Border Gateway Protocol (BGP4)

AFNOG 2001, 2002, 2004, 2005

Border Gateway Protocol (BGP)

- Rappels : bases du routage
- Briques élémentaires
- Exercices
- Bases du protocole BGP
- Exercices
- Attributs de routes BGP
- Calcul du meilleur chemin
- Exercices

Border Gateway Protocol (BGP)...

- Topologies typiques avec BGP
- Politiques de routage
- Exercices
- Redondance / Partage de charge
- Etat de l'art (BCP, Best Current Practices)

Le routage : quelques bases

Routage IP

- Chaque routeur (ou machine) décide comment acheminer un paquet
- L'expéditeur n'a pas à connaître le chemin jusqu'à la destination
- L'expéditeur doit seulement déterminer le prochain saut (next-hop).
  - Ce processus est répété jusqu'à arriver à la destination
- La table de routage est consultée afin de déterminer le prochain saut

Routage IP

- Routage par préfixe (Classless routing)
  - une route est composée de
    - la destination
    - l'adresse du prochain routeur (next-hop)
    - le masque de réseau permet de déterminer la taille de l'espace d'adressage concerné (-> préfixe)
- Choix du préfixe le plus long
  - pour une destination donnée, il faut prendre la route la plus spécifique (le préfixe le plus grand)
  - exemple: adresse destination 35.35.66.42
    - la table de routage contient 35.0.0.0/8, 35.35.64.0/19 and 0.0.0.0/0

Routage IP

- Route par défaut (default route)
  - indique où expédier un paquet si la table de routage ne contient pas une route spécifique
  - c'est une configuration courant : la plupart des machines disposent d'une (et une seule) route par défaut
  - autre nom : passerelle par défaut (default gateway)

