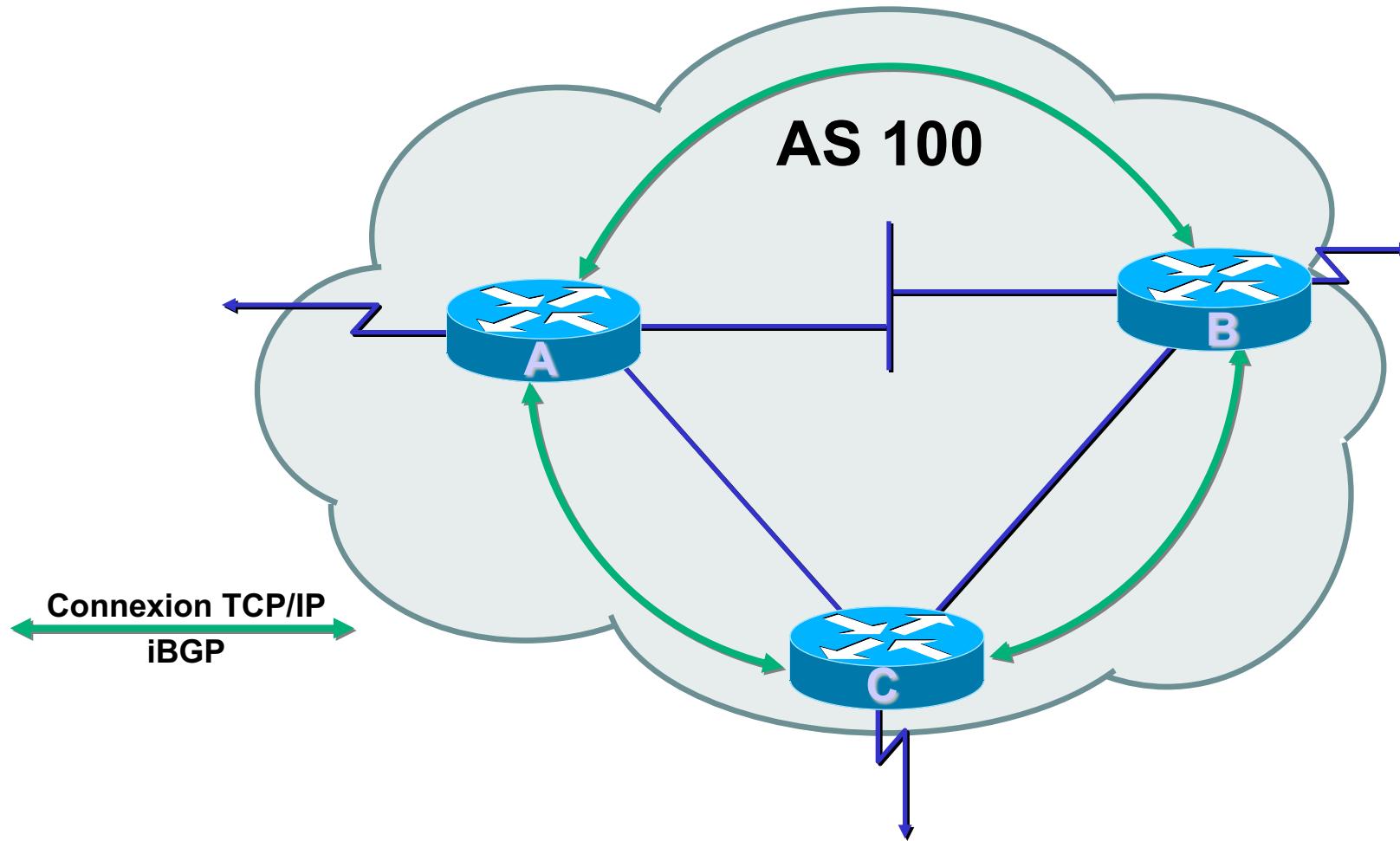
- Les sessions BGP sont établies en utilisant la commande BGP “neighbor” du routeur
 - Lorsque les numéros d'AS sont différents il s'agit d'une session BGP Externe (eBGP)

Configuration de sessions BGP



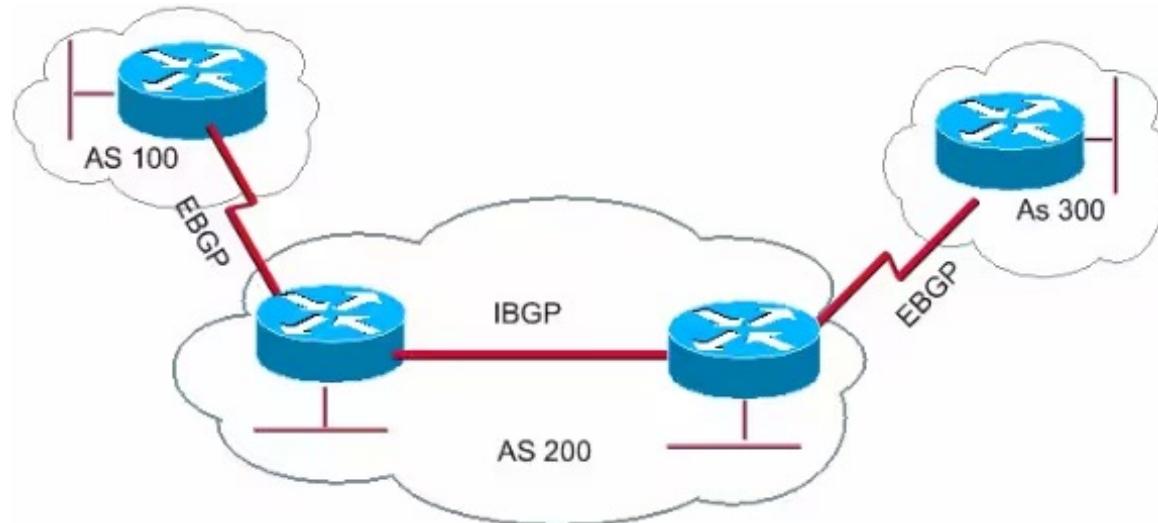
- Les sessions BGP sont établies en utilisant la commande BGP “neighbor” du routeur
 - Numéros d'AS différents -> BGP Externe (eBGP)
 - Numéros d'AS identiques -> BGP Interne (iBGP)

Configuration de sessions BGP



- Chaque routeur iBGP doit établir une session avec tous les autres routeurs iBGP du même AS

eBGP et iBGP



Si un AS comporte plusieurs speakers BGP, il peut servir de service de transit pour d'autres AS. Comme le montre le schéma suivant de cette section, AS200 est un AS de transit pour AS100 et AS300.

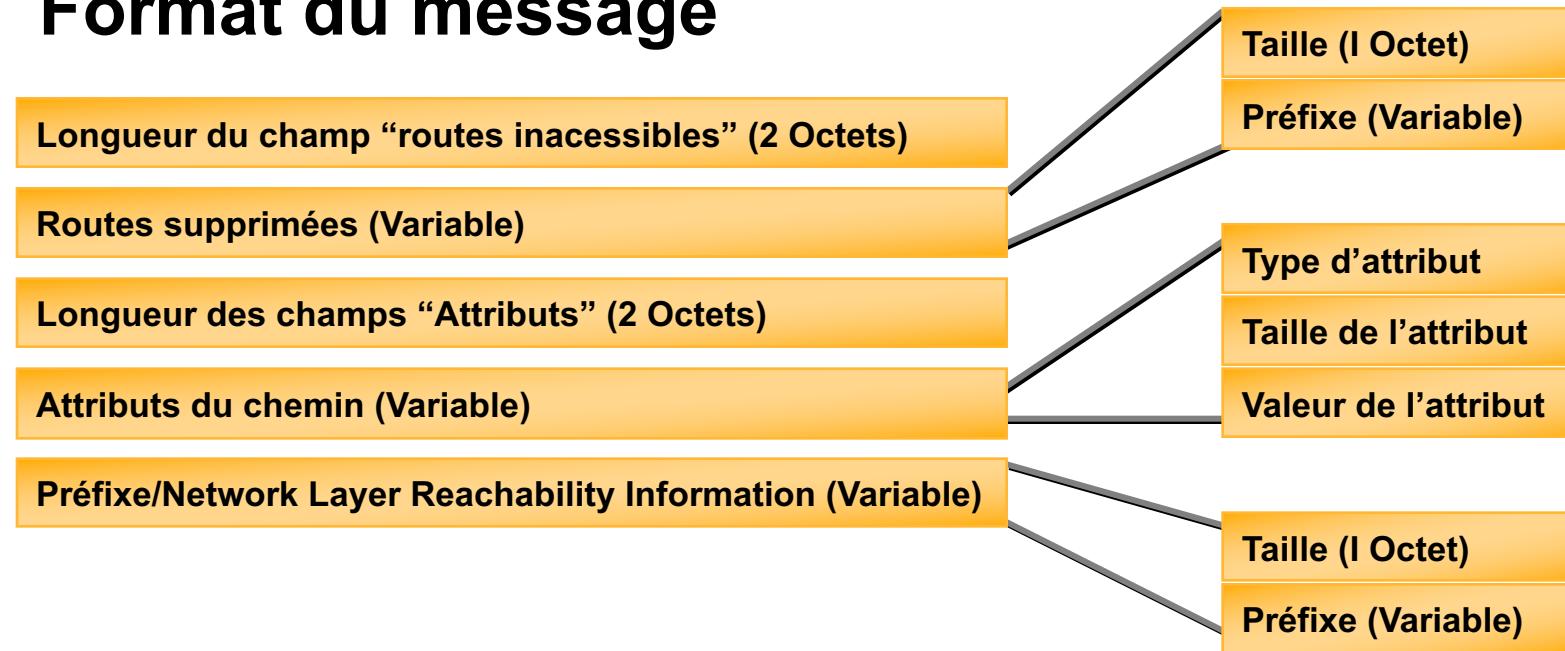
Afin d'envoyer l'information aux AS externes, l'accessibilité des réseaux doit être garantie. Afin d'assurer l'accessibilité des réseaux, les processus suivants sont exécutés :

- Interconnexion BGP interne (iBGP) entre les routeurs au sein d'un AS
- Redistribution des informations BGP aux IGP qui s'exécutent dans l'AS

Quand BGP s'exécute entre des routeurs qui appartiennent à deux AS différents, on parle de BGP extérieur (eBGP). Quand BGP s'exécute entre des routeurs du même AS, on parle d'iBGP.

Messages de mise à jour BGP

Format du message



- Une mise à jour BGP permet d'annoncer une route (et une seule) à un voisin, ou bien de supprimer plusieurs routes qui ne sont plus accessibles [note : depuis quelques années, une mise à jour BGP peut concerner plusieurs prefixes]
- Chaque message contient des attributs comme : origine, chemin d'AS, Next-Hop, ...

Mises à jour BGP — Préfixes/NLRI

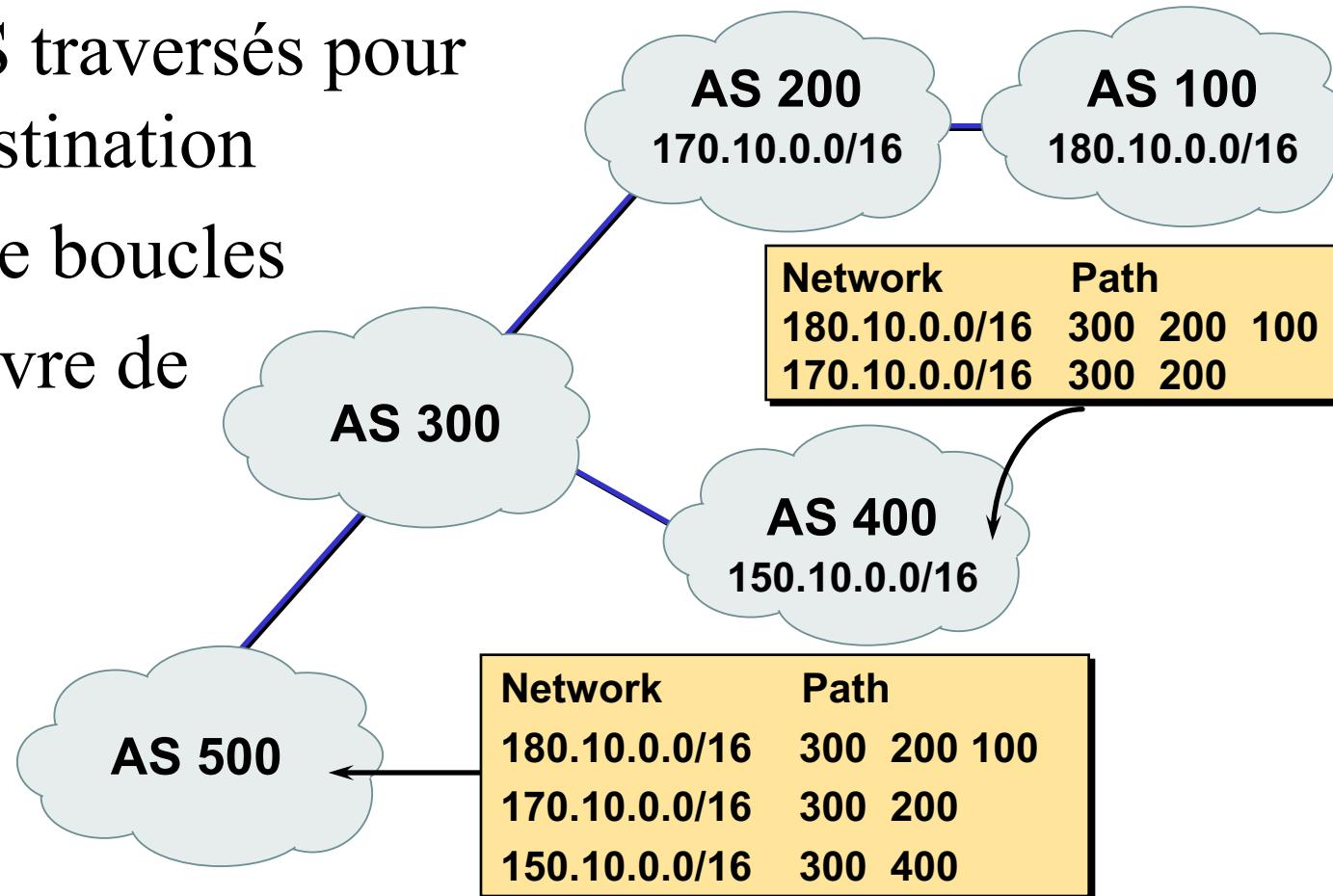
- NLRI = Network Layer Reachability Information = Préfixes
- Permet d'annoncer l'accessibilité d'une route
- Composé des informations suivantes :
 - Préfixe réseau
 - Longueur du masque