Les routes spécifiques sont utilisées en premier

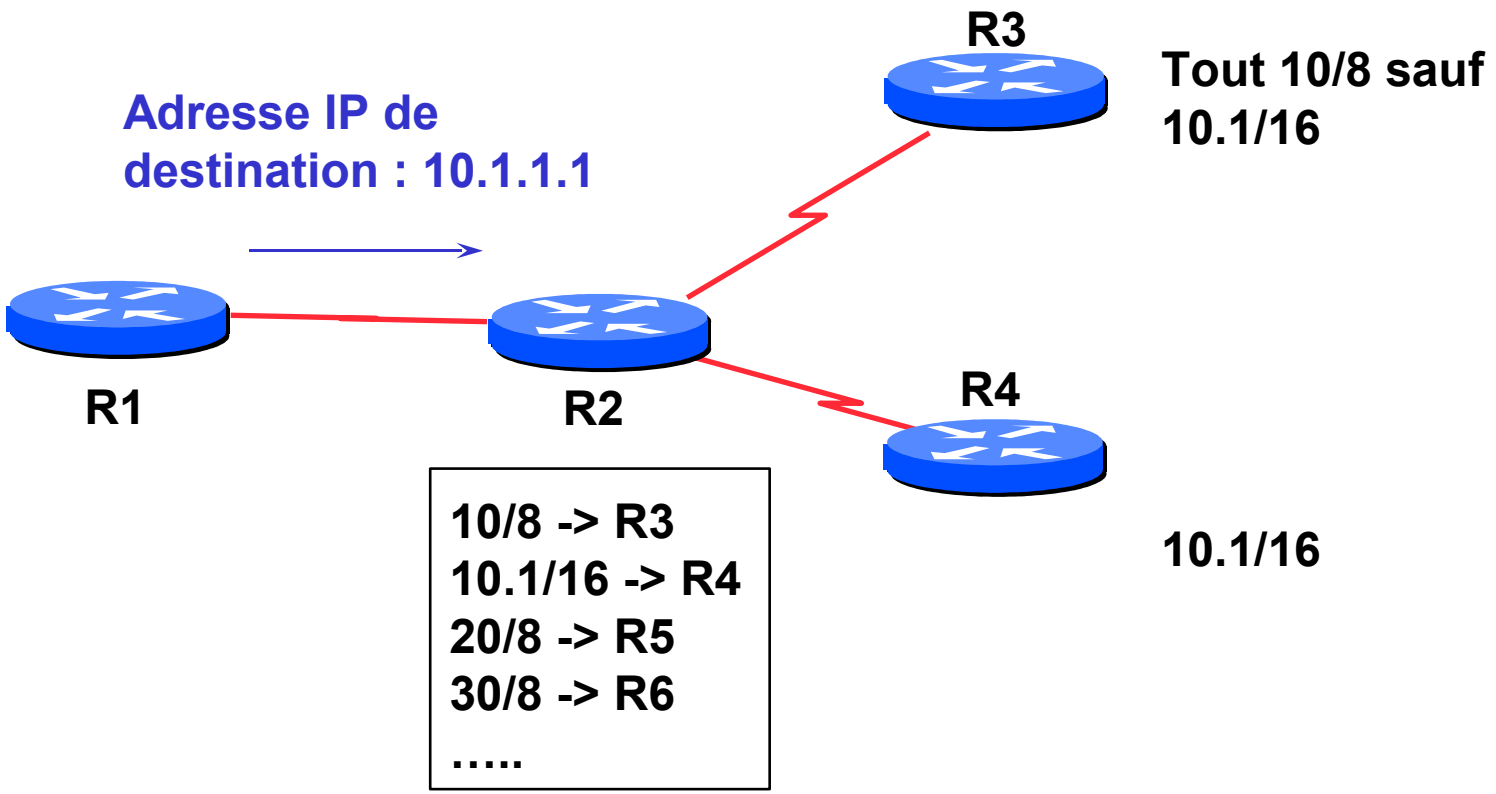


Table de routage IP de R2

Les routes spécifiques sont utilisées en premier

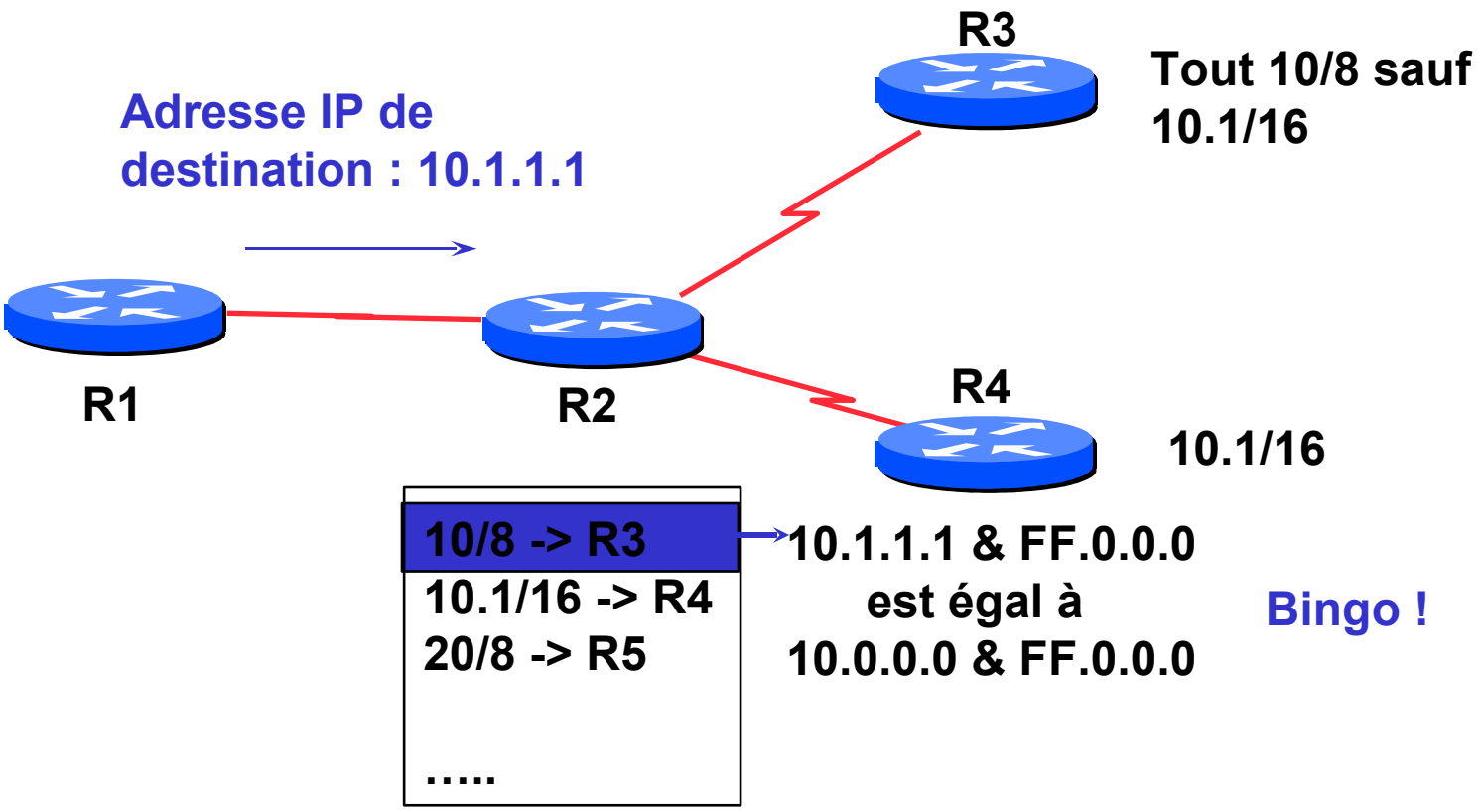


Table de routage IP de R2

Les routes spécifiques sont utilisées en premier

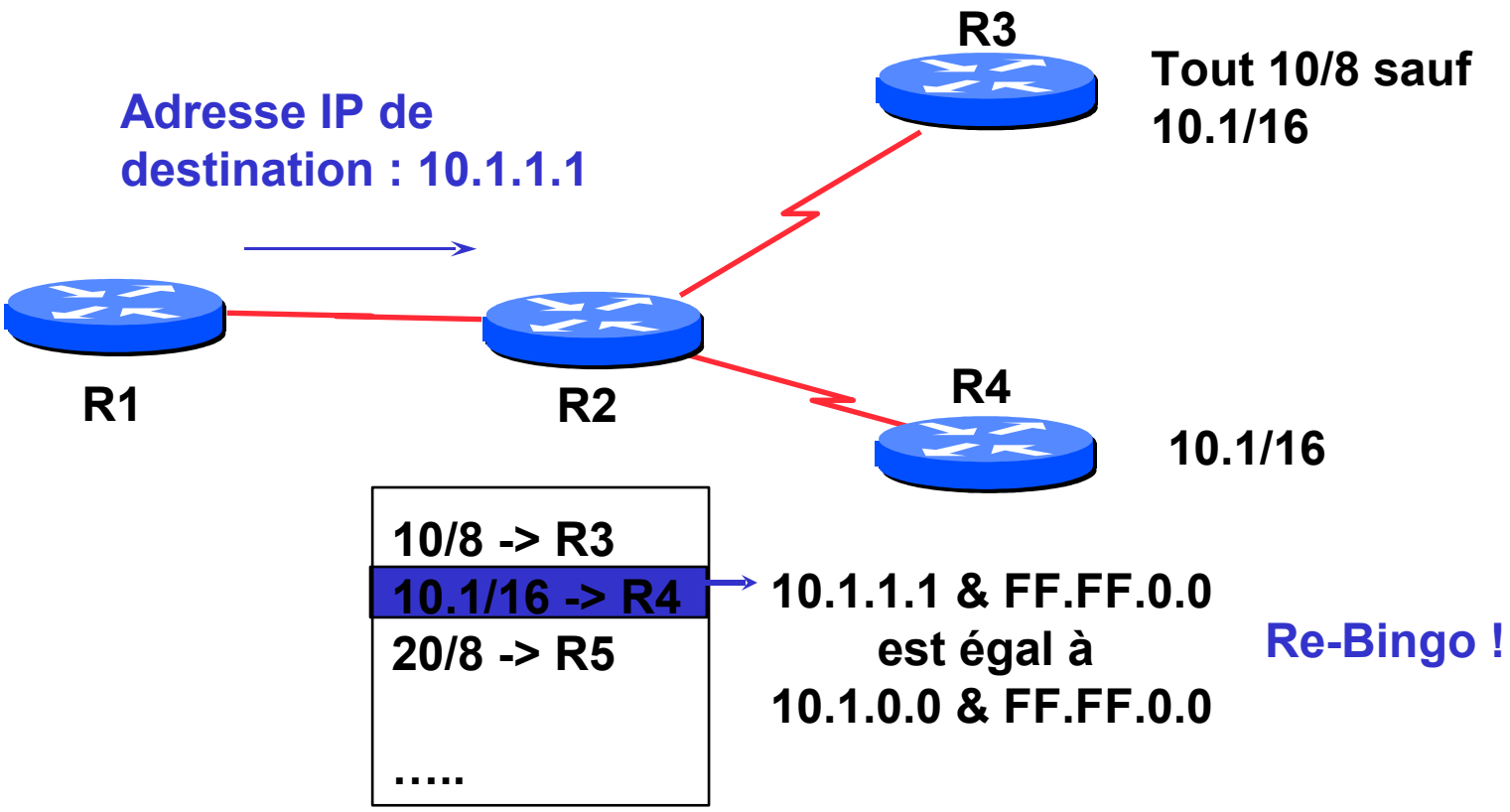


Table de routage IP de R2

Les routes spécifiques sont utilisées en premier

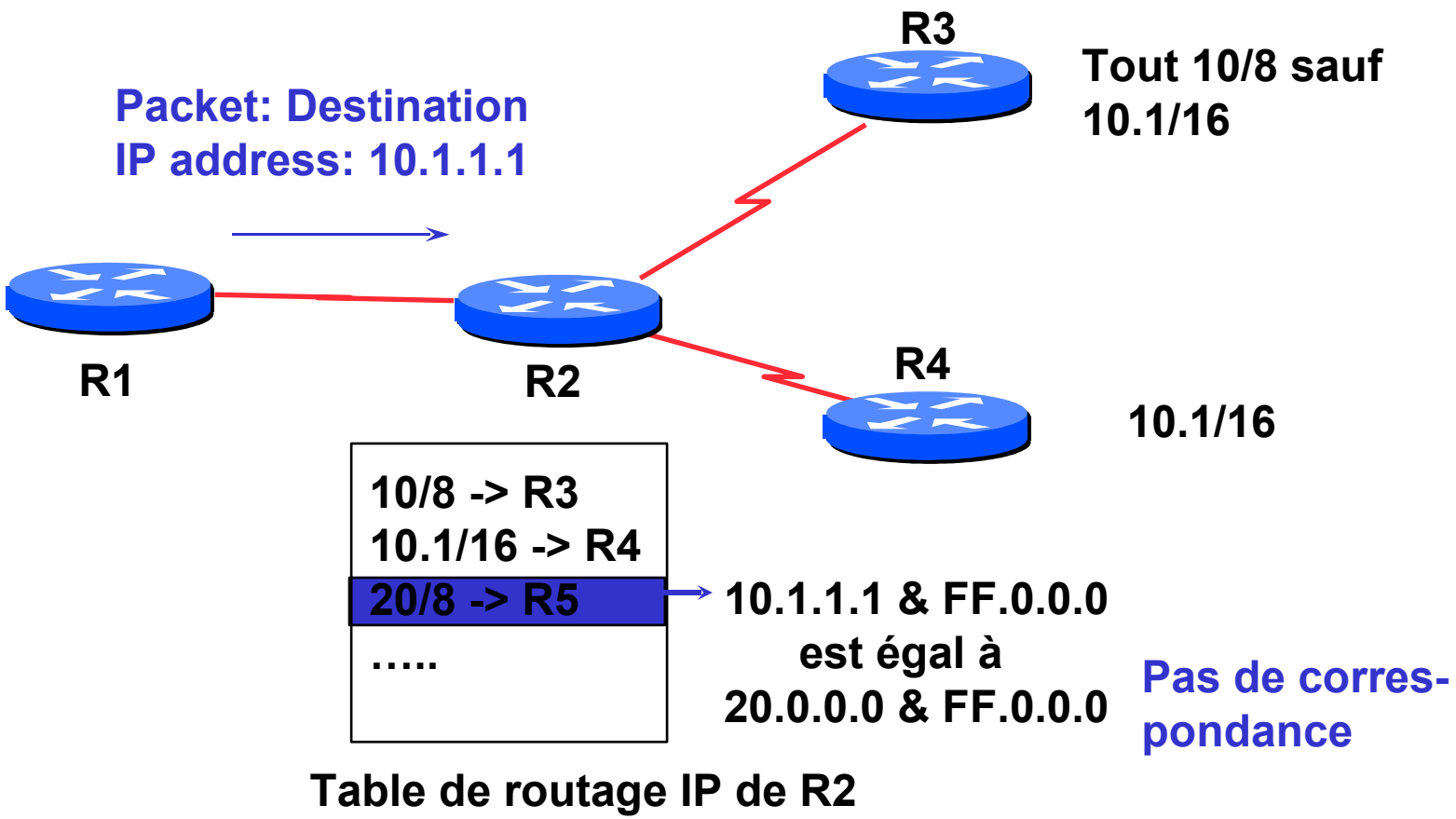


Table de routage IP de R2

Les routes spécifiques sont utilisées en premier

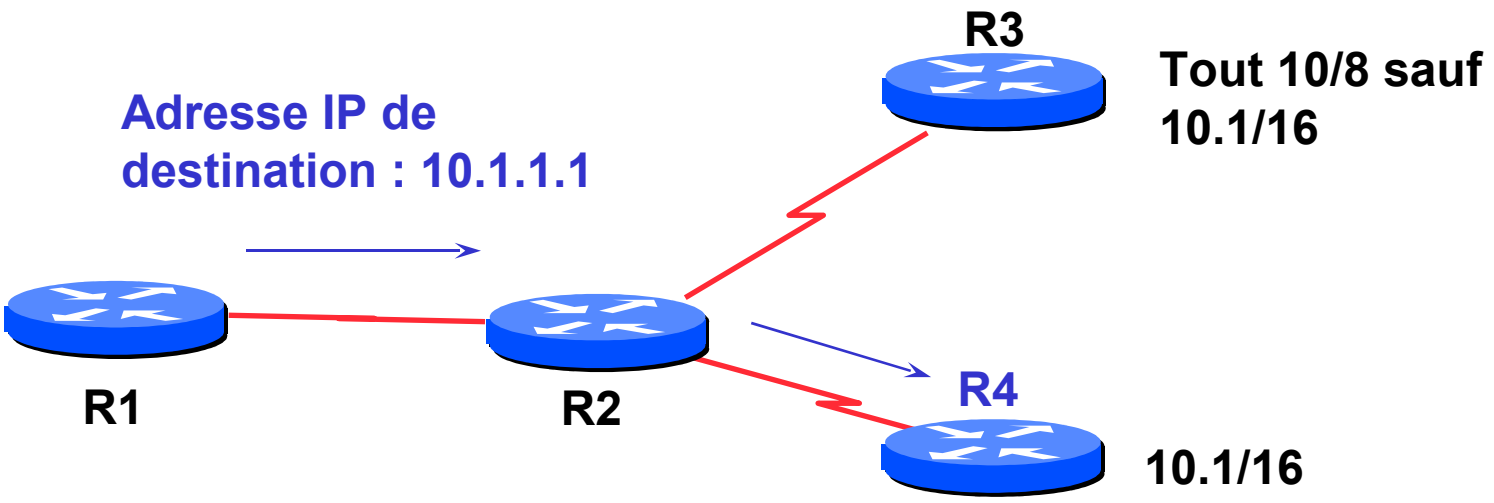


Table de routage IP de R2

Les routes spécifiques sont utilisées en premier

- On utilise toujours la route la plus spécifique (celle qui correspond au plus petit volume d'adresses IP)
- La route par défaut est notée 0.0.0.0/0
  - ce qui permet d'utiliser l'algorithme décrit ci-dessus
  - Il y a toujours correspondance. C'est la route la moins spécifique.

Routage dynamique

- Les routeurs déterminent leur table de routage automatiquement à partir des informations reçues des autres routeurs
- Les routeurs s'échangent les information de topologie en utilisant divers protocoles
- Les routeurs calculent ensuite un ou plusieurs "next-hops" pour chaque destination en essayant d'emprunter le meilleur chemin

Table d'acheminement

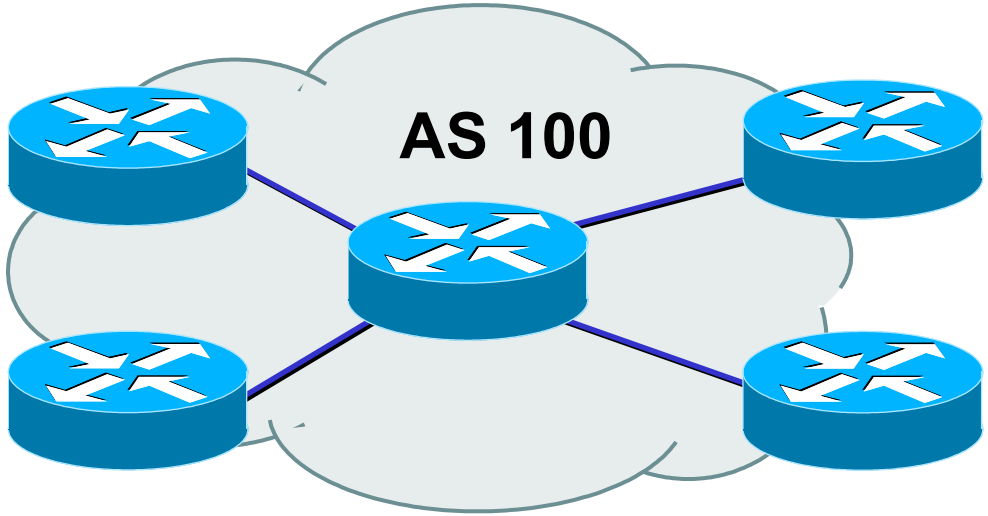
- En anglais : forwarding table
- Permet de déterminer comment acheminer un paquet dans le routeur
- Construite à partir de la table de routage
  - Les meilleurs routes sont choisies dans la table de routage
- Effectue une recherche pour déterminer le prochain saut et l'interface de sortie
- Commute le paquet sur l'interface de sortie avec l'encapsulation adéquate (ex : PPP, FR, POS)

Briques élémentaires

Briques élémentaires

- Système autonome - Autonomous System (AS)
- Type de routes
- IGP/EGP
- DMZ (zone démilitarisée)
- Politique
- Trafic sortant
- Trafic entrant

Système autonome (AS)

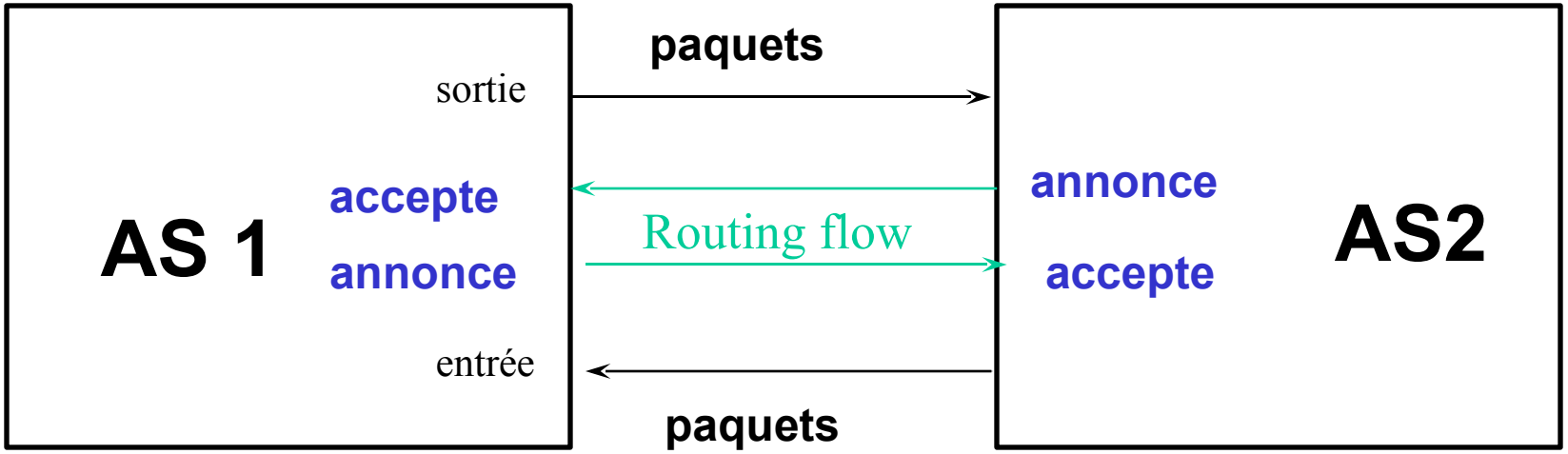


- Ensemble de réseaux partageant la même politique de routage
- Utilisation d'un même protocole de routage
- Généralement sous une gestion administration unique
- Utilisation d'un IGP au sein d'un même AS

Système autonome (AS)...

- Caractérisé par un numéro d'AS
- Il existe des numéros d'AS privés et publics
- Exemples :
  - Prestataire de services Internet
  - Clients raccordés à plusieurs prestataires
  - Quiconque souhaite établir une politique de routage spécifique

Flux de routes et de paquets



Pour que AS1 et AS2 puissent communiquer :

- AS1 annonce des routes à AS2
- AS2 accepte des routes de AS1
- AS2 annonce des routes à AS1
- AS1 accepte des routes de AS2

Trafic en sortie

- Paquets qui quittent le réseau
  - Choix de la route (ce que les autres vous envoient)
  - Acceptation d'une route (ce que vous acceptez des autres)
  - Politique et configuration (ce que vous faites des annonces des autres)
  - Accords de transit et d'échange de trafic

Trafic entrant

- Paquets entrant dans votre réseau
- Ce trafic dépend de :
  - Ce que vous annoncez à vos voisins
  - Votre adressage et plan d'AS
  - La politique mise en place par les voisins (ce qu'ils acceptent comme annonces de votre réseau et ce qu'ils en font)

Types de routes

- Routes statiques
  - configurées manuellement
- Routes “connectées
  - crées automatiquement quand une interface réseau est “active”
- Routes dites “intérieures”
  - routes au sein d’un AS
  - routes apprises par un IGP
- Routes dites “extérieures”
  - routes n’appartenant pas à l’AS local
  - apprises par un EGP

Politique de routage

- Définition de ce que vous acceptez ou envoyez aux autes
  - connexion économique, partage de charge, etc...
- Accepter des routes de certains FAI et pas d'autres
- Envoyer des routes à certains FAI et pas à d'autres
- Préferrer les routes d'un FAI plutôt que d'un autre

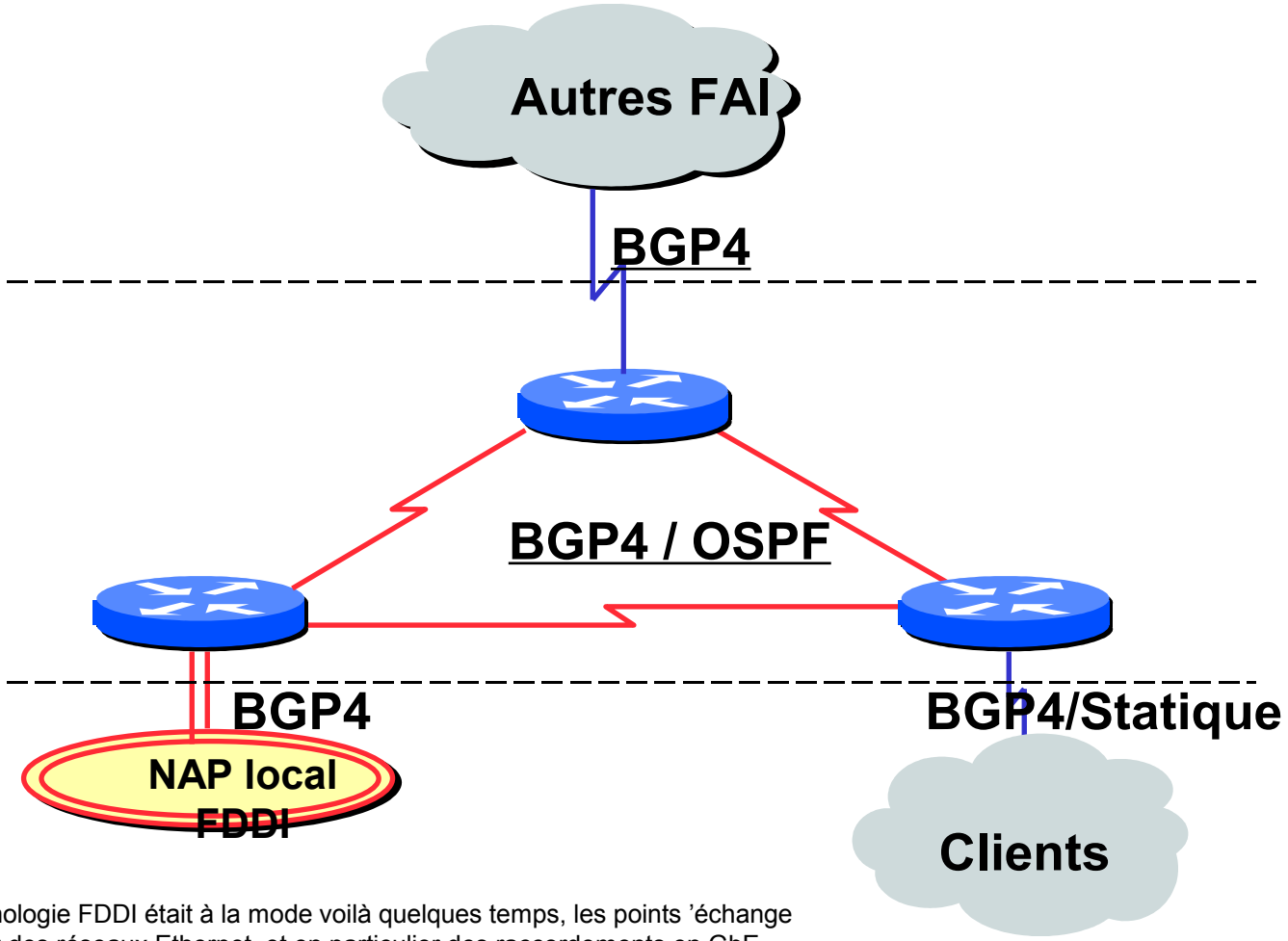
Pourquoi a-t-on besoin d'un EGP ?

- S'adapter à un réseau de grande taille
  - hiérarchie
  - limiter la portée des pannes
- Définir des limites administratives
- Routage politique
  - contrôler l'accessibilité des préfixes (routes)

Protocoles intérieurs vs. extérieurs

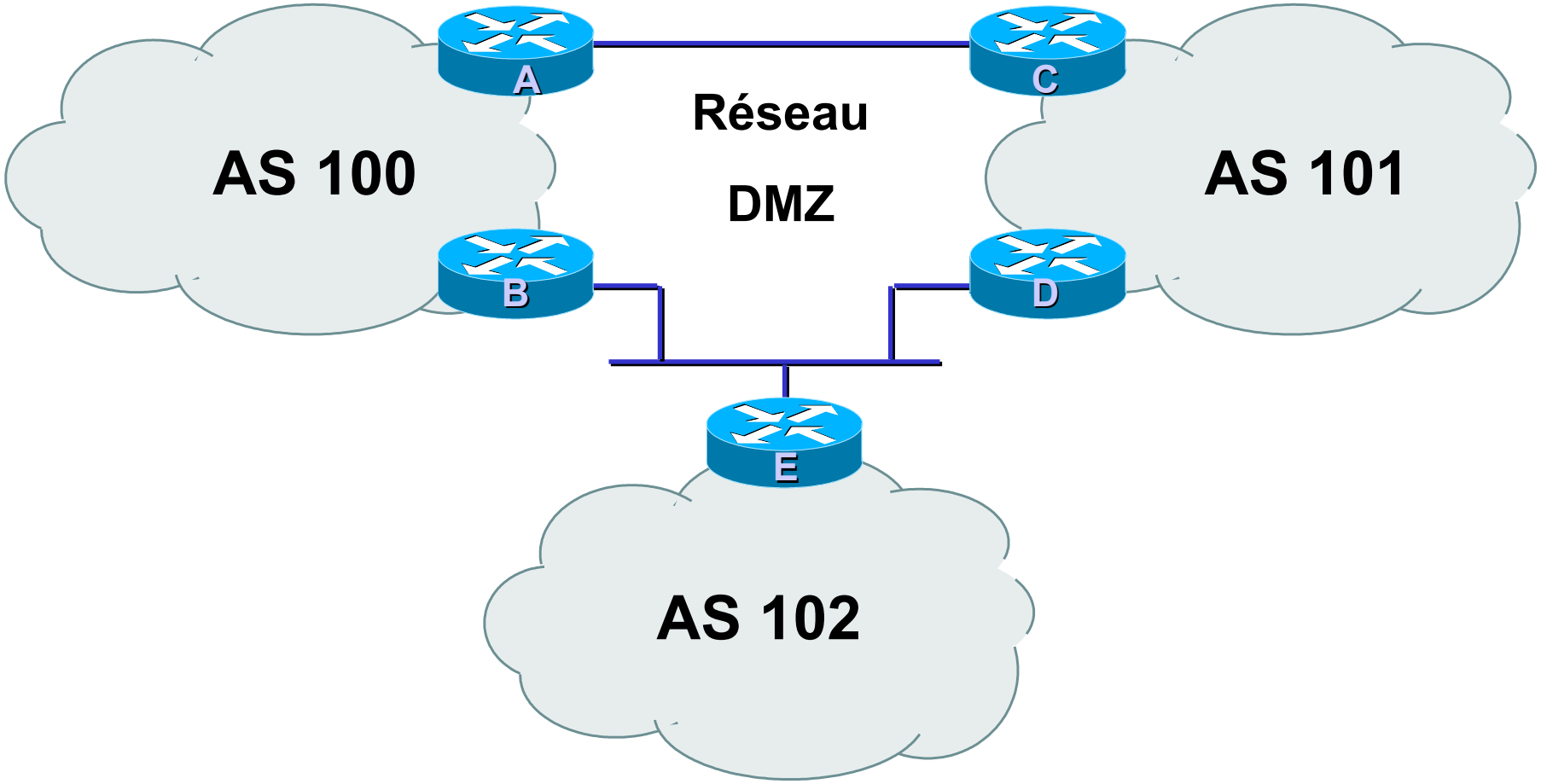
- Intérieurs (IGP)
  - Découverte automatique
  - Confiance accordée aux routeurs de l'IGP
  - Les routes sont diffusées sur l'ensemble des routeurs de l'IGP
- Extérieurs (EGP)
  - Voisins explicitement déclarézs
  - Connexion avec des réseaux tiers
  - Mettre des limites administratives

Hiérarchie dans les protocoles



Note: la technologie FDDI était à la mode voilà quelques temps, les points ’échange utilisent plutôt des réseaux Ethernet, et en particulier des raccordements en GbE ou 10 GbE.