Mise à jour BGP — Attributs

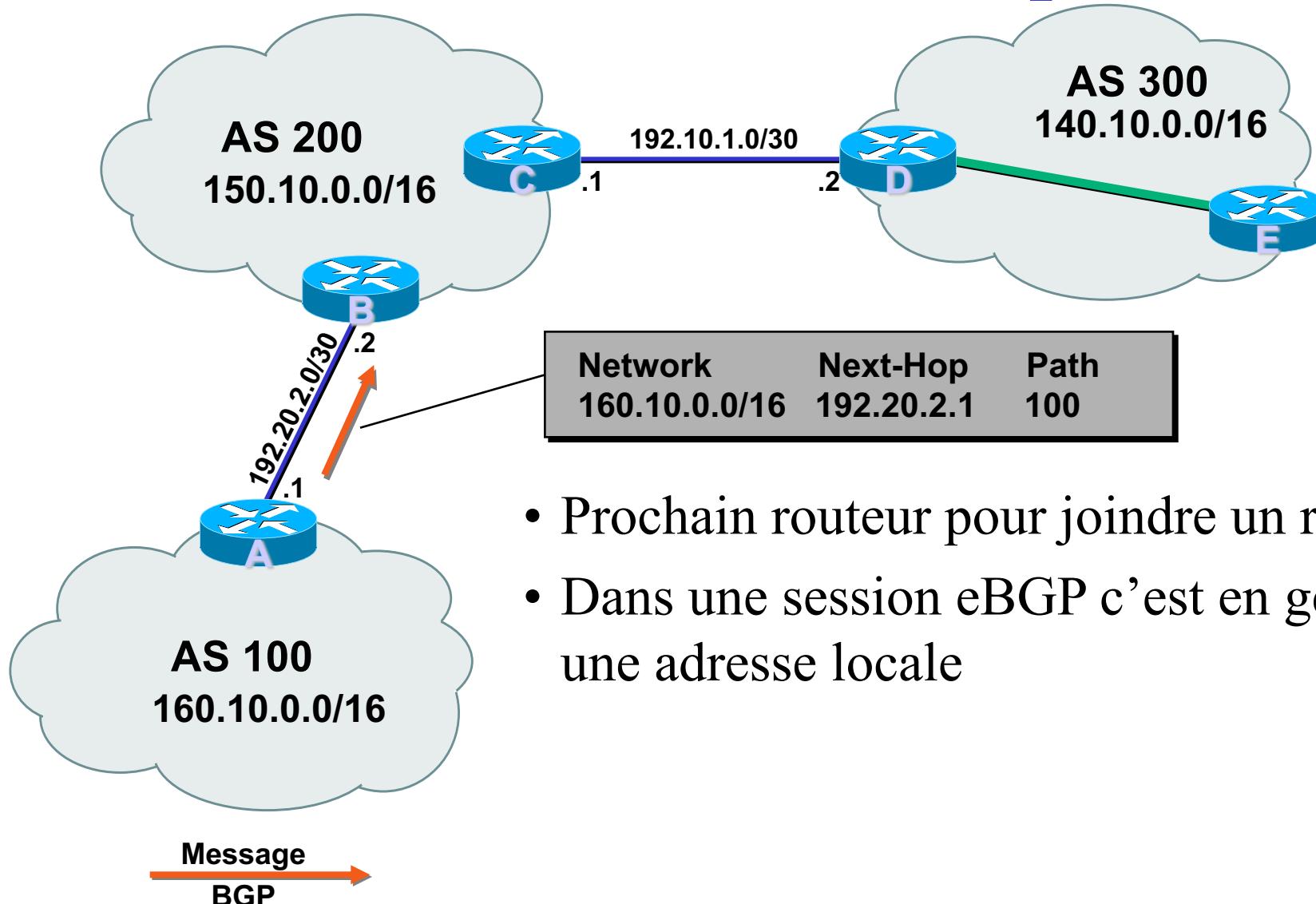
- Permet de transporter des informations liées au préfixe
 - Chemin d'AS
 - Adresse IP du “next-hop”
 - Local preference (préférence locale)
 - Multi-Exit Discriminator (MED)
 - Community (communauté)
 - Origin (origine de la route)
 - Aggregator (IP d'origine si aggrégation)

Attribut “chemin d’AS”

- Liste les AS traversés pour arriver à destination
- Détection de boucles
- Mise en œuvre de politiques

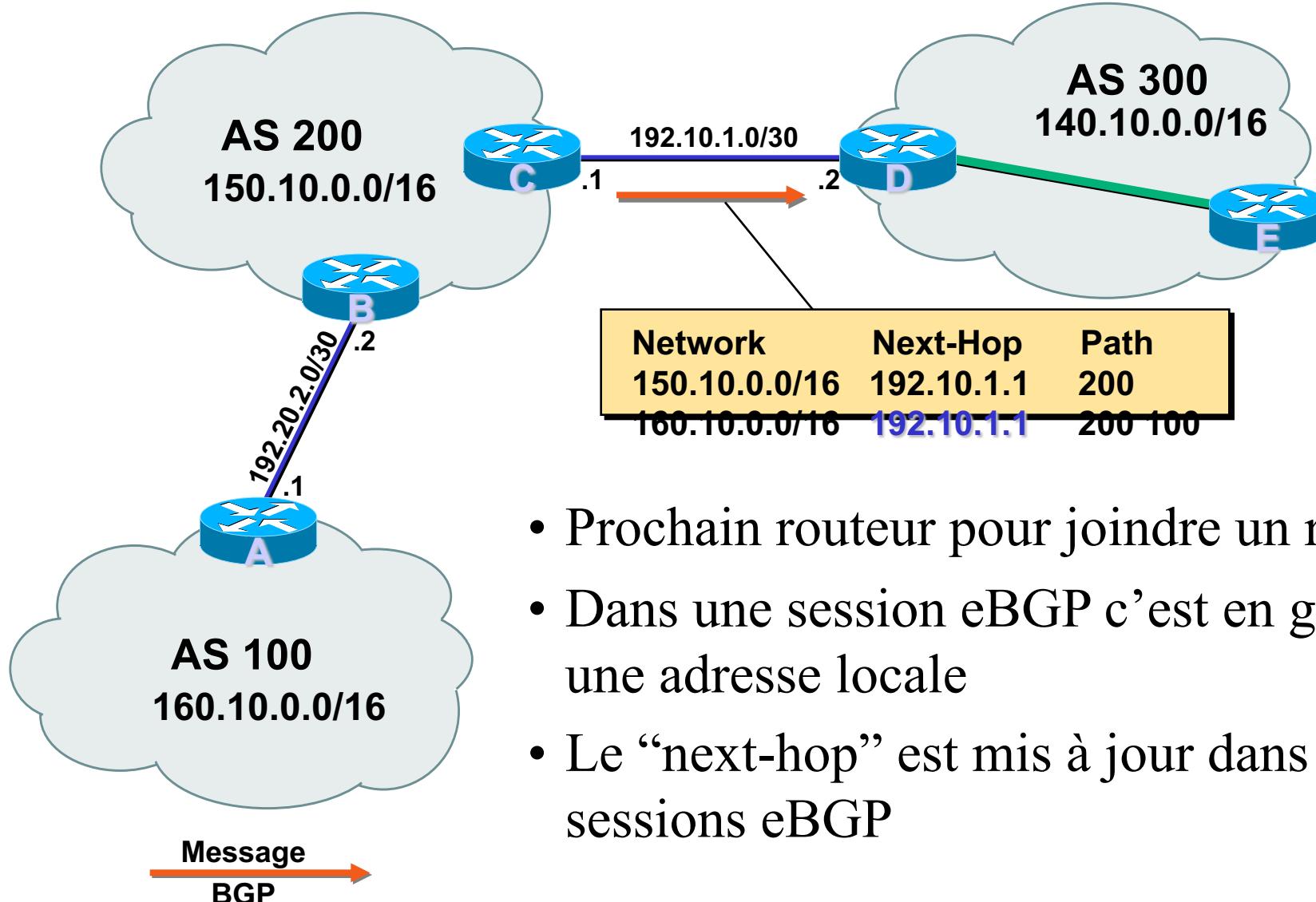


Attribut “Next-Hop”



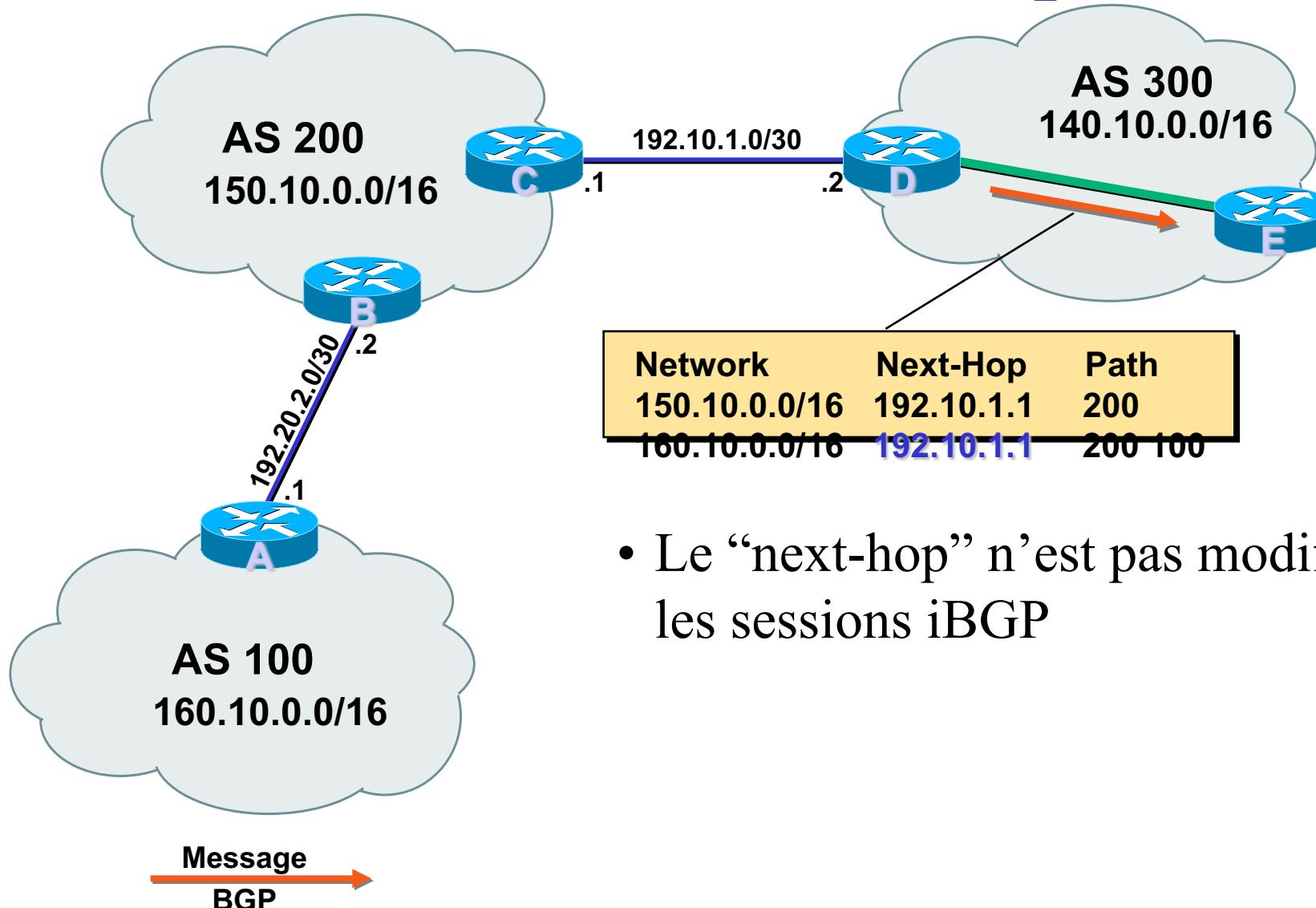
- Prochain routeur pour joindre un réseau
- Dans une session eBGP c'est en général une adresse locale

Attribut “Next-Hop”



- Prochain routeur pour joindre un réseau
- Dans une session eBGP c'est en général une adresse locale
- Le “next-hop” est mis à jour dans les sessions eBGP

Attribut “Next-Hop”



Attribut “Next-Hop” (suite)

- Les adresses des “next-hops” doivent circuler dans l’IGP
- Recherche récursive des routes
- Permet de concevoir la topologie BGP indépendamment de la topologie physique du réseau
- En interne les bonnes décisions de routage sont faites par l’IGP

Mises à jour BGP — Suppression de routes

- Permet de retirer un réseau de la liste des réseaux accessibles
- Chaque route supprimée est composée de :
 - son Préfixe
 - la longueur du masque