Zone démilitarisée (DMZ)



- Le réseau démilitarisé est partagé entre plusieurs AS

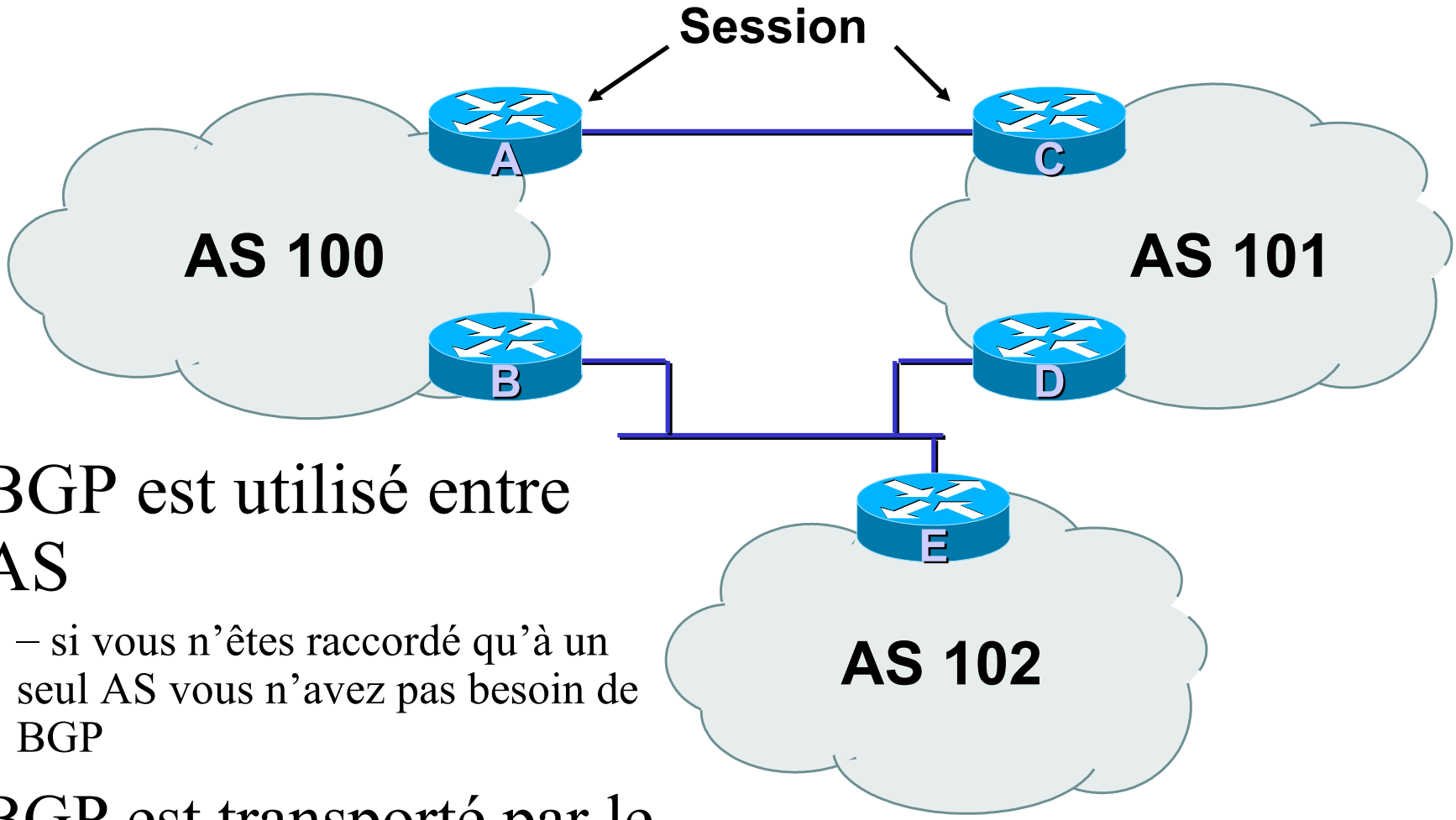
Gestion de l'adressage (FAI)

- Il faut réserver des adresses IP pour son propre usage
- Des adresses IP sont également allouées aux clients
- Il faut prendre en considération la croissance de l'activité
- Le prestataire "upstream" attribuera les adresses d'interconnexion dans ses blocs

Bases de BGP

- Bases concernant le protocole
- Vocabulaire
- Messages
- Exploitation d'un routeur BGP
- Types de sessions BGP (eBGP/iBGP)
- Comment annoncer les routes

Principes de base du protocole



- BGP est utilisé entre AS
  - si vous n'êtes raccordé qu'à un seul AS vous n'avez pas besoin de BGP
- BGP est transporté par le protocole TCP

Principes de base (2)

- Les mises à jours sont incrémentielles
- BGP conserve le chemin d'AS pour atteindre un réseau cible
- De nombreuses options permettent d'appliquer une politique de routage

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

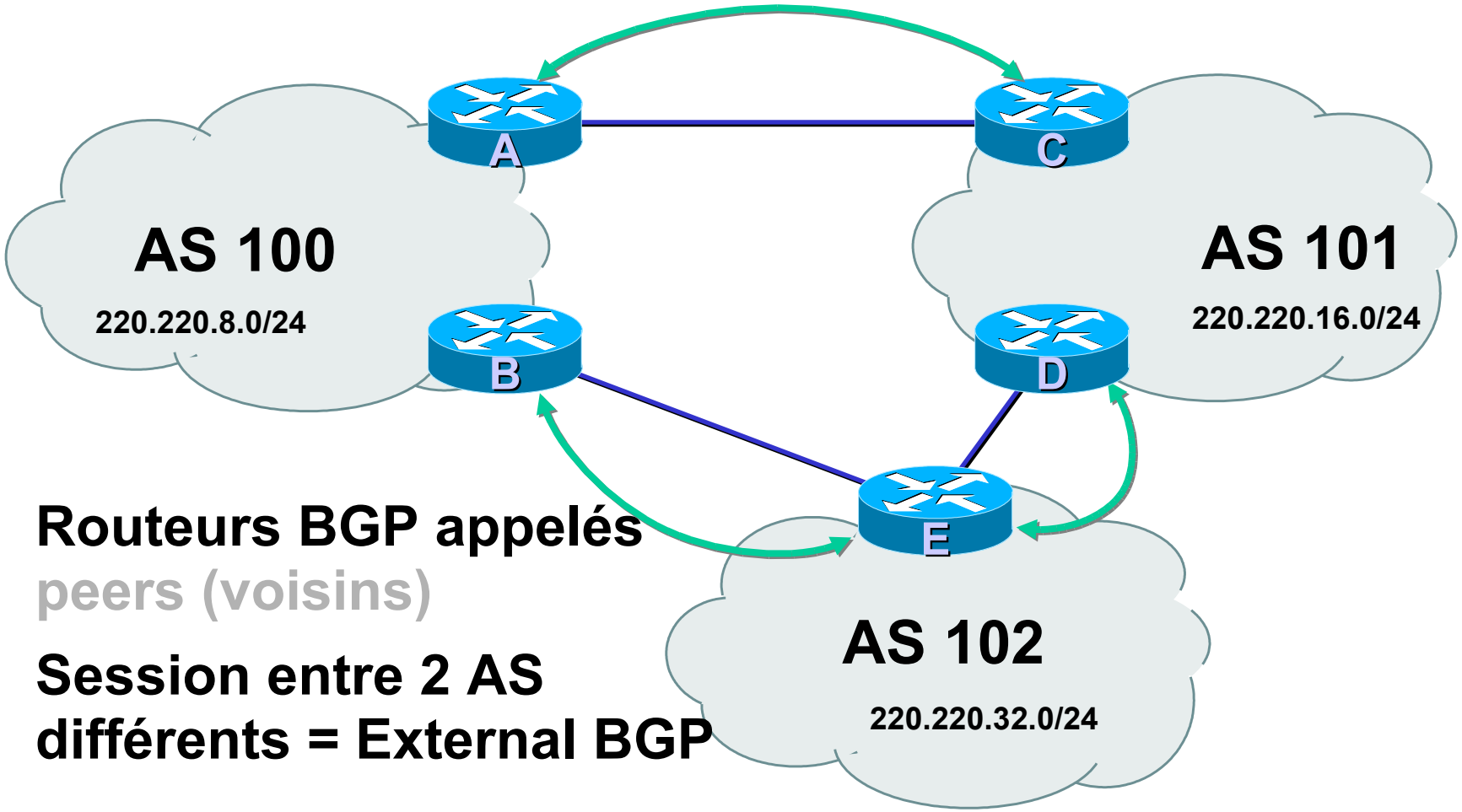
Vocabulaire (2)

- 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

Bases de BGP ...

- Chaque AS est le point de départ d'un ensemble de préfixes (NLRI)
- Les préfixes sont échangés dans les sessions BGP
- Plusieurs chemins possibles pour un préfixe
- Choix du meilleur chemin pour le routage
- Les attributs et la configuration "politique" permettent d'influencer ce choix du meilleur chemin

Sessions BGP



Routeurs BGP appelés peers (voisins)

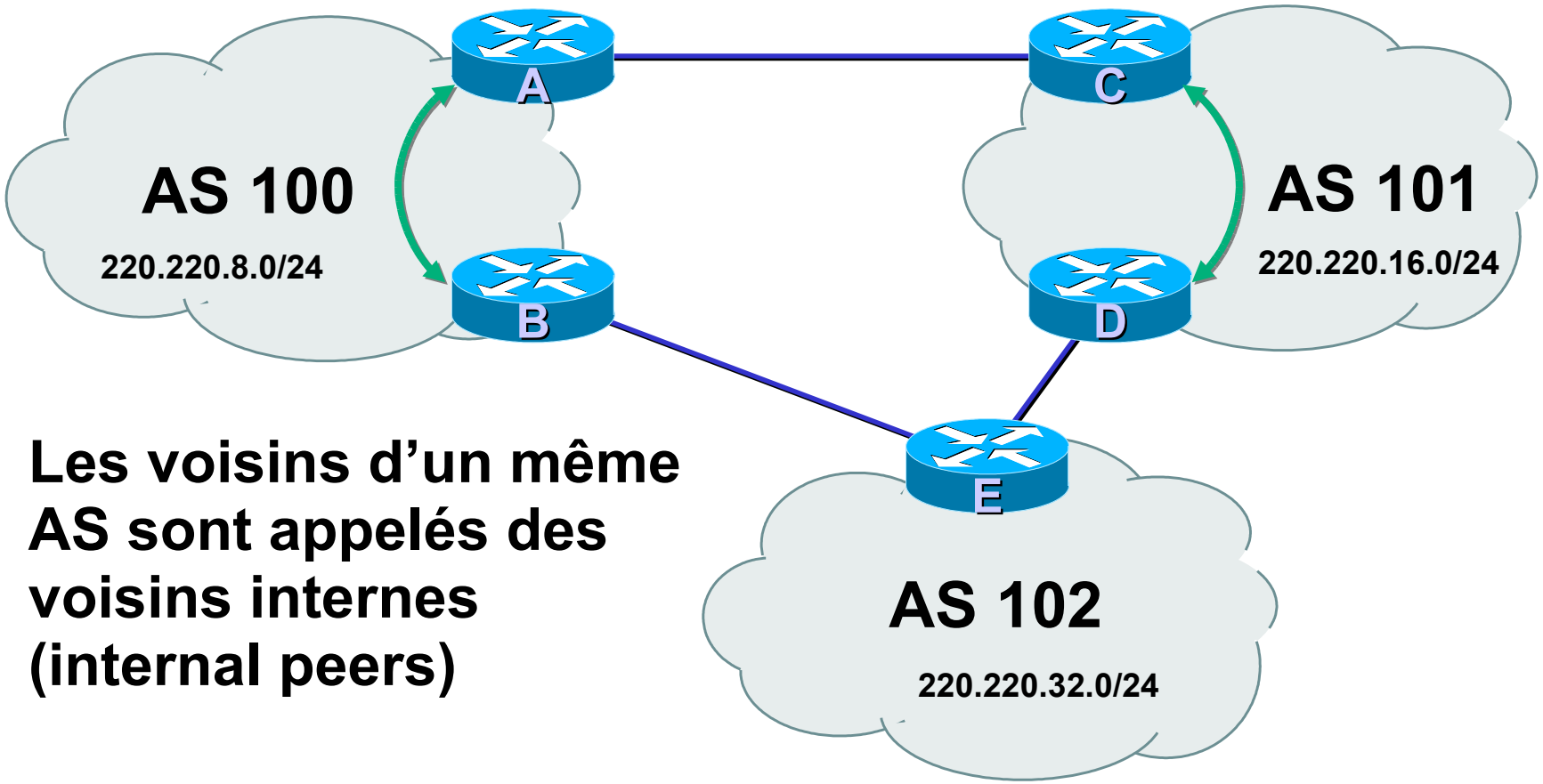
Session entre 2 AS différents = External BGP

Connexion TCP/IP

eBG-

Note: les voisins eBGP doivent être directements raccordés.

Sessions BGP



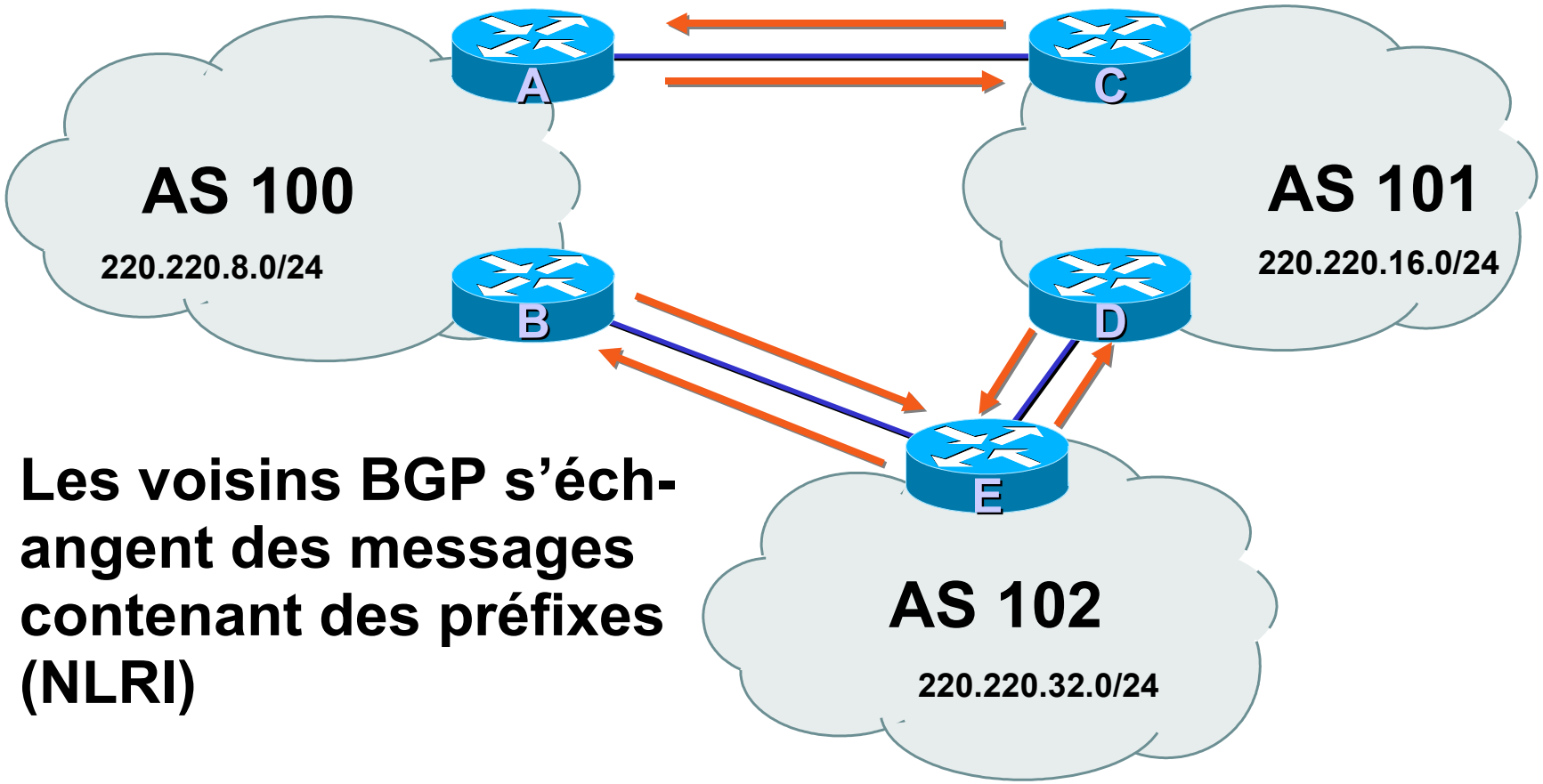
Les voisins d'un même AS sont appelés des voisins internes (internal peers)

Connexion TCP/IP

iBGP

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

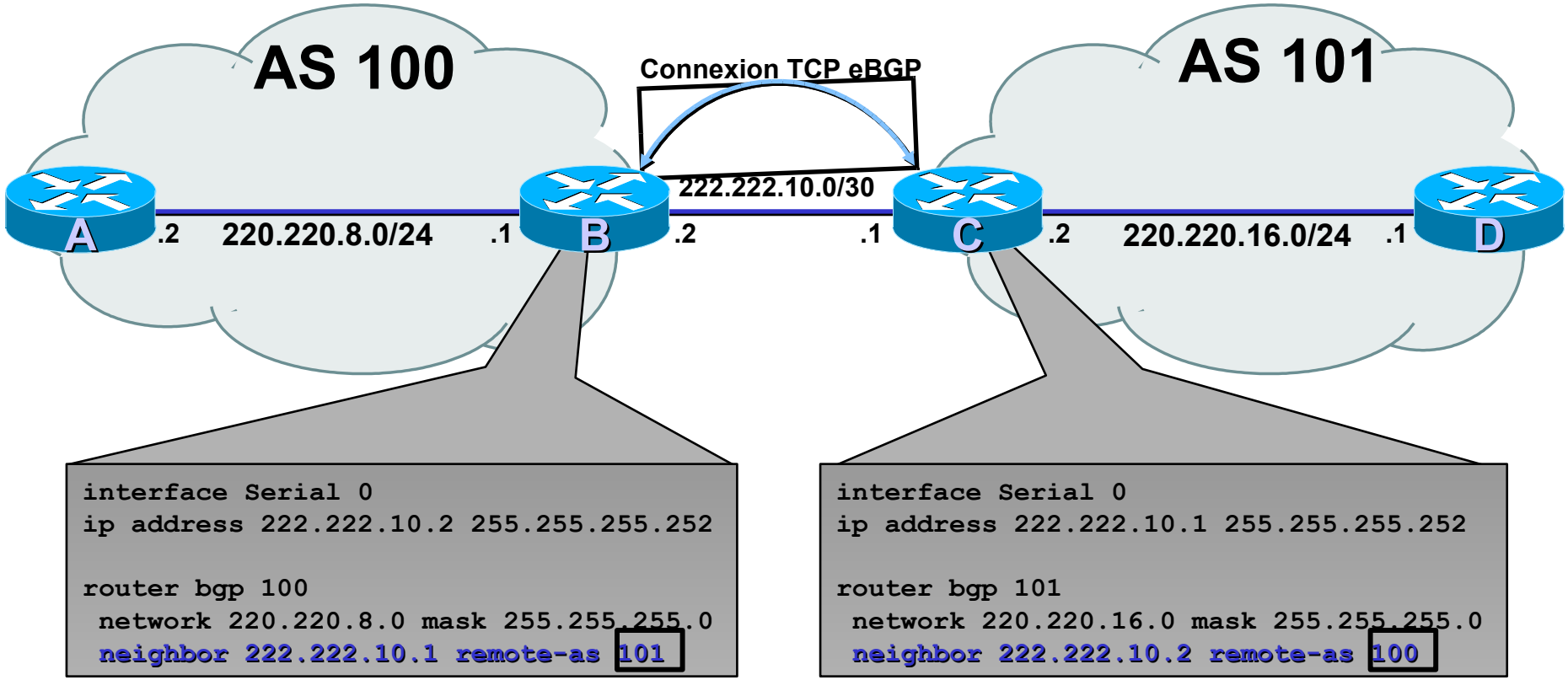
Sessions BGP



Les voisins BGP s'échangent des messages contenant des préfixes (NLRI)