Mises à jour BGP - Suppression de routes

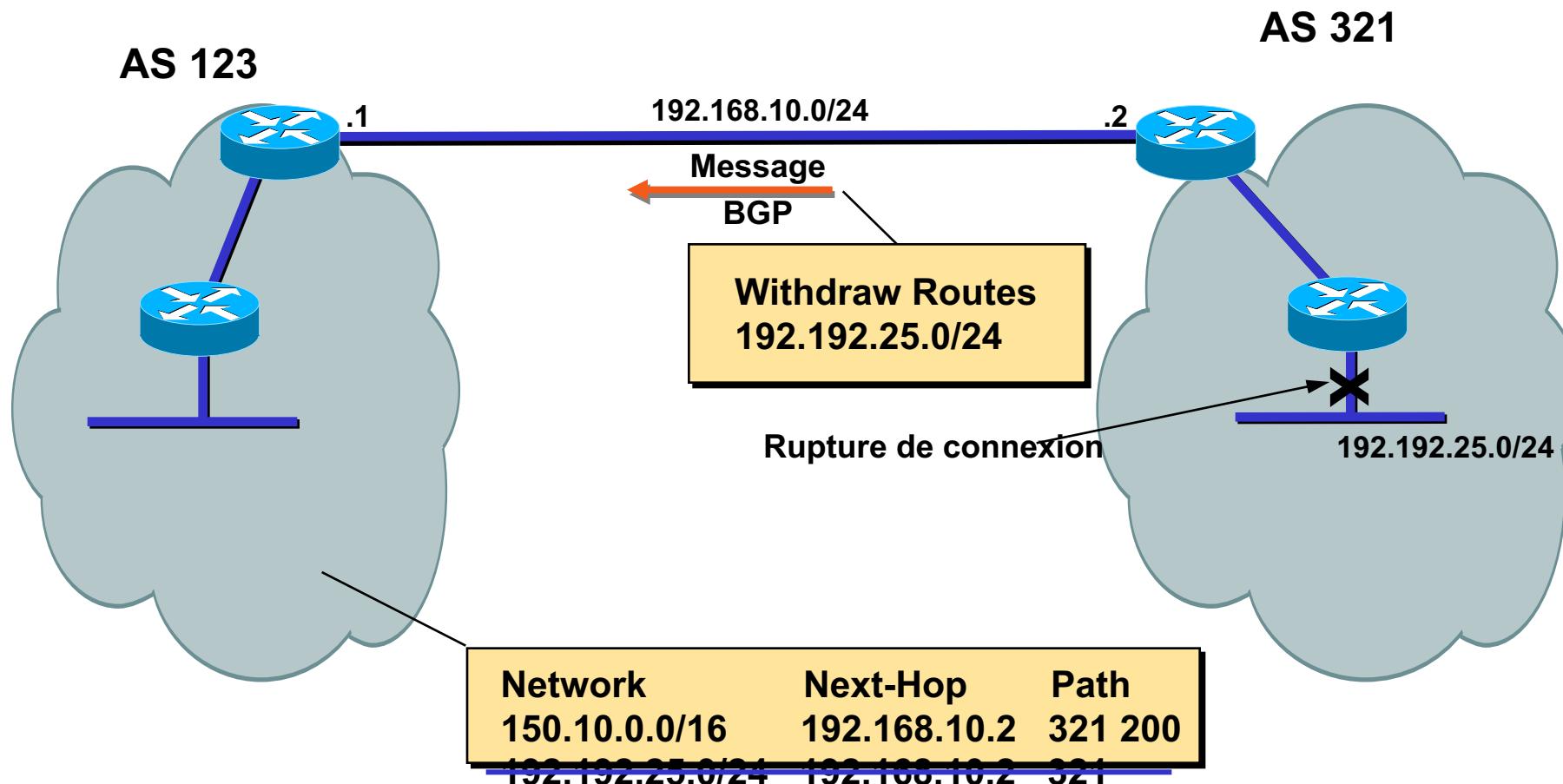


Table du routeur BGP

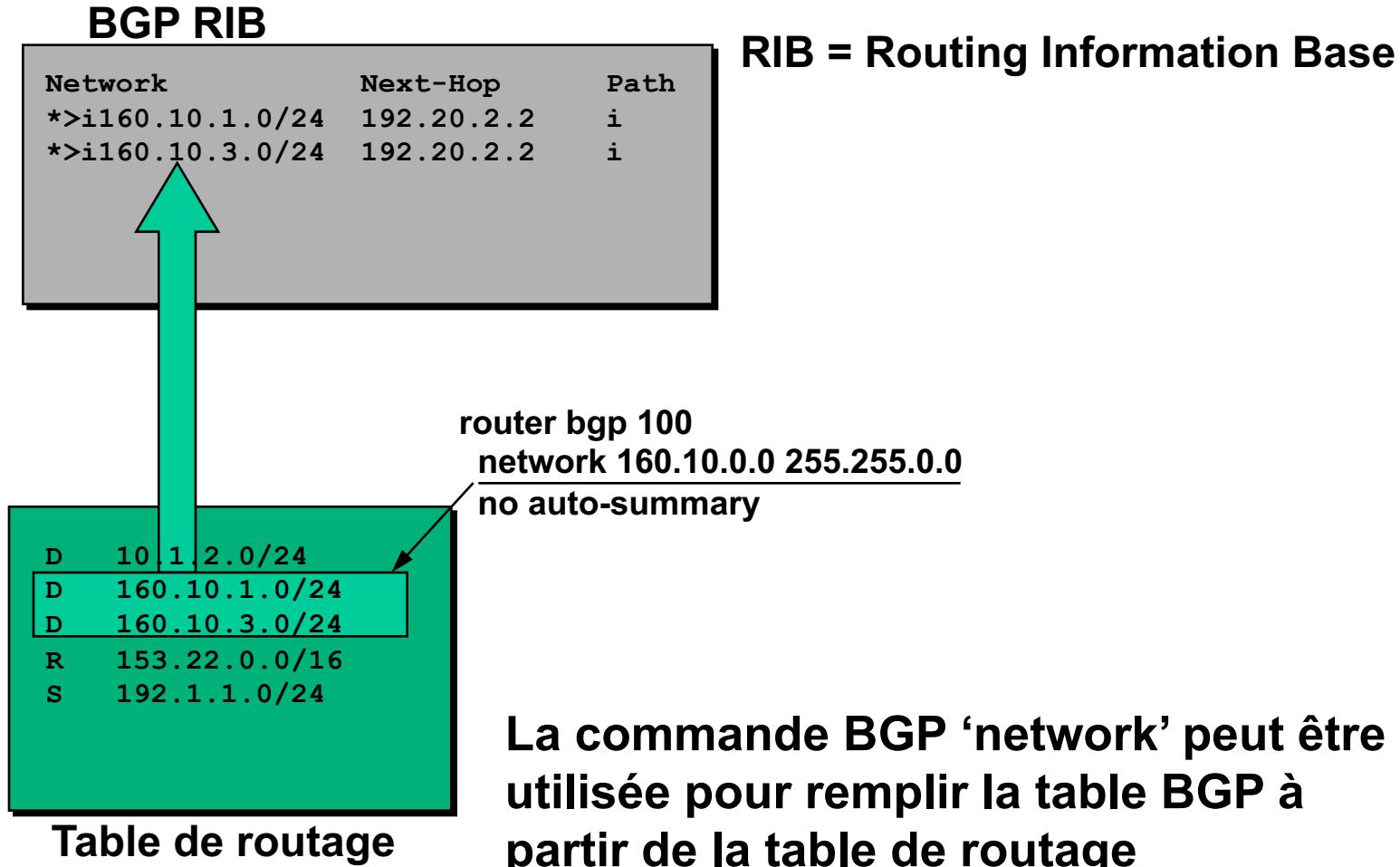


Table du routeur BGP

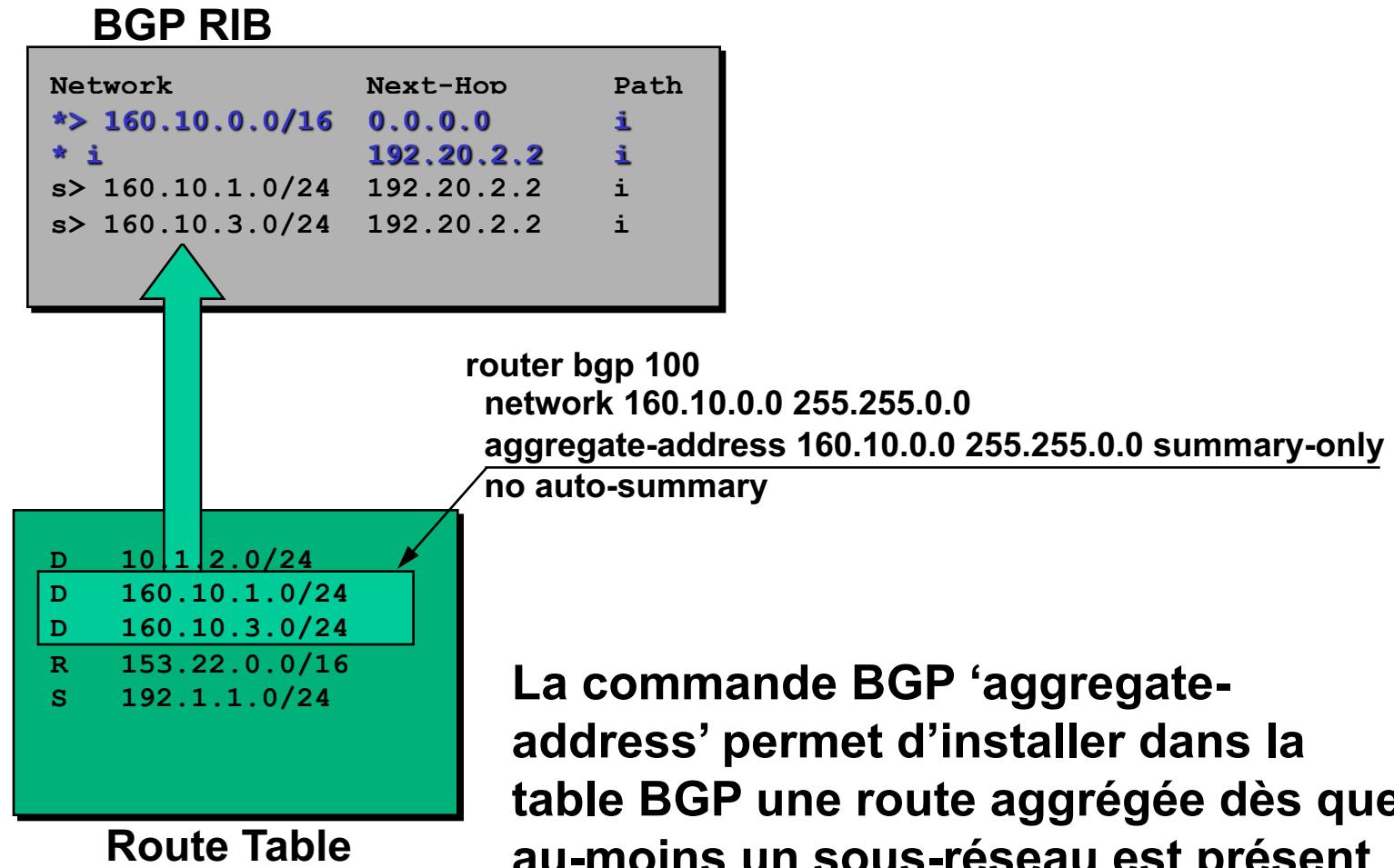


Table du routeur BGP

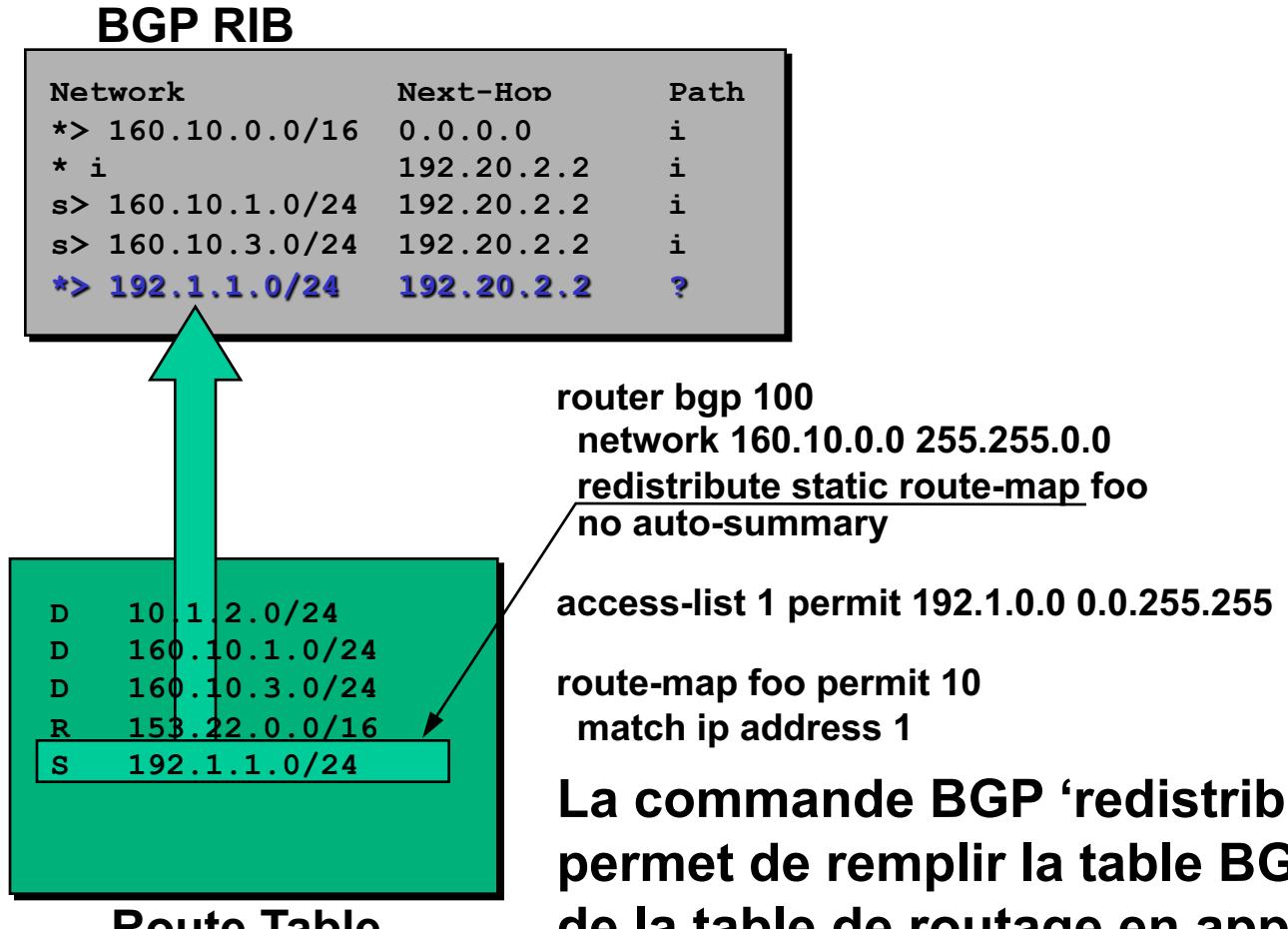
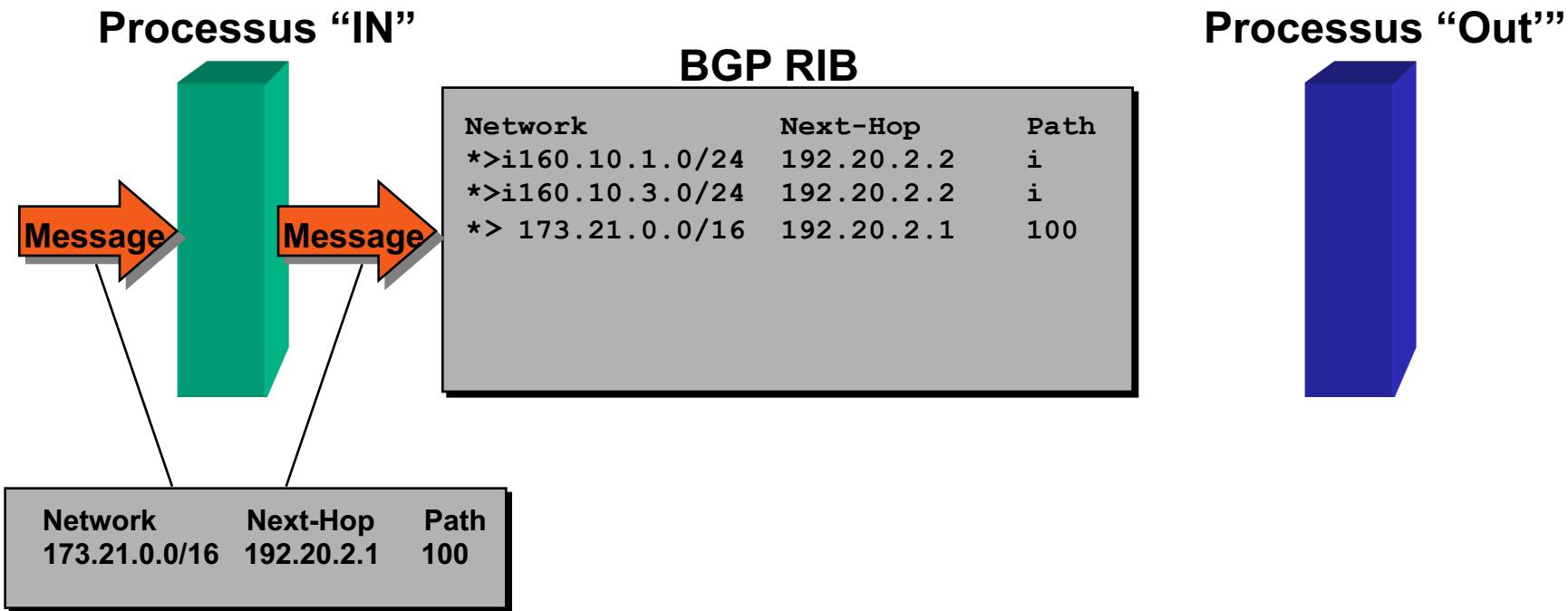
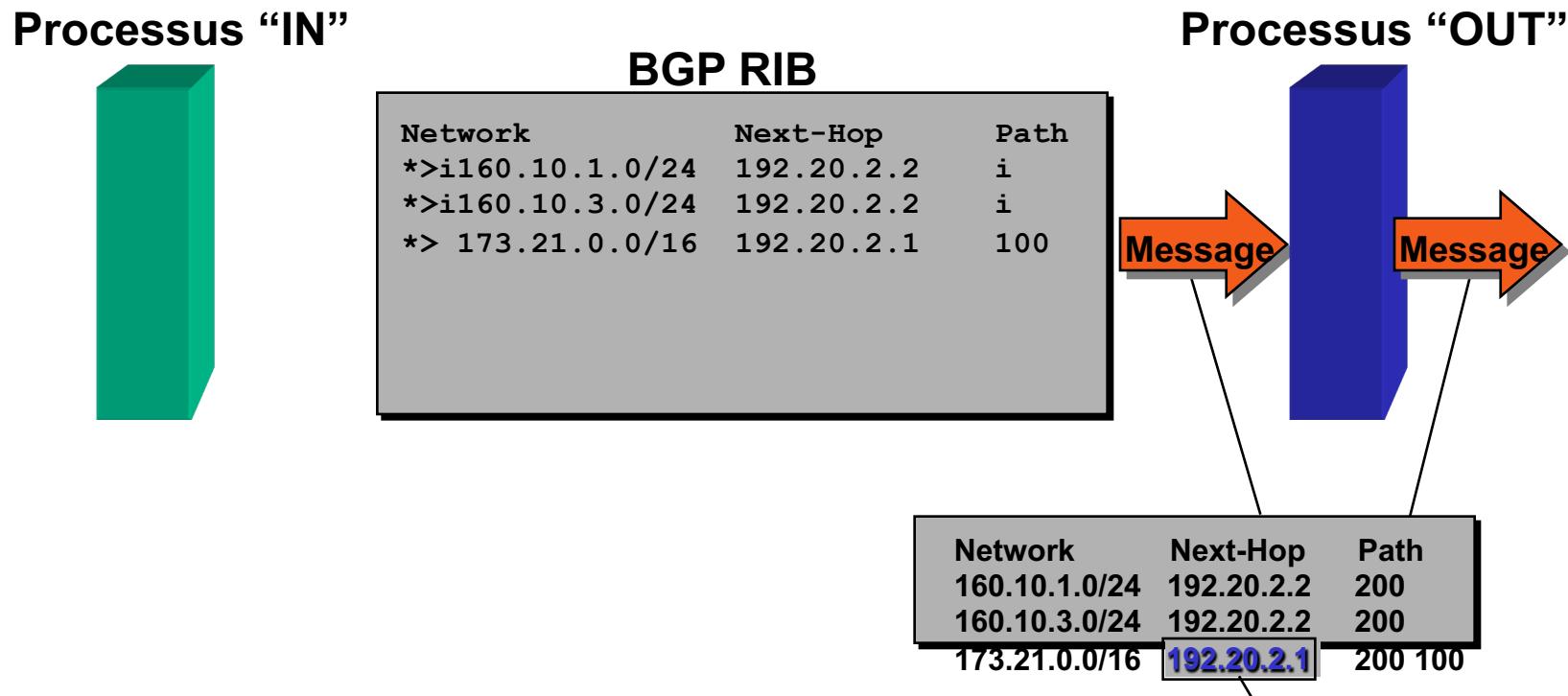