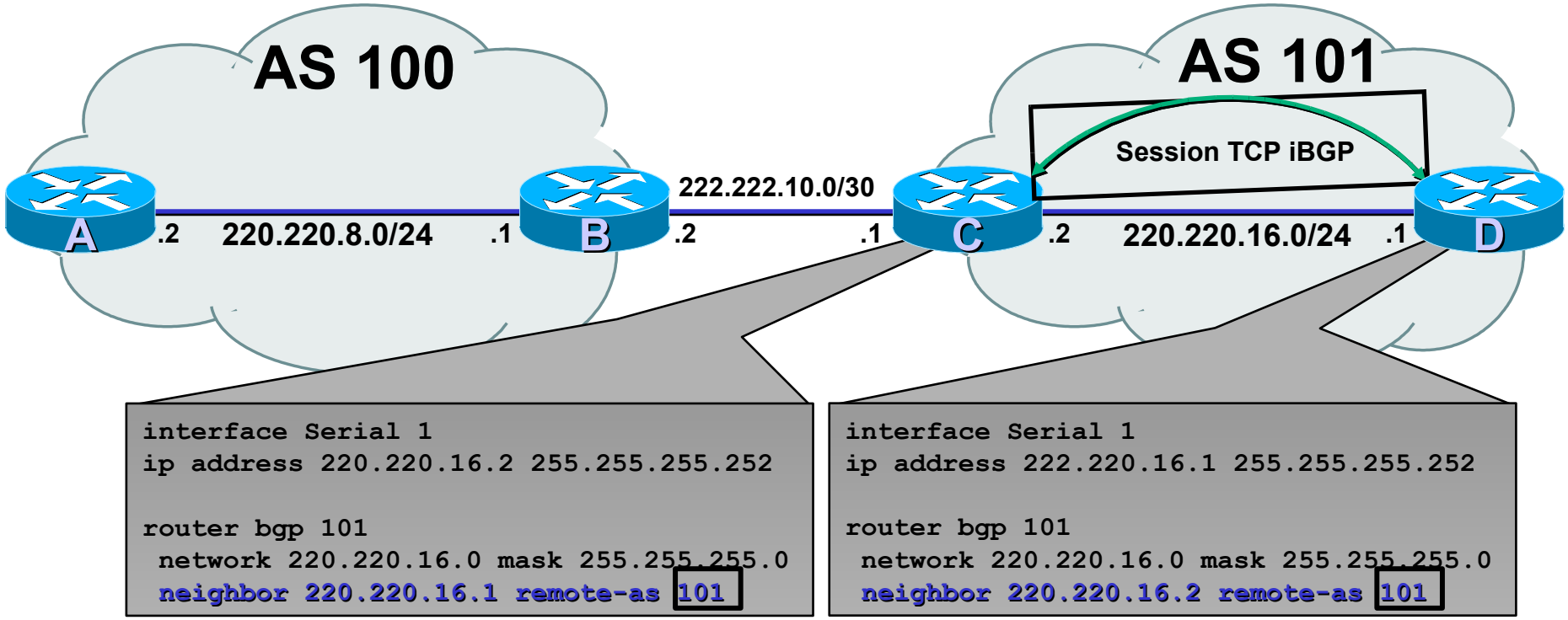
Message de mise à jour BGP

Configuration de sessions BGP



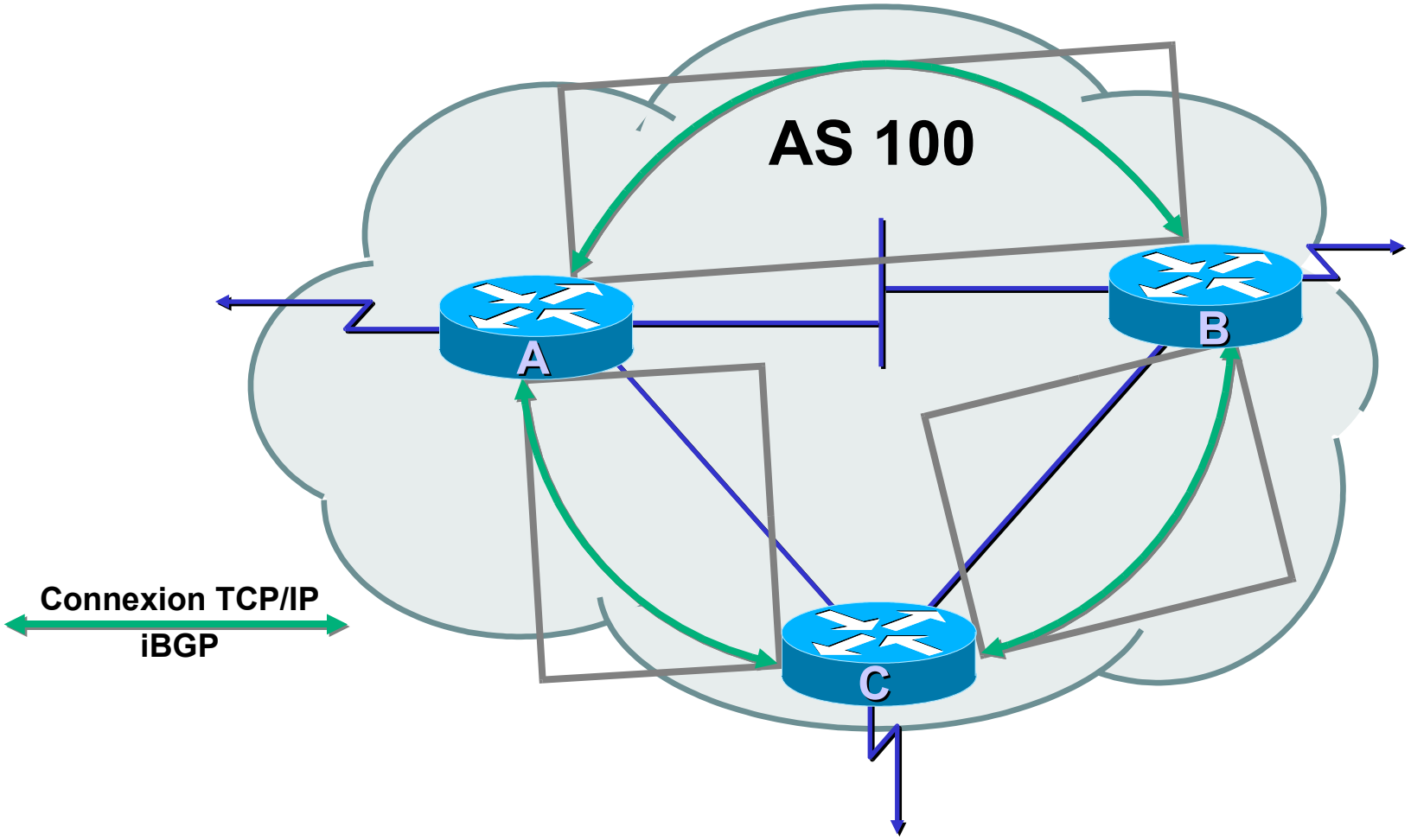
- **Les sessions BGP sont établies en utilisant la commande BGP "neighbor" du routeur**
  - Lorsque les numéros d'AS sont différents il s'agit d'une session BGP Externe (eBGP)

Configuration de sessions BGP



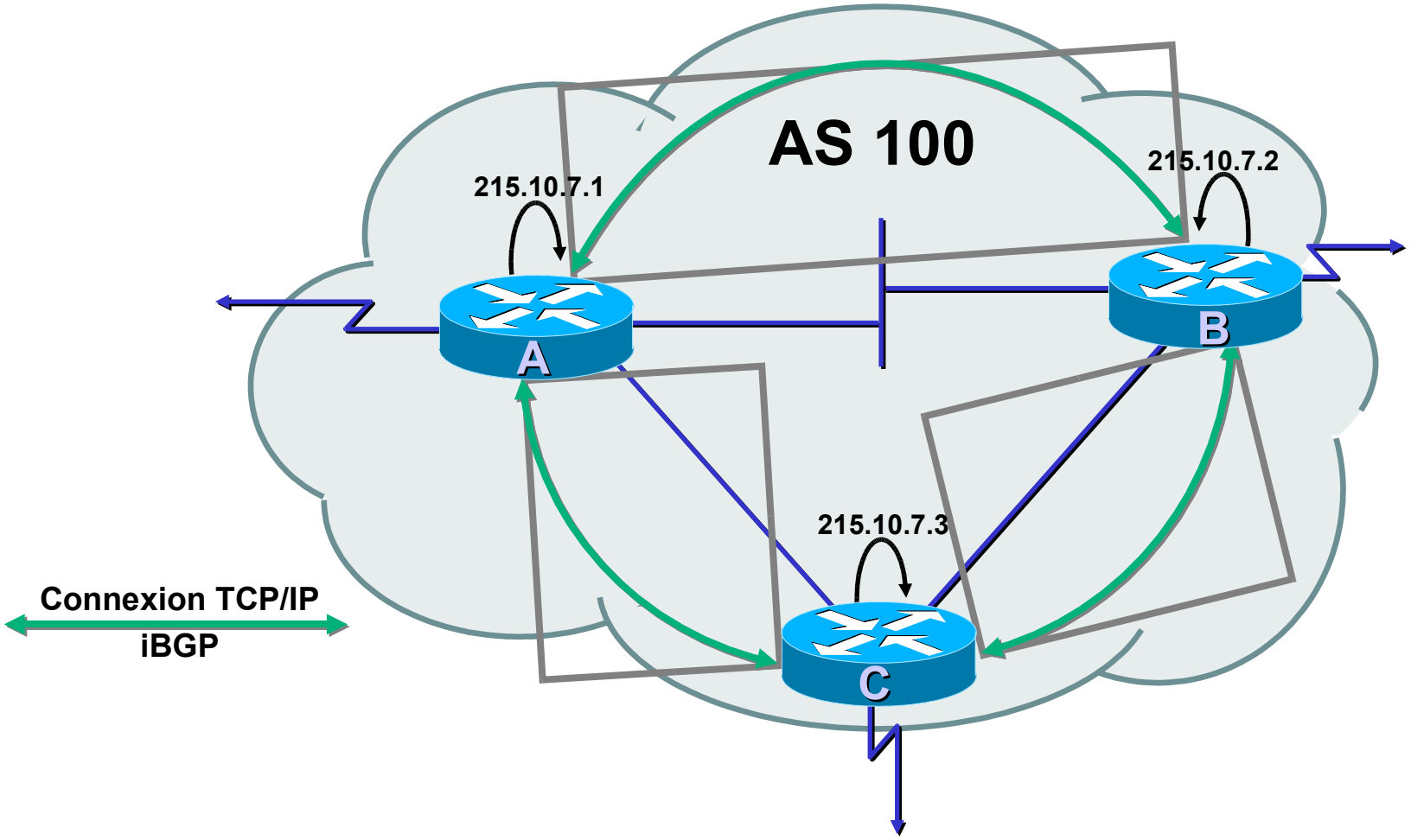
- **Les sessions BGP sont établies en utilisant la commande BGP "neighbor" du routeur**
  - Numéros d'AS différents -> BGP Externe (eBGP)
  - Numéros d'AS identiques -> BGP Interne (iBGP)

Configuration de sessions BGP



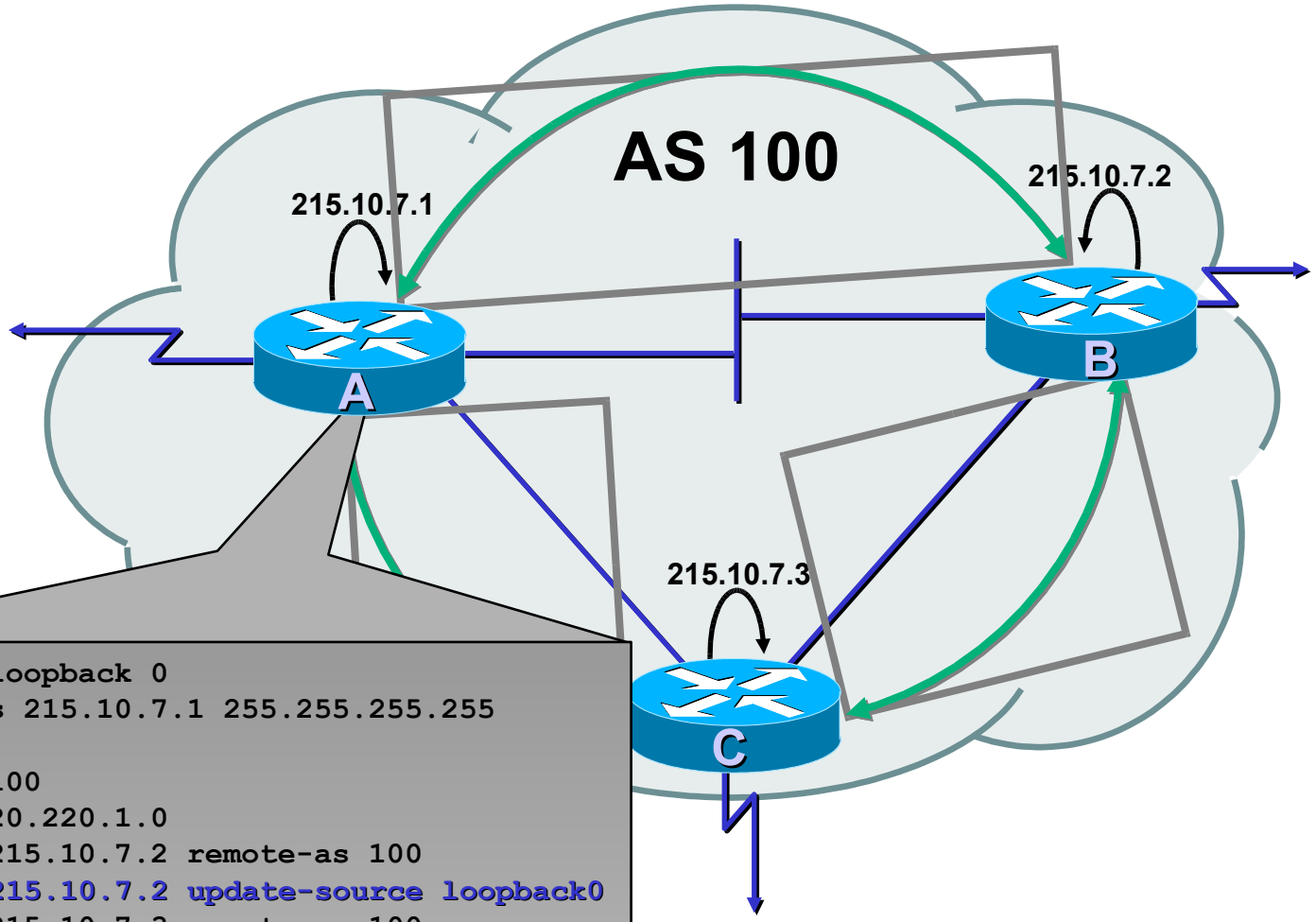
- Chaque routeur iBGP doit établir une session avec tous les autres routeurs iBGP du même AS

Configuration de sessions BGP



- Il est recommandé d'utiliser des interfaces Loopback sur les routeurs comme extrêmités des sessions iBGP

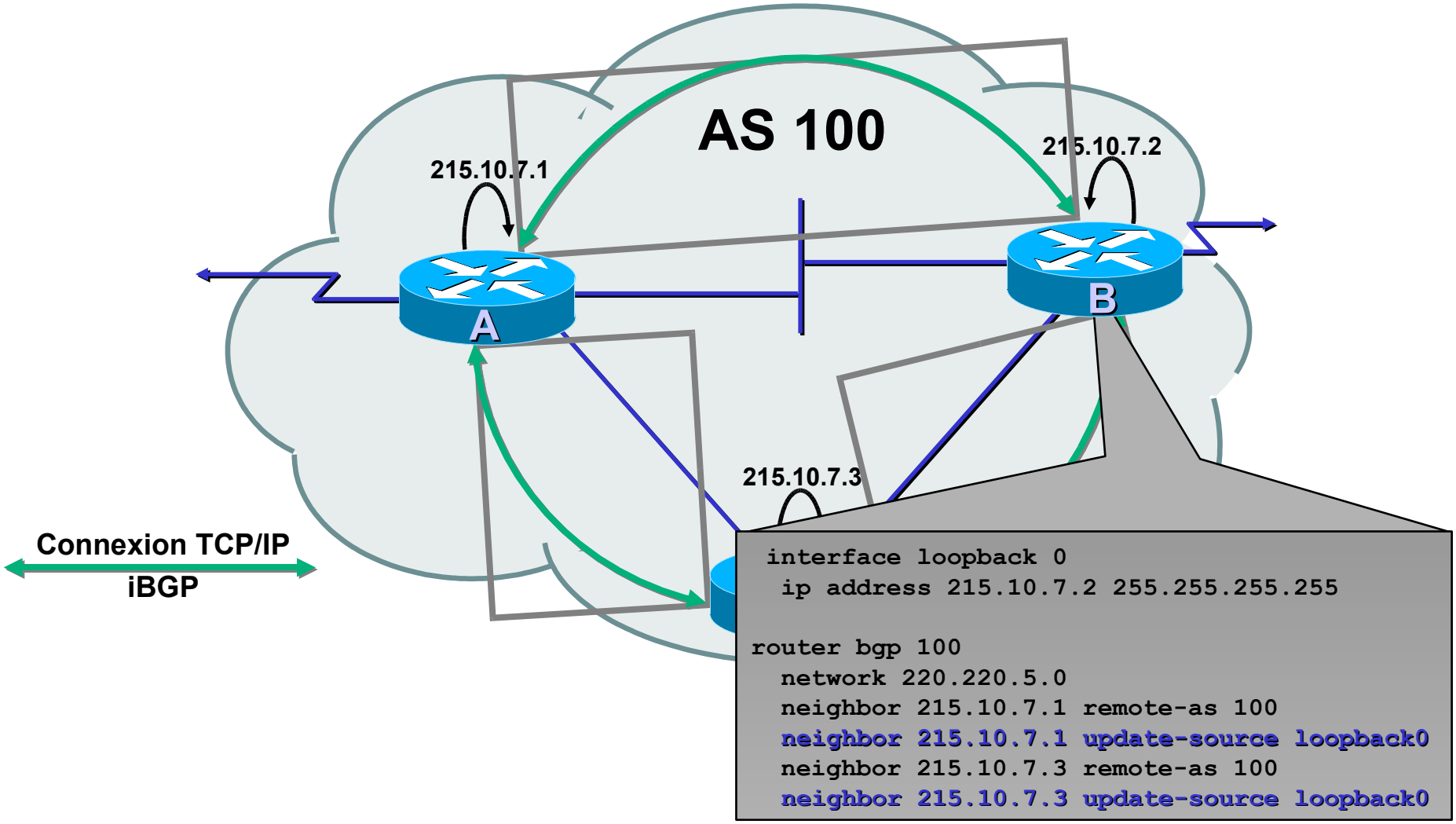
Configuration des sessions BGP



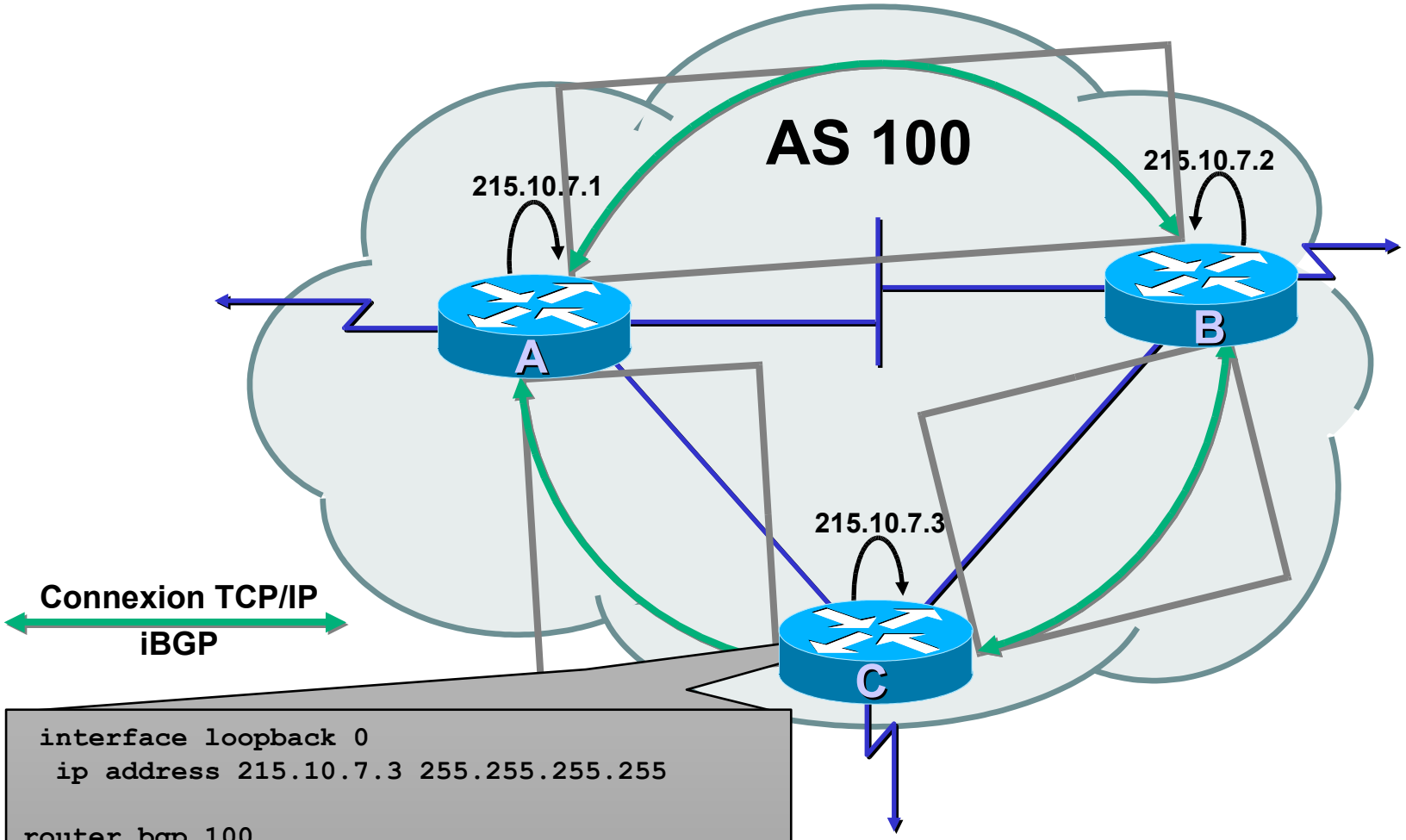

```
 interface loopback 0
  ip address 215.10.7.1 255.255.255.255

router bgp 100
  network 220.220.1.0
  neighbor 215.10.7.2 remote-as 100
  neighbor 215.10.7.2 update-source loopback0
  neighbor 215.10.7.3 remote-as 100
  neighbor 215.10.7.3 update-source loopback0
```