Table du routeur BGP



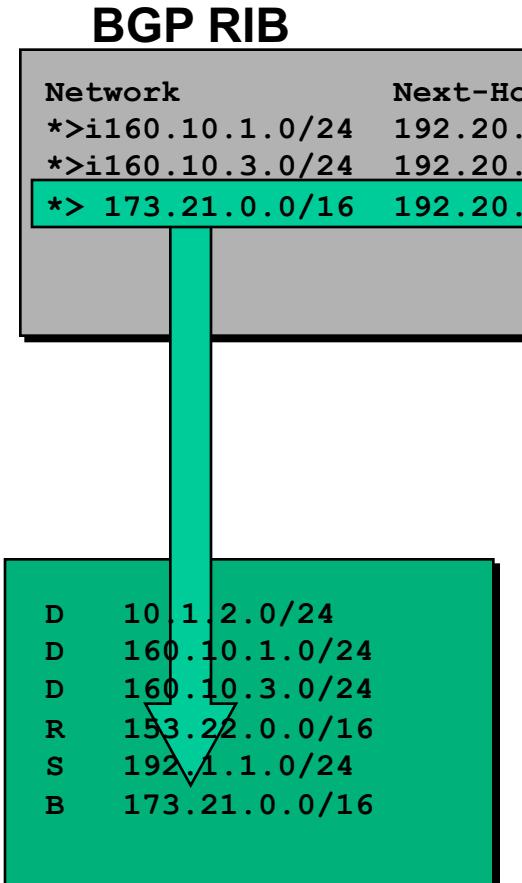
- Le processus BGP “in” (entrée)
 - reçoit les messages des voisins
 - place le ou les chemins sélectionnés dans la table BGP
 - le meilleur chemin (best path) est indiqué avec le signe “>”

Table du routeur BGP



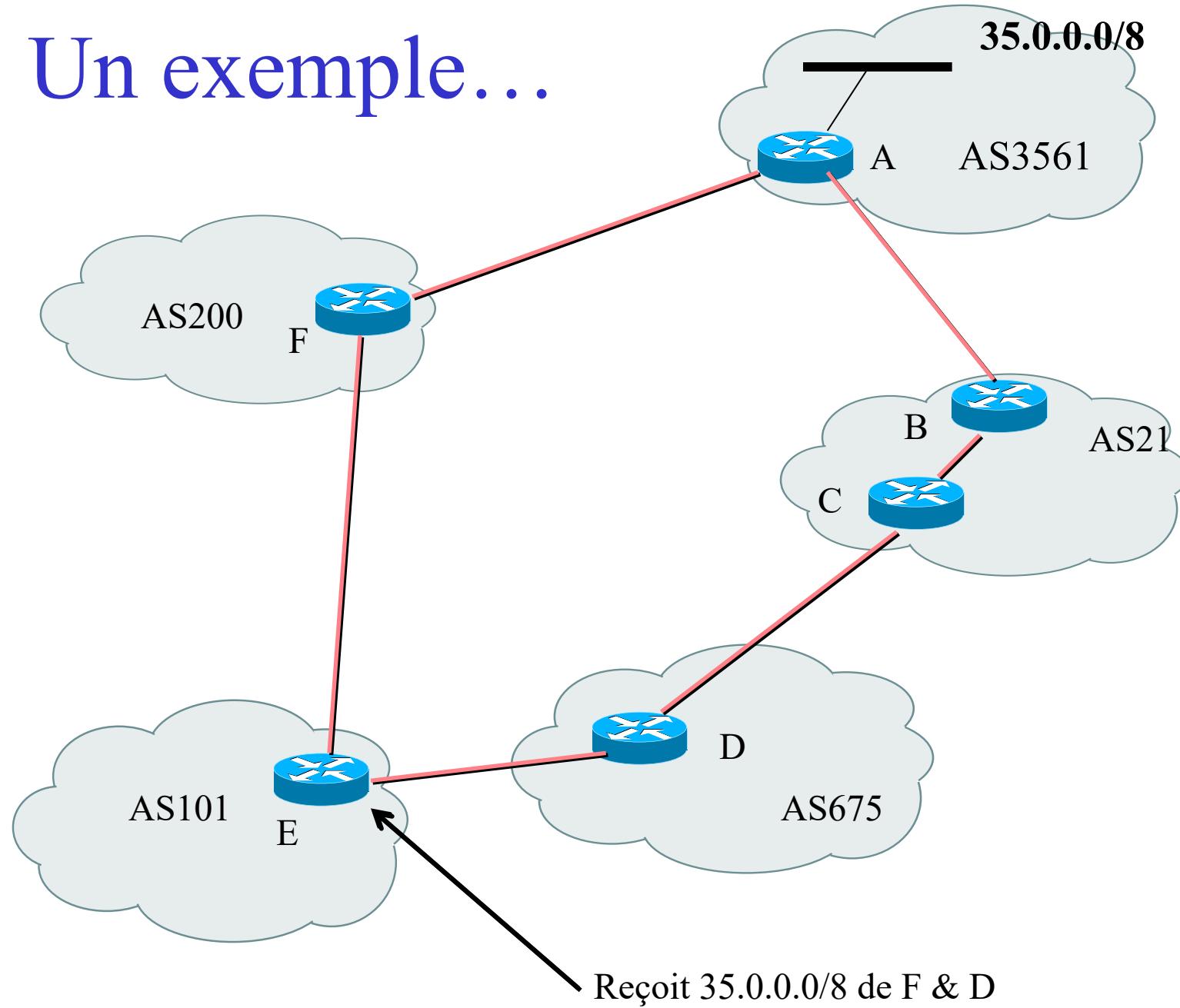
- Le processus BGP “out” (sortie)
 - message construit à partir des informations de la table BGP
 - modification du message selon configuration
 - envoi du message aux voisins

Table du routeur BGP



- **Le meilleur chemin est installé dans la table de routage du routeur si :**
 - Le préfixe et sa taille sont uniques
 - la valeur “distance” du protocole est la plus faible

Un exemple...



Commandes BGP de base

Configuration

```
router bgp <AS-number>
neighbor <ip address> remote-as <as-number>
no auto-summary
```

Consultation d'informations

```
show ip bgp summary
show ip bgp neighbors
```

Résumé BGP neighbour status

```
Router6>sh ip bgp sum
BGP router identifier 10.0.15.246, local AS number 10
BGP table version is 16, main routing table version 16
7 network entries using 819 bytes of memory
14 path entries using 728 bytes of memory
2/1 BGP path/bestpath attribute entries using 248 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1795 total bytes of memory
BGP activity 7/0 prefixes, 14/0 paths, scan interval 60 secs
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
10.0.15.241	4	10	9	8	16	0	0	00:04:47	2
10.0.15.242	4	10	6	5	16	0	0	00:01:43	2
10.0.15.243	4	10	9	8	16	0	0	00:04:49	2
...									

Version BGP Mises à jour envoyées et reçues Mises à jour en attente

Résumé Table BGP

```
Router6>sh ip bgp
BGP table version is 30, local router ID is 10.0.15.246
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal,
r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i10.0.0.0/26	10.0.15.241	0	100	0	i
*>i10.0.0.64/26	10.0.15.242	0	100	0	i
*>i10.0.0.128/26	10.0.15.243	0	100	0	i
*>i10.0.0.192/26	10.0.15.244	0	100	0	i
*>i10.0.1.0/26	10.0.15.245	0	100	0	i
*> 10.0.1.64/26	0.0.0.0	0		32768	i
*>i10.0.1.128/26	10.0.15.247	0	100	0	i
*>i10.0.1.192/26	10.0.15.248	0	100	0	i
...					

Pour insérer une route...

- Commande ***network*** ou redistribution
network <*ipaddress*> **mask** <*netmask*>
redistribute <*protocol name*>
- Il faut que la route soit présente dans la table de routage du routeur pour qu'elle soit insérée dans la table BGP

Utilisation de “redistribute static”

- Exemple de configuration

```
router bgp 109
    redistribute static
        ip route 198.10.4.0 255.255.254.0 serial0
```

- La route statique doit exister avant que la redistribution ne fonctionne
- L'origine de la route sera “*incomplete*”, mais il est possible de le changer avec une “route-map”
- A utiliser avec prudence !

Utilisation de “redistribute”

- Attention avec les redistributions
 - redistribute <protocole> signifie que toutes les routes du <protocole> seront transférées dans le protocole courant
 - cette solution doit être contrôlée (volumétrie)
 - à éviter dans la mesure du possible
 - préférer l'utilisation de “route-maps” et avec un contrôle administratif très strict

Utilisation de la commande “network”

- Exemple de configuration

```
network 198.10.4.0 mask 255.255.254.0  
ip route 198.10.0.0 255.255.254.0 serial 0
```

- La route doit être présente dans la table de routage pour qu'il y ait une annonce BGP
- Origine de la route : IGP

Choix pour les sessions iBGP

- Les sessions iBGP ne doivent pas être liées à la topologie du réseau
- L'IGP transporte les adresses de Loopback

router ospf <ID>

network <loopback-address> 0.0.0.0

- Utiliser les adresses Loopback pour les sessions iBGP

router bgp <AS1>

neighbor <x.x.x.x> remote-as <AS1>

neighbor <x.x.x.x> update-source loopback0

La technologie MPLS et son implantation

PLAN

Introduction
Principes
Applications
Déploiement

Introduction

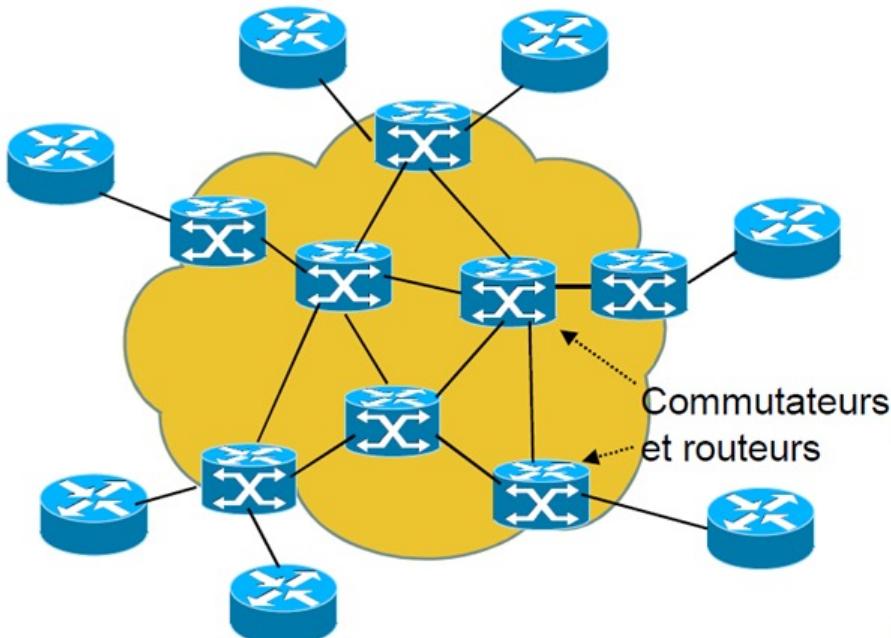
Constat

- Caractéristiques des réseaux IP actuels
 - Routage hop-by-hop suivant l'adresse destination
 - Pas de Traffic Engineering
 - Pas de QoS
 - Routage lent et consommateur en CPU
- Le **Multi Protocol Label Switching**
 - Création de chemins (LSP: label switch path)
 - Commutation de labels insérés entre la couche 2 et 3
 - Traffic Engineering et QoS

- Avantages
 - Routage à l'entrée du réseau
 - Rapidité dans le cœur de réseau
- L'intelligence se trouve aux extrémités du réseau

Les approches possibles des opérateurs

Approche d'intégration : MPLS



- **Une approche intermédiaire d'intégration a été proposée**
 - Un seul plan d'adressage IP
 - Les équipements internes de l'opérateur sont des équipements hybrides
 - Ils peuvent fonctionner comme des routeurs IP
 - Ils peuvent aussi fonctionner dans un mode proche du circuit virtuel en commutant selon un numéro de label
- **Le routage n'est plus forcément corrélé à l'adresse destinataire IP**
 - L'ingénierie de trafic est possible
- **IP et MPLS sont des technologies proches conçues pour interopérer**

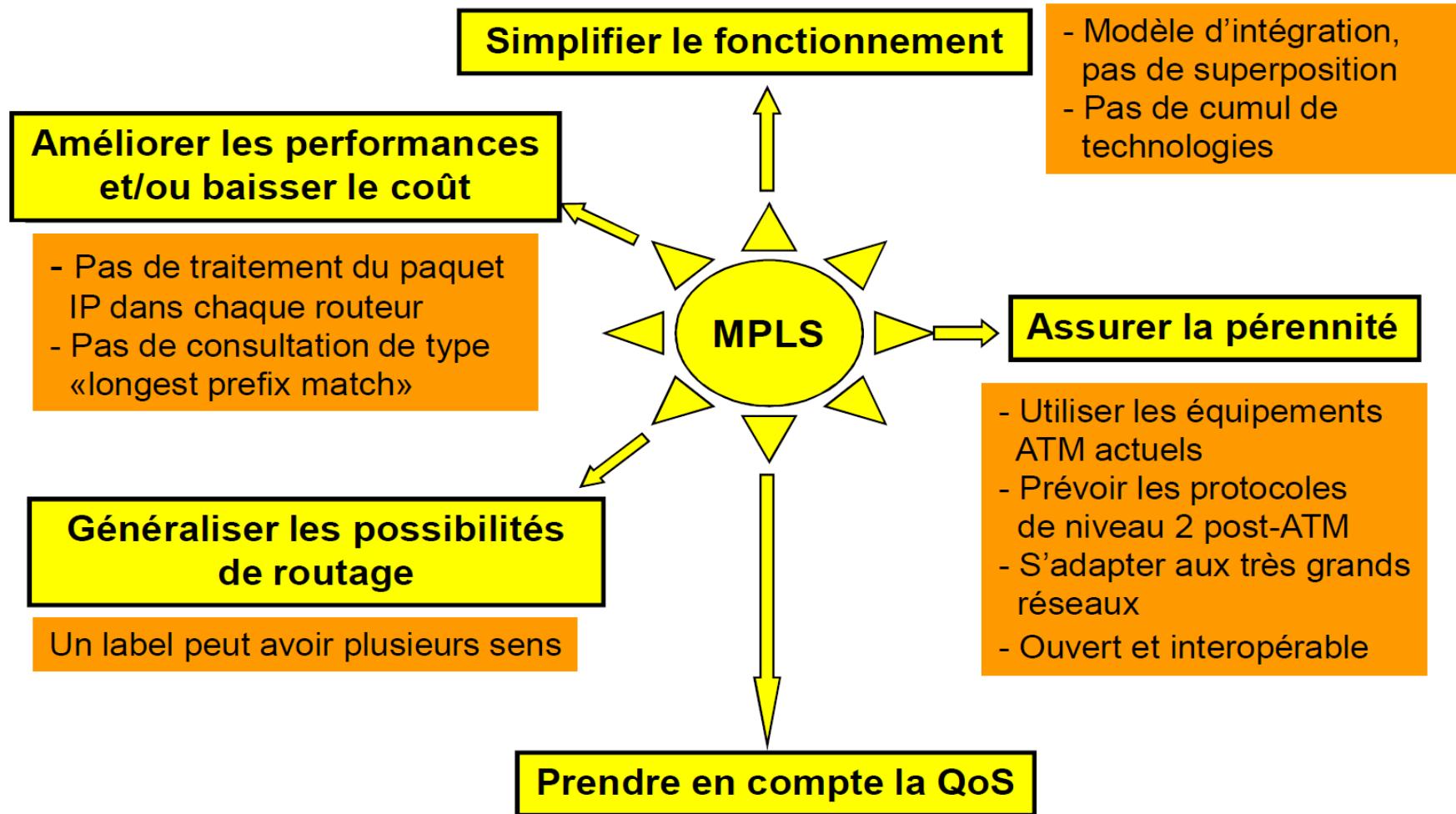
Objectifs

- Améliorer les performances
- Proposer de la Qualité de Service
- Réduire la taille des tables de routage
- Fonctionner avec n'importe quel protocole de couche 2

Pour les opérateurs :

- Nouveaux services
- Tout en s'appuyant sur l'infrastructure en place

Objectifs de MPLS



Principles

Principes de MPLS

- **Des chemins prédefinis relient les points d'extrémité du réseau**
 - Les LSP (Label Switched Path)
 - Un LSP est unidirectionnel
 - Les LSP sont établis par un protocole de signalisation en suivant la route déterminée par les protocoles de routage IP
 - Les LSP peuvent être établis à l'initiative de l'administrateur (proche des circuits virtuels permanents)
 - Ou ils sont établis automatiquement à l'initiative d'un point d'extrémité du réseau dès qu'il apprend par les protocoles de routage l'existence d'un nouveau préfixe IP
- **Les équipements MPLS s'appellent des LSR (Label Switch Router)**
- **A l'entrée du réseau, le 1er LSR (« Ingress LSR ») analyse le paquet IP**
 - Il choisit alors le LSP et insère un label devant le paquet IP
- **Les équipements suivants (les LSR du cœur de réseau) relaient le paquet en se basant seulement sur le label**
- **Le LSR de sortie (« Egress LSR) retire le label**
 - Dans certaines implémentations, c'est l'avant dernier LSR qui retire le label
- **A la sortie le paquet est routé selon le fonctionnement IP traditionnel**

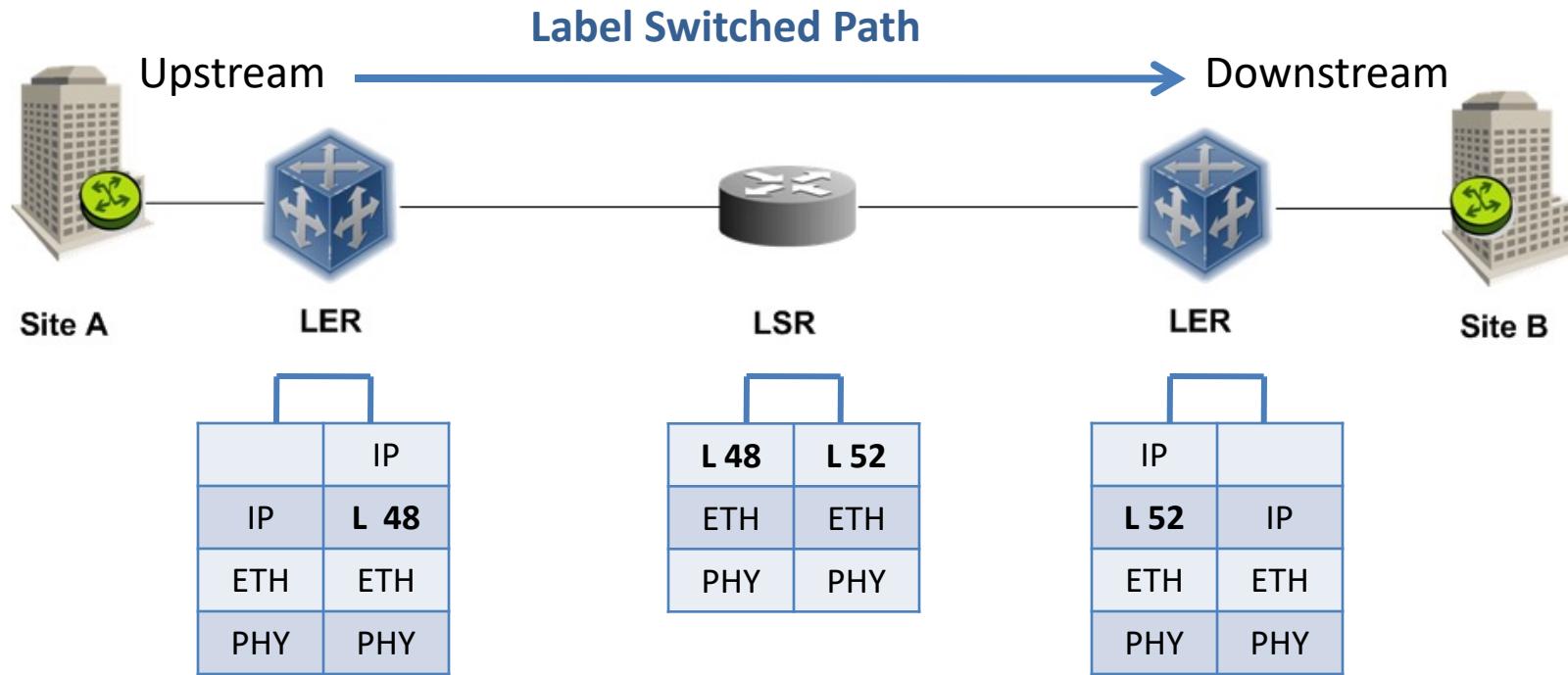
- MPLS
 - Multi Protocol Label Switching
- Associer
 - La puissance de commutation du niveau 2
 - La flexibilité du routage de niveau 3

Principes du MPLS

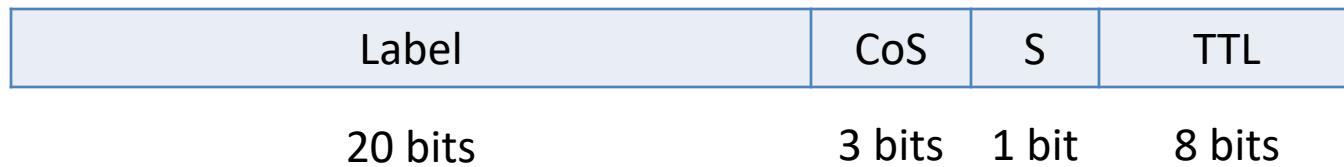
1. La commutation de label
2. Les Forwarding Equivalent Class (FEC)
3. Distribution des labels
4. Rétention des labels
5. Etablissement d'un LSP
6. LSP Tunneling

Principe du MPLS

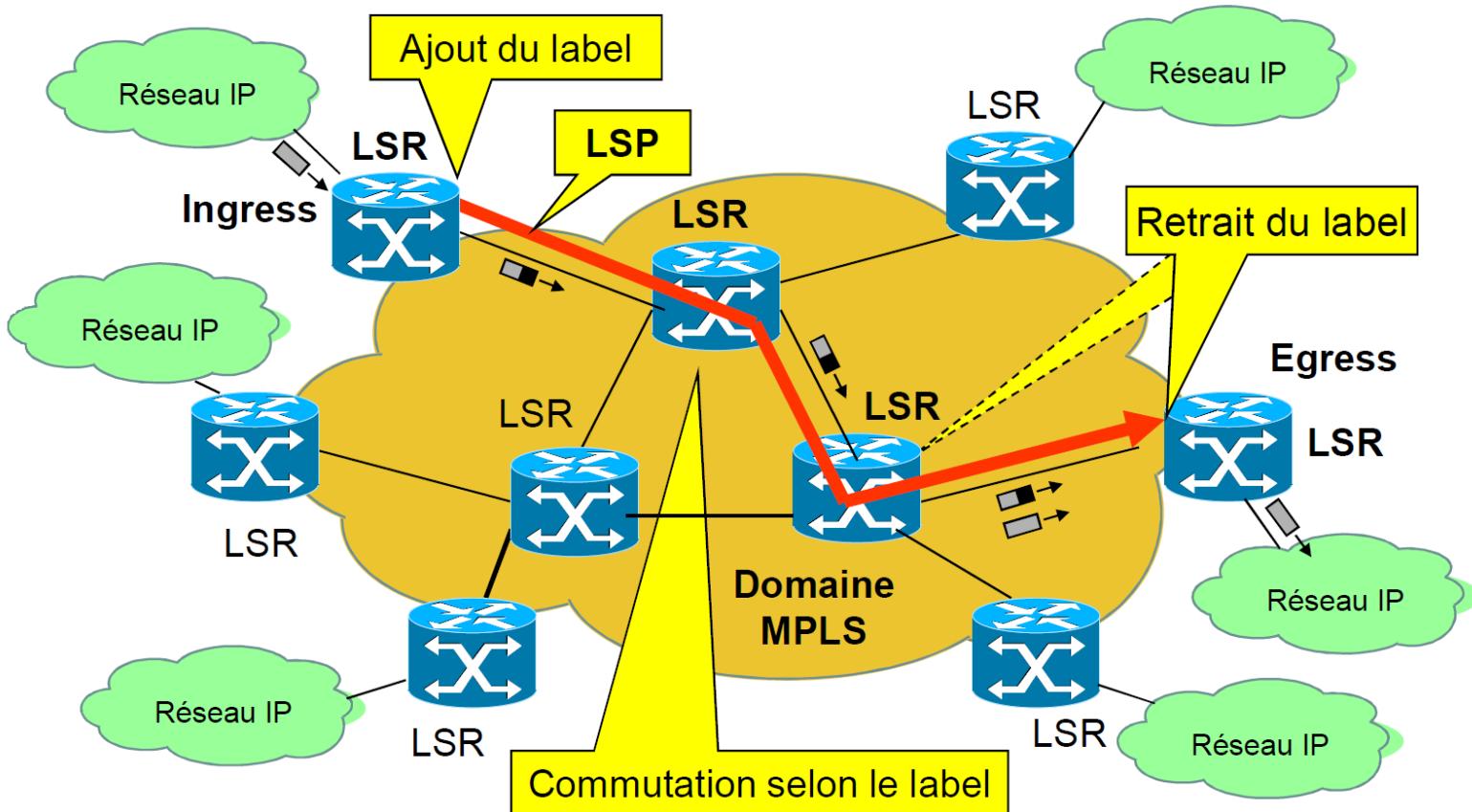
- La commutation de labels



- L'entête MPLS



Principes de MPLS



ACRONYMES

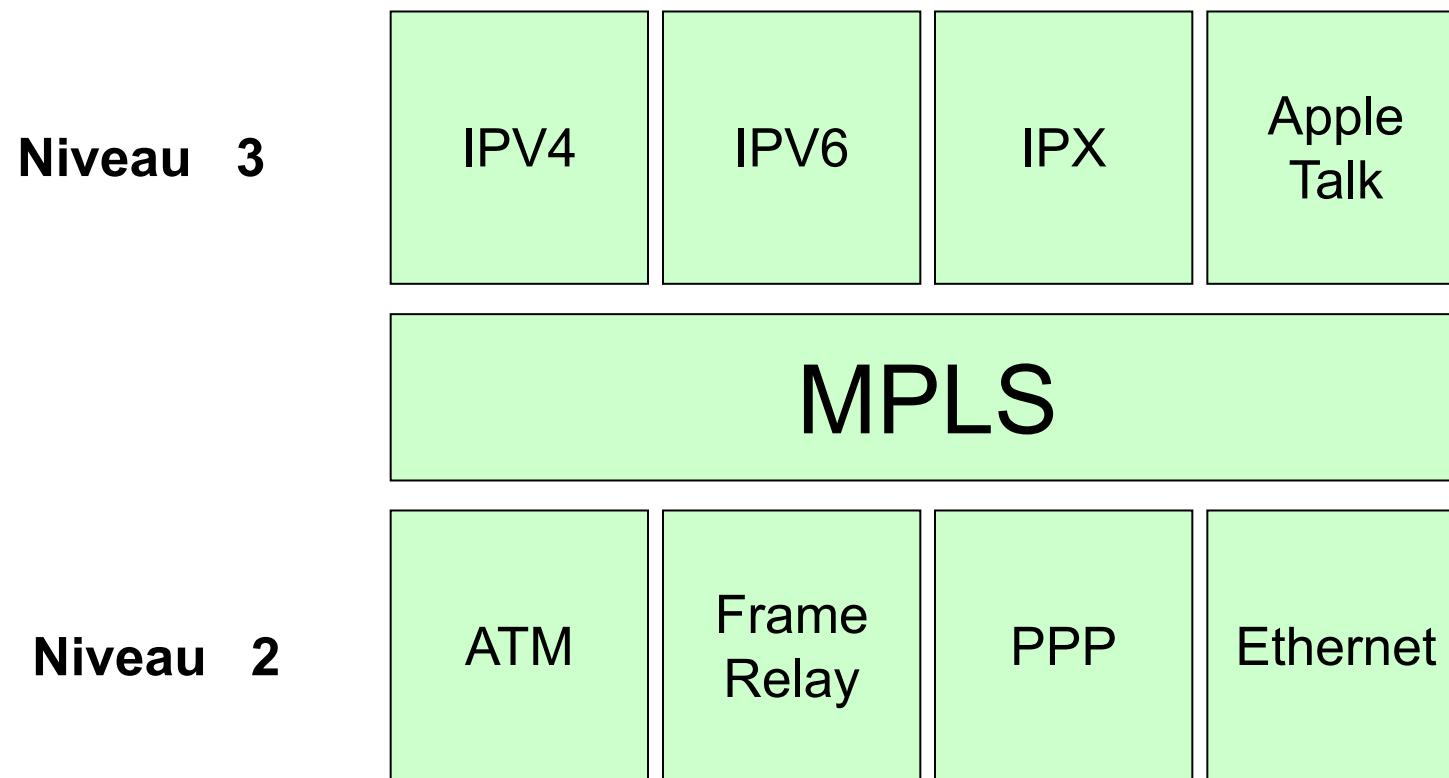
- **MPLS** Multiple Protocol Label Switching;
- **iLER** ingress Label Edge Router (routeur d'entrée)
- **eLER** egress Label Edge Router (routeur de sortie)
- **LSR** Label Switch Router
- **LIB** Label Information Base
- **FIB** Forwarding Information Base
- **LSP** Label Switch Path
- **FEC** Forward Equivalence Class, also Functional Equivalent Class
- **LDP** Label Distribution Protocol

MPLS points clés

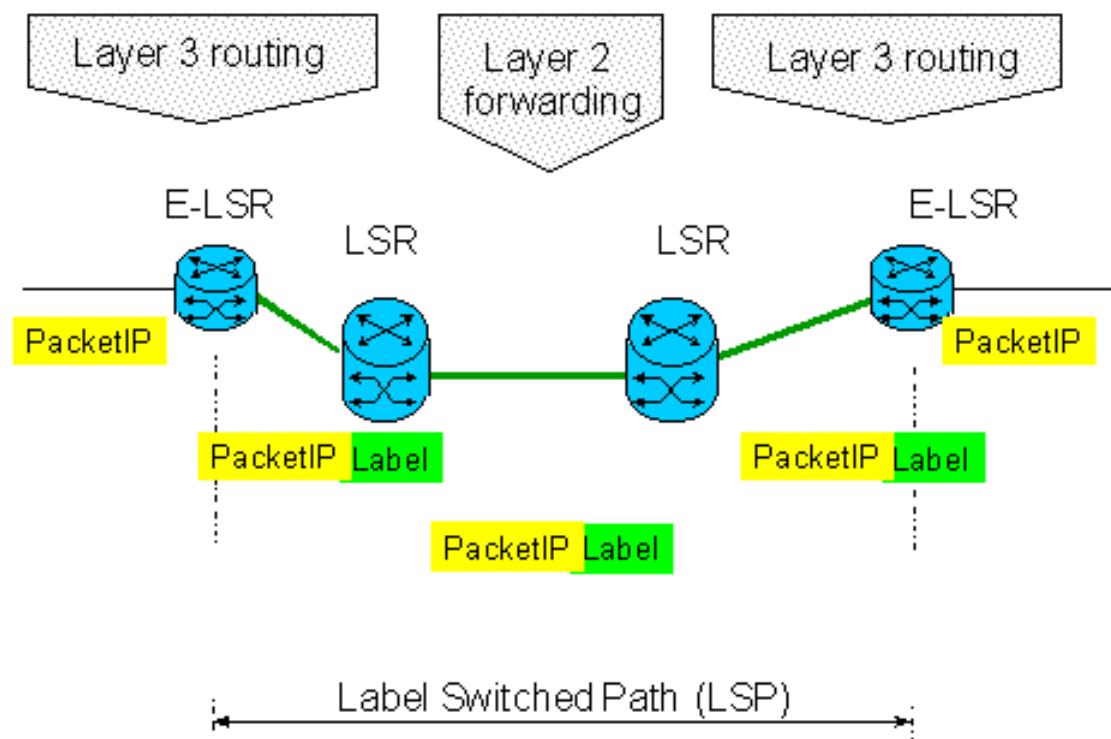
- MPLS permet un **mariage** facile d'IP avec des technologies de niveau 2 telle ATM par ajout une couche très simple sur IP.
- Les routeurs avec le logiciel adequat LDP (Label Distribution Protocol), c'est à dire supportant MPLS, vont mettre en oeuvre la gestion de label MPLS contenu dans le "Shim Header".
- Le trafic de base IP est présenté au LER (label edge router), où les labels sont mis en place, ces paquets sont envoyés aux travers des LSP (label switch path) aux LSR (label switch router) où les labels sont traités.
- En sortie du réseau (the egress to the network) le LER enlève les MPLS labels et délivre les paquets IP.
- Le traffic au travers de plusieurs réseaux, **peut-être "tunneled" en empilant les labels.**

Normalisation

Un protocole « multi-potocole »



Normalisation



MPLS

- Multi Protocol
 - Simplification du cœur de réseau
 - Déploiement de nouveaux services et protocoles
 - Indépendant du protocole utilisé pour la couche inférieure

LABEL

- Label switching
 - Encapsulation de datagrammes dans un label MPLS
- Signification d' un label
 - Un chemin
 - Une source
 - Une destination
 - Une application
 - Une QoS

LABEL

- Un **label** a une signification locale entre 2 LSR adjacents et mappe le flux de trafic entre le LSR amont et le LSR aval.
- A chaque bond le long du LSP, un label est utilisé pour chercher les informations de routage (next hop, lien de sortie, encapsulation, queueing et scheduling) et les actions à réaliser sur le label : insérer, changer ou retirer.
- Labels dans les différentes technologies ATM, Frame Relay, PPP, Ethernet et HDLC.
 - Pour les réseaux Ethernet, un champ appelé shim a été introduit entre la couche 2 et la couche 3.
 - Sur 32 bits, il a une signification d'identificateur local d'une FEC. **20 bits contiennent le label**, un champ de **3 bits appelé Classe of Service (CoS) sert actuellement pour la QoS**, **un bit S pour indiquer s'il y a empilement de labels** et un dernier champ, **le TTL sur 8 bits** (même signification que pour IP).
 - L'empilement des labels permet en particulier d'associer plusieurs contrats de service à un flux au cours de sa traversée du réseau MPLS.

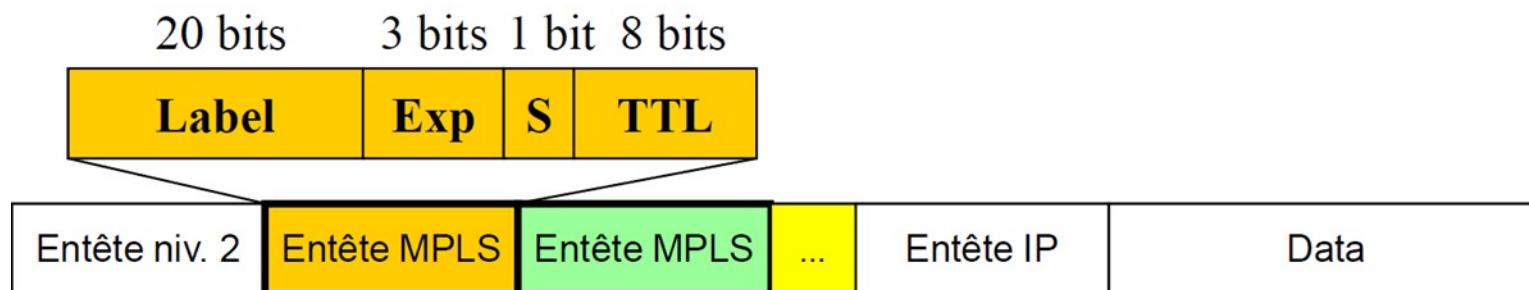
LABEL

- Le label MPLS (header MPLS)



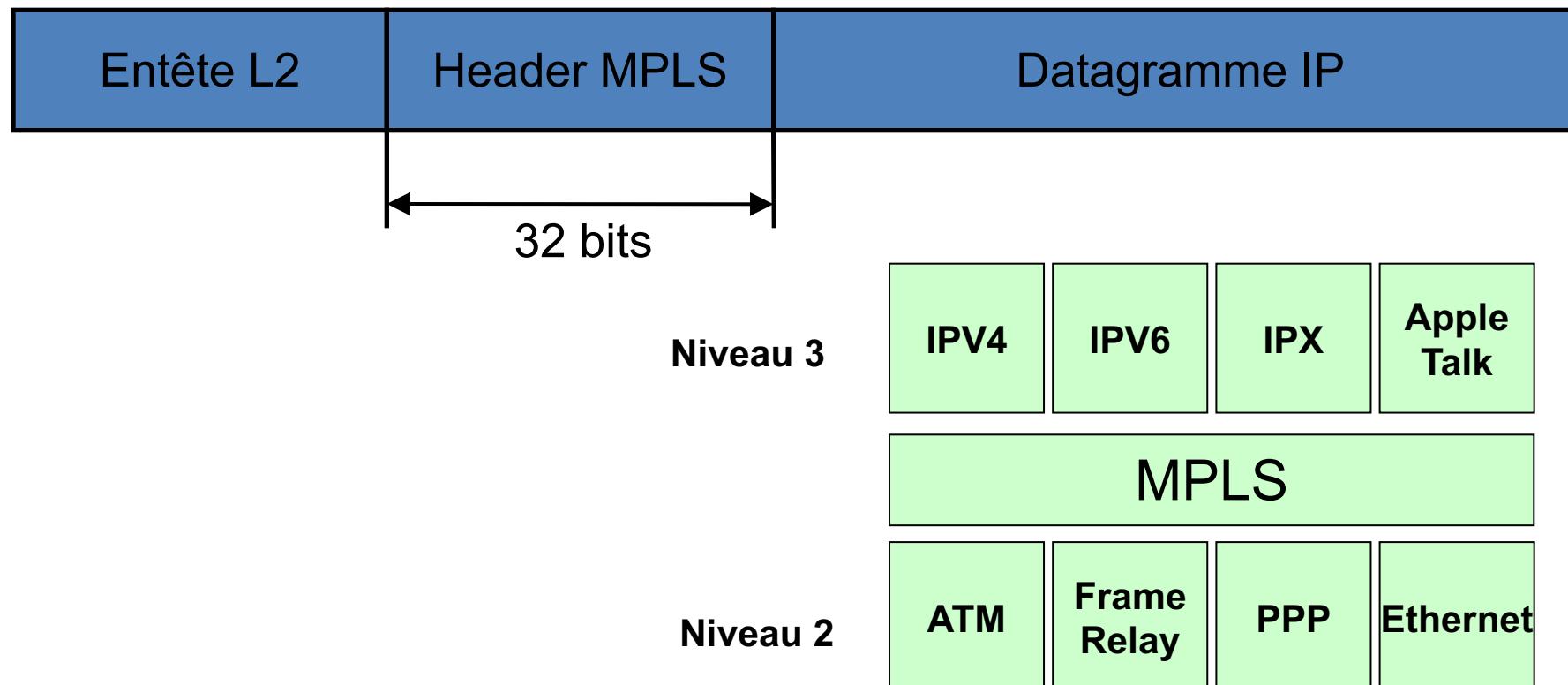
Labels MPLS

- **L'en-tête MPLS est inséré entre le niveau 2 et l'en-tête IP**
 - Label : numéro sur 20 bits (Valeurs 0 à 16 réservées)
 - S sert à gérer des labels hiérarchisés (Stack)
 - Marque le dernier label avant l'entête IP
 - Exp peut être utilisé pour traiter la QoS : files d'attente et rejet
 - Fonctionnement conforme à DiffServ
 - TTL a le même rôle que dans IP (détection de boucles)
- **Opérations sur les labels**
 - Swap (dans les LSRs), push (dans ingress LSR), pop (dans egress LSR)
- **Le label peut éventuellement être implicite**
 - Par exemple une longueur d'onde

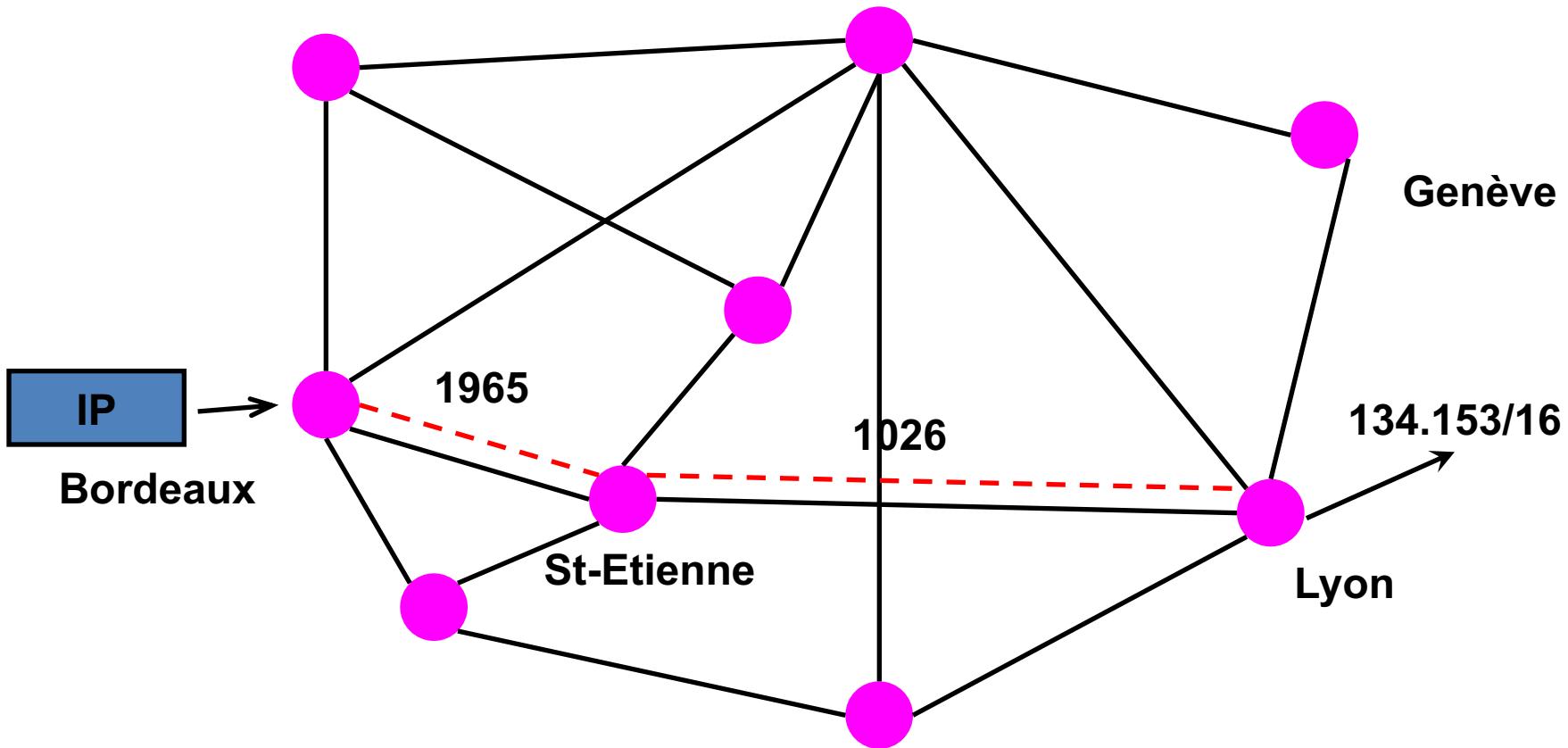


LABEL

- Encapsulation MPLS

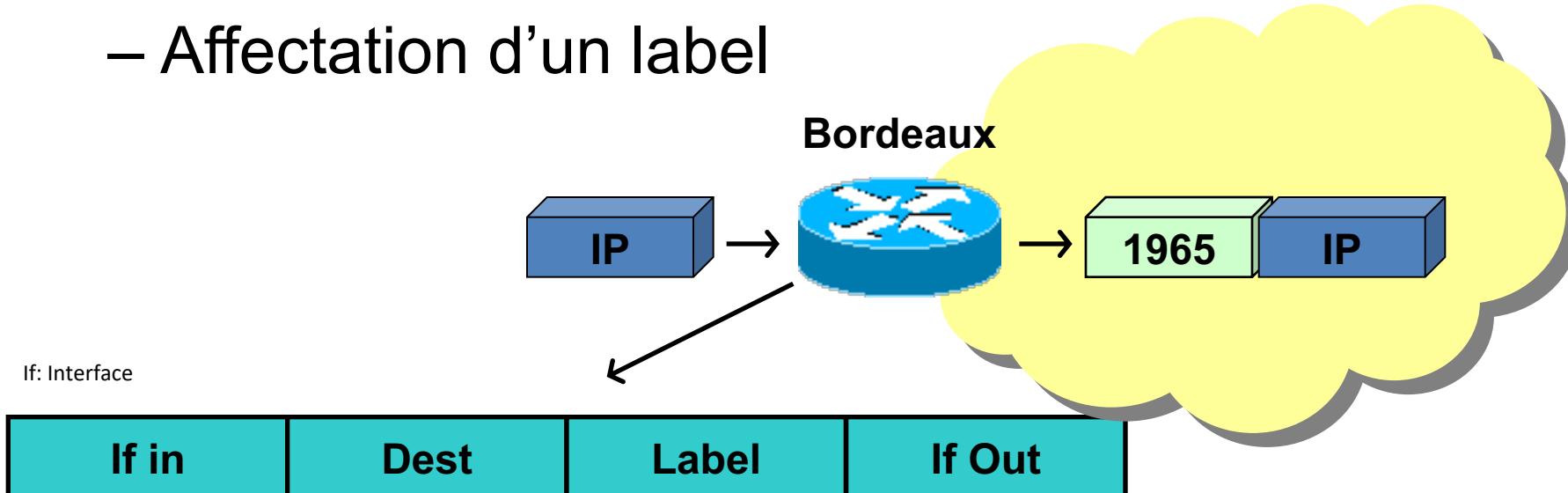


Principes

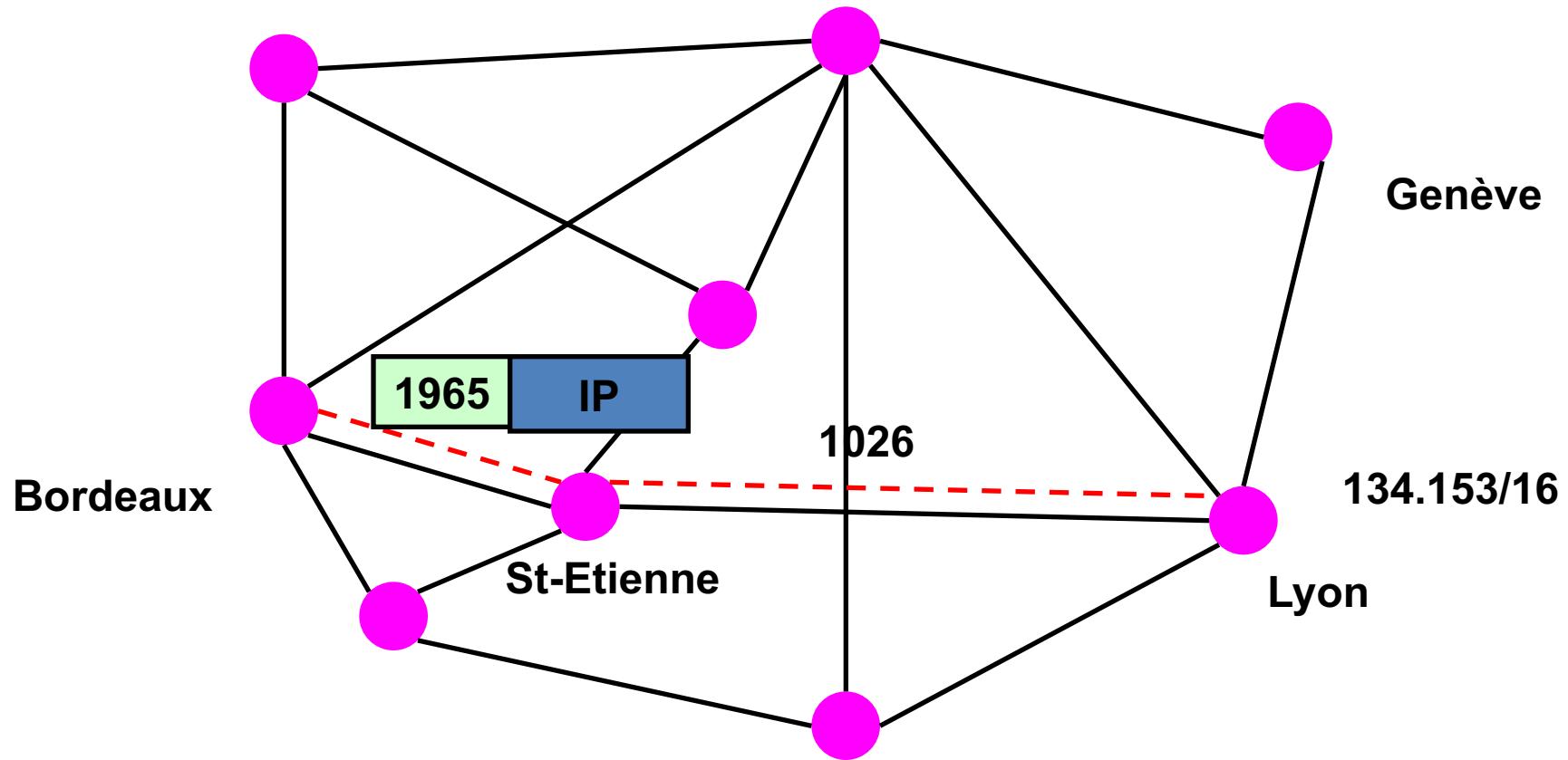


Principes

- Etape 1 : Push
 - Entré dans le réseau
 - Affectation d'un label

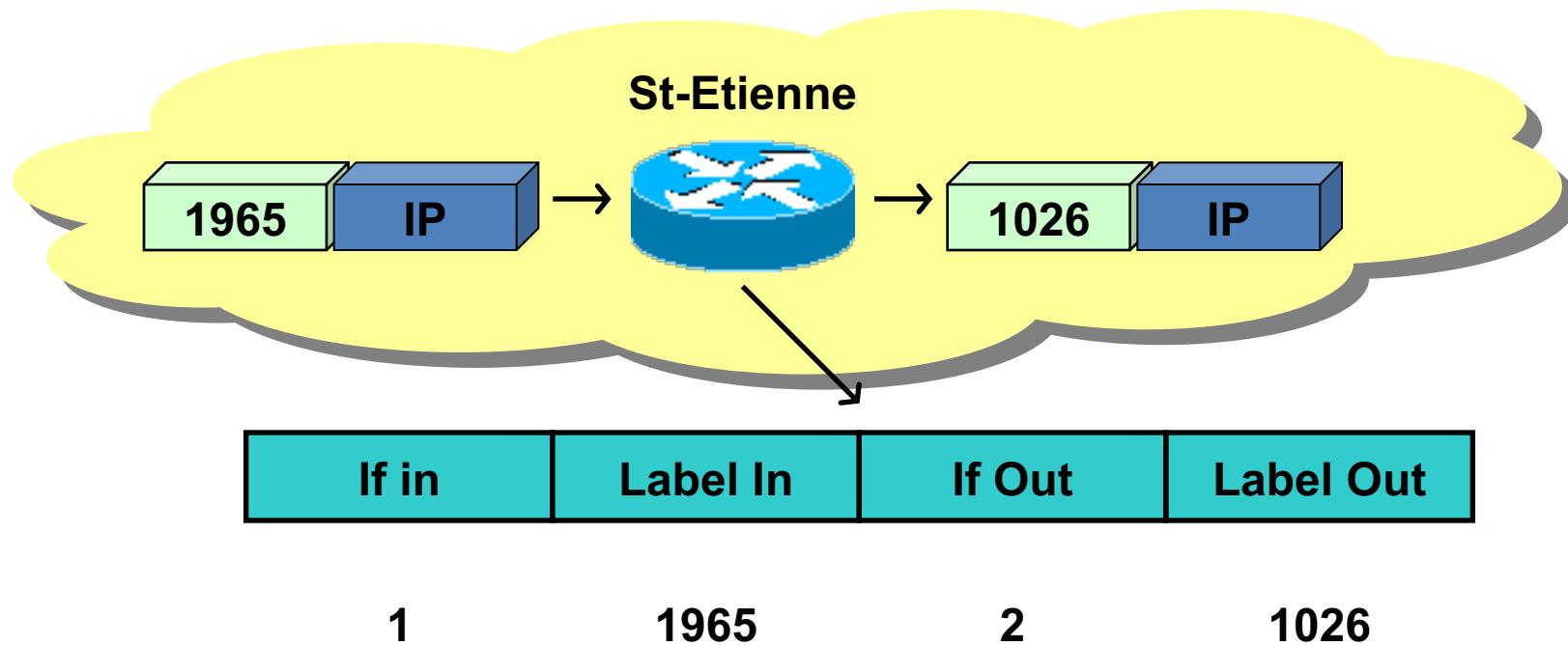


Principes

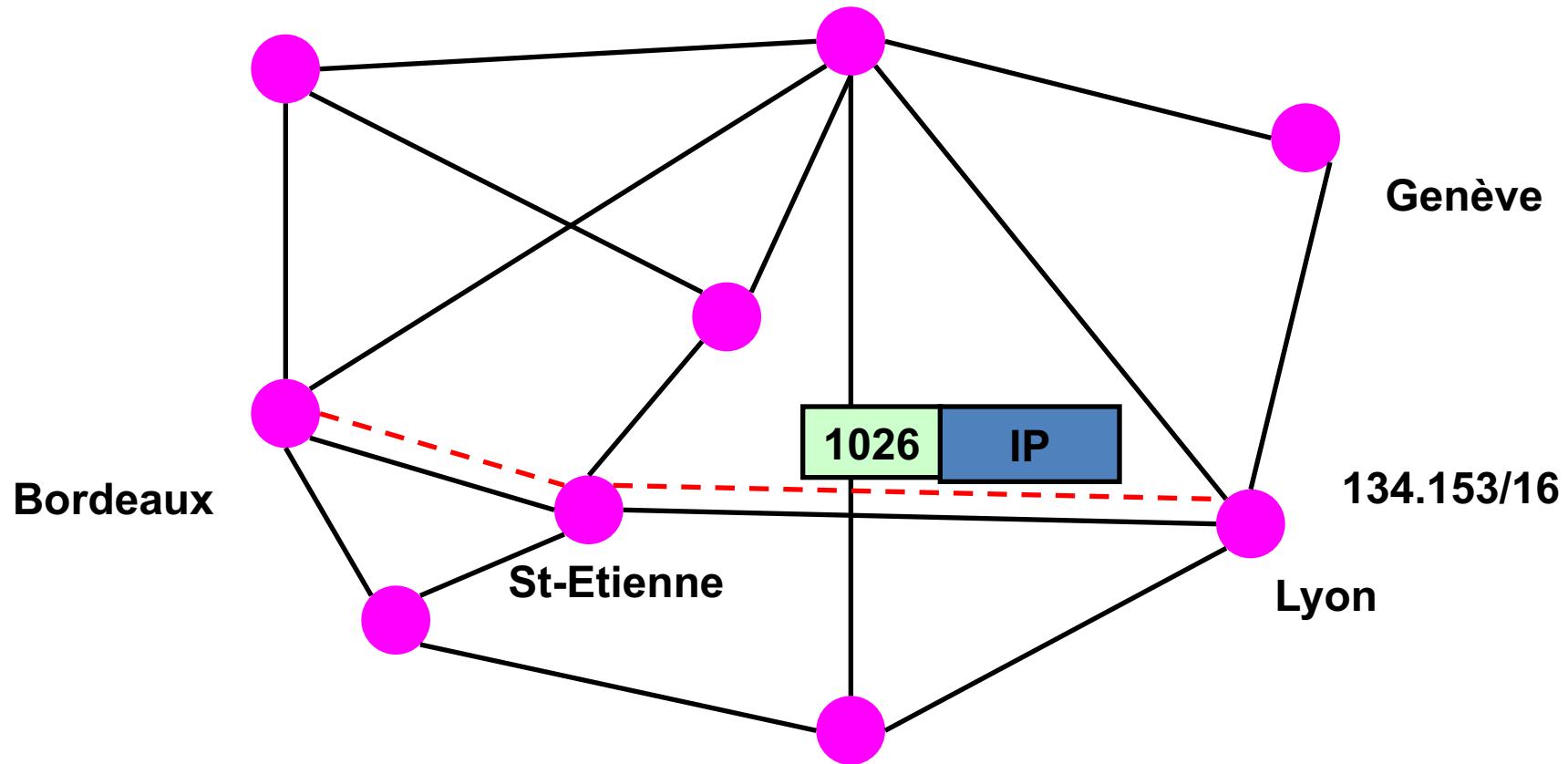


Principes

- Etape 2 à n-1 : Swap
 - Echange du label suivant la table de commutation du LSR

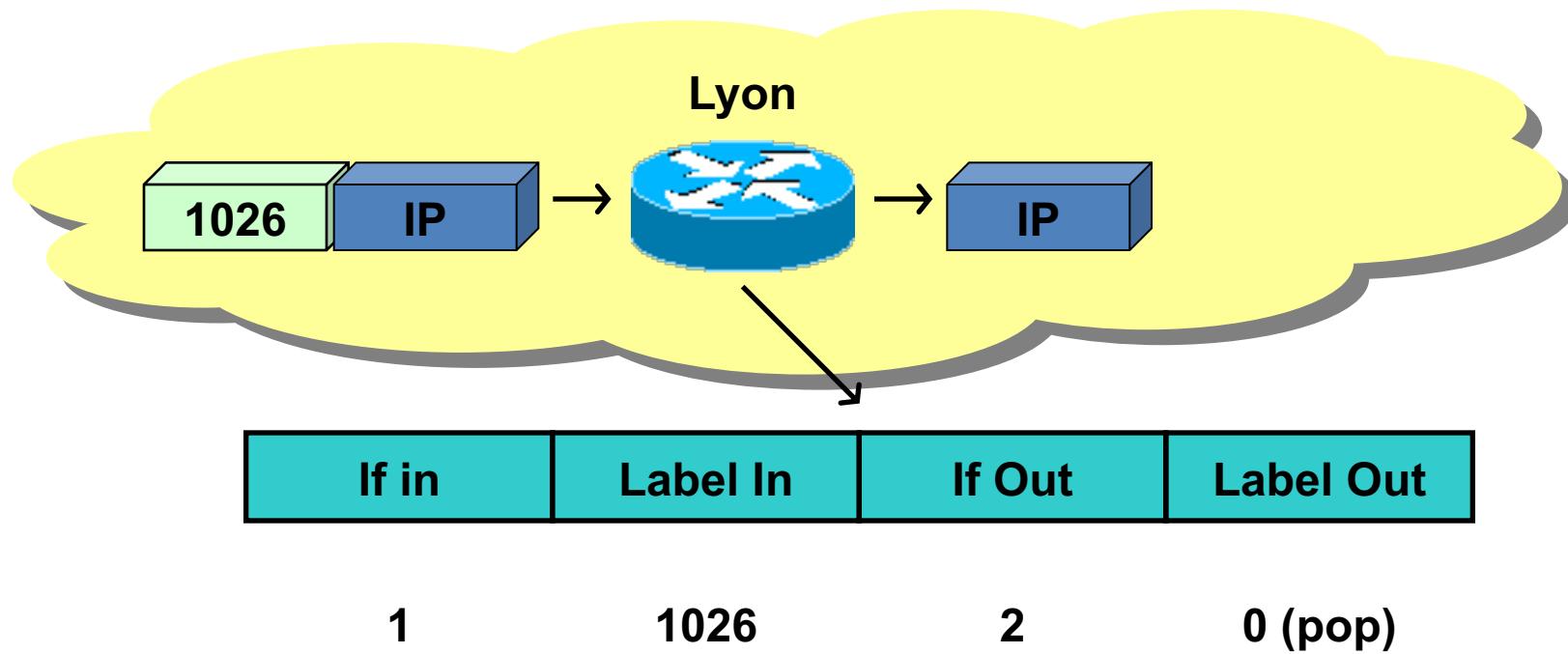


Principes

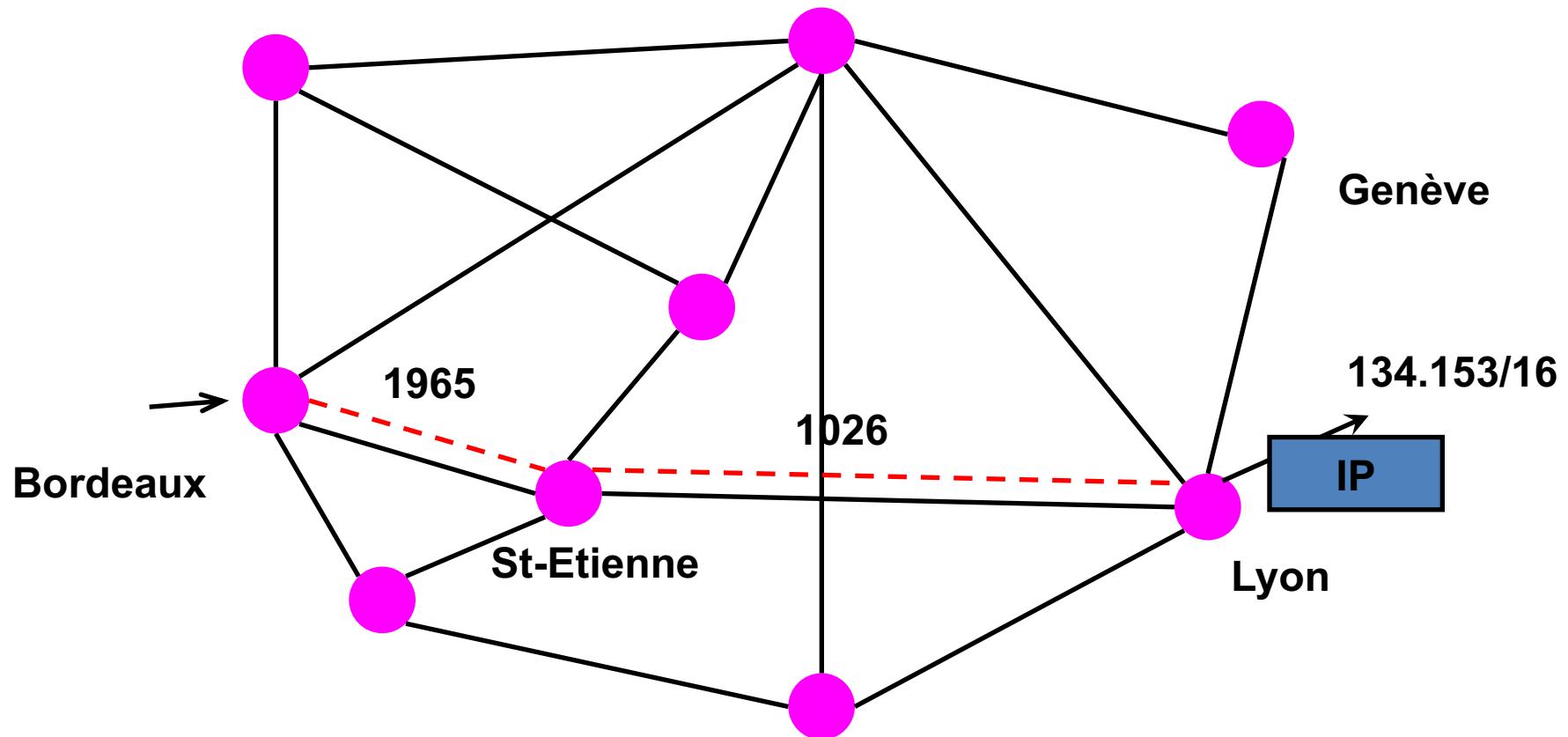


Principes

- Etape n : Pop
 - Décapsulation du label à l'avant-dernier saut



Principes



Les FEC

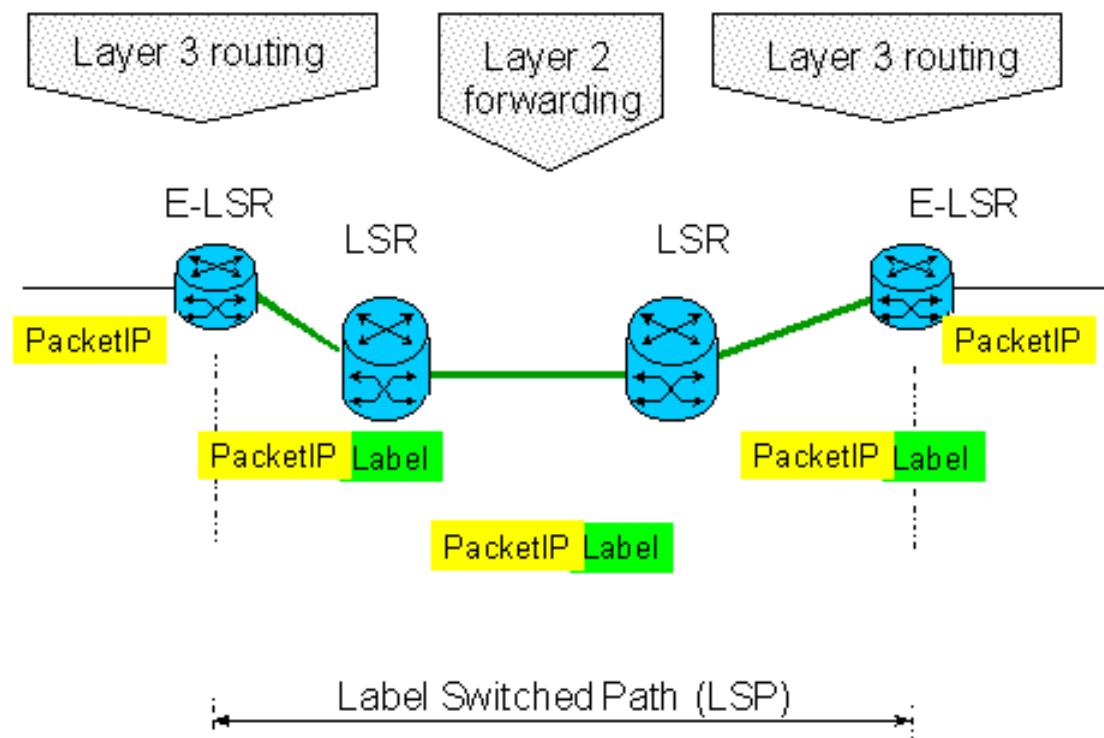
- Comment attribuer un label à un paquet IP ?
 - A l'entrée dans le réseau : classer les paquets dans des FEC
 - Classement par **Forwarding Equivalent Class** suivant :
 - Adresse IP source
 - Adresse IP destination
 - Paramètres de QoS
 - Choix d'une FEC s'appuie sur le protocole de routage interne (igp tel OSPF) et sur le protocole de distribution de label
 - **Une FEC est associée à un LABEL local dans chaque LSR**

- FEC
 - “Forward Equivalence Class”
 - L’ensemble des paquets qui subissent
 - . la même décision de routage (“next hop”)
 - . le même traitement (par ex., même file d’attente)
 - Par exemple, sous IP les paquets ayant
 - . le même préfixe d’adresse de destination
 - . la même adresse de destination
 - . idem + même numéro de port (source ou/et dest.)
- La création d’une FEC peut être basée sur :
 - le type de flux (QoS : RSVP, etc)
 - informations de routage
 - . par ex., tous les paquets convergeant vers le même point

Rétention des labels

- Mode « **Liberal** »
 - Un LSR conserve tous les labels annoncés, même ceux non utilisés
 - Convergence rapide
 - Consommateur en mémoire
 - Utilisé en mode « Downstream Unsolicited »
- Mode « **Conservative** »
 - Les LSR conservent seulement les labels reçus des next-hop pour chaque FEC
 - Convergence plus lente
 - Peu de consommation mémoire
 - Utilisé en mode « Downstream On Demand »

Label Switch Path (LSP)



Label Switch Path (LSP)

- LSP
 - Label Switched Path
 - Construction de LSP grâce à RSVP (allocation Bande Passante)
 - Mise en place de contrainte dans RSVP pour optimiser des ressources
- Contraintes
 - Bande passante
 - Couleur de lien
 - Préemption...