Configuration des sessions BGP



Configuration des sessions BGP

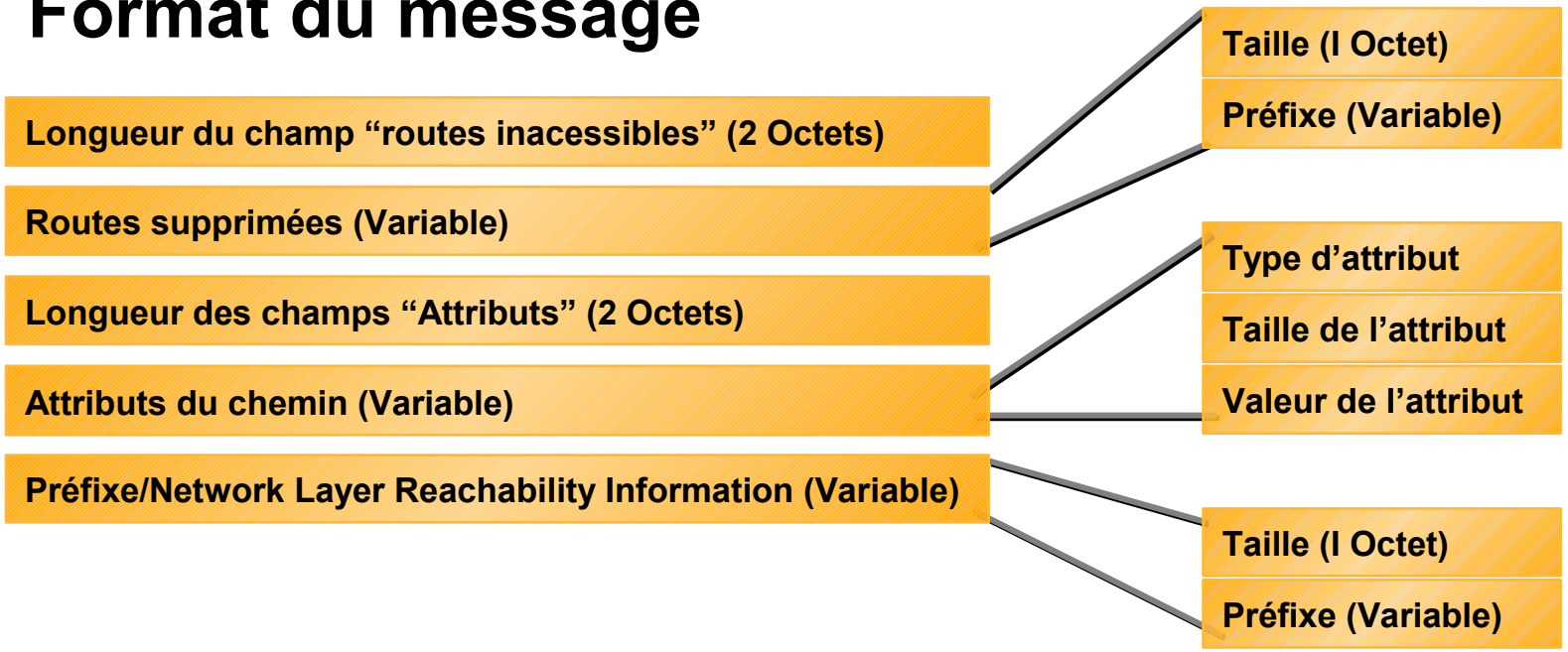



```
 interface loopback 0
  ip address 215.10.7.3 255.255.255.255

router bgp 100
  network 220.220.1.0
  neighbor 215.10.7.1 remote-as 100
  neighbor 215.10.7.1 update-source loopback0
  neighbor 215.10.7.2 remote-as 100
  neighbor 215.10.7.2 update-source loopback0
```

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 préfixes]
- Chaque message contient des attributs comme : origine, chemin d’AS, Next-Hop, ...

Mises à jour BGP — Préfixes/NLRI

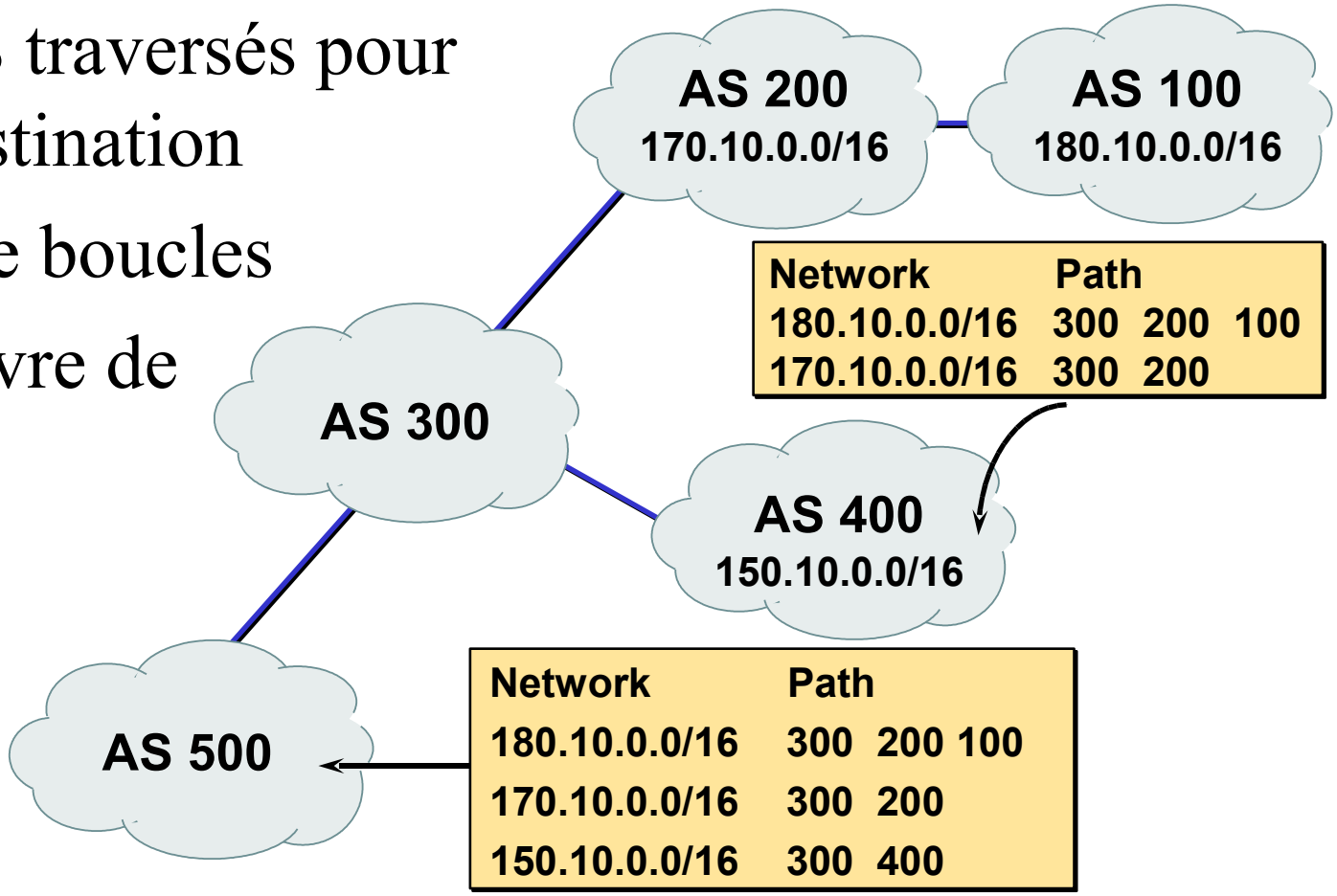
- NLRI = Network Layer Reachability Information = Préfixes
- Permet d'annoncer l'accessibilité d'une route
- Composé des informations suivantes :
  - Préfixe réseau
  - Longueur du masque

Mise à jour BGP — Attributs

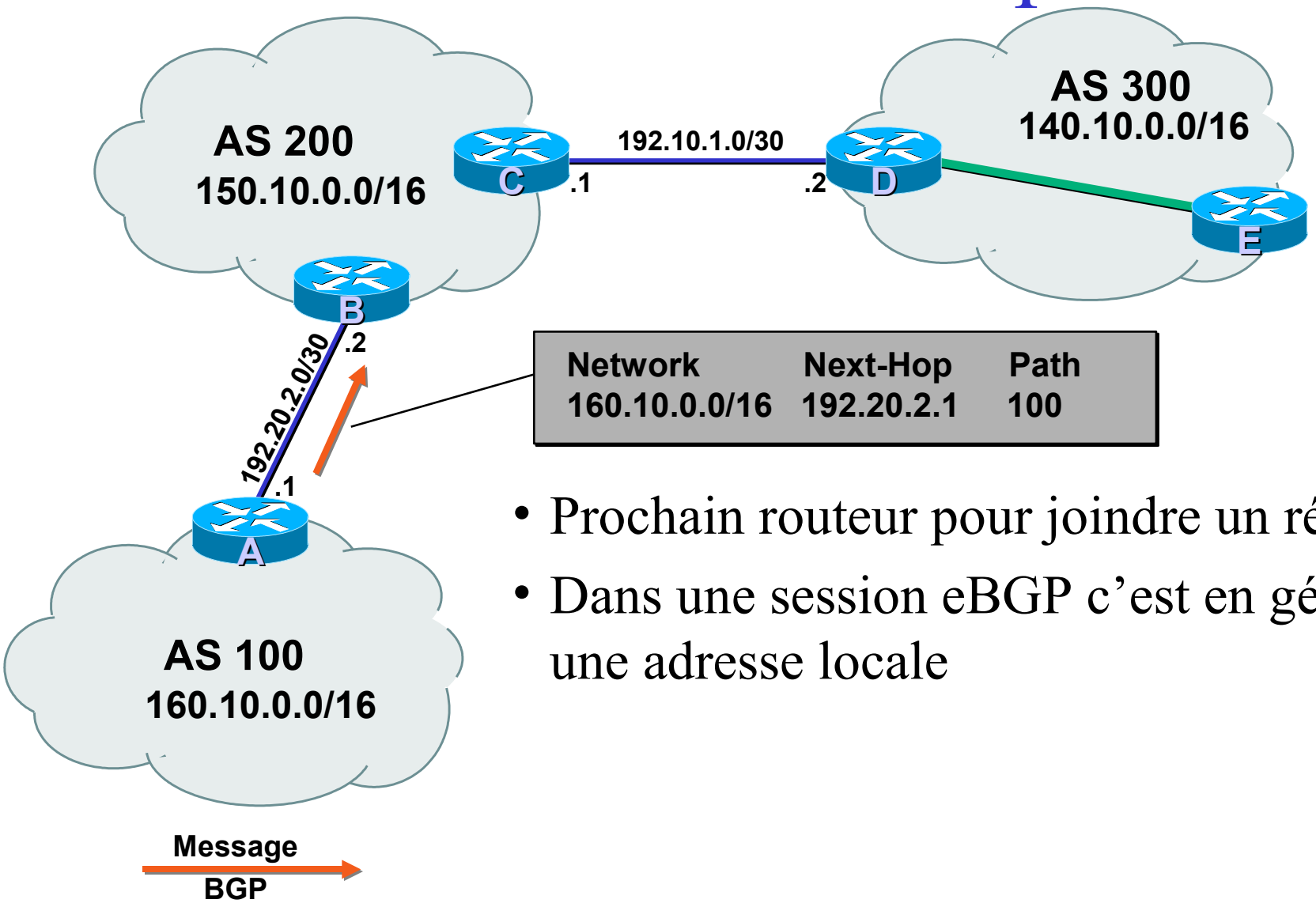
- Permet de transporter des informations liées au préfixe
  - Chemin d'AS
  - Adresse IP du "next-hop"
  - Local preference (préférence locale)
  - Multi-Exit Discriminator (MED)
  - Community (communauté)
  - Origin (origine de la route)
  - Aggregator (IP d'origine si aggrégation)

Attribut “chemin d’AS”

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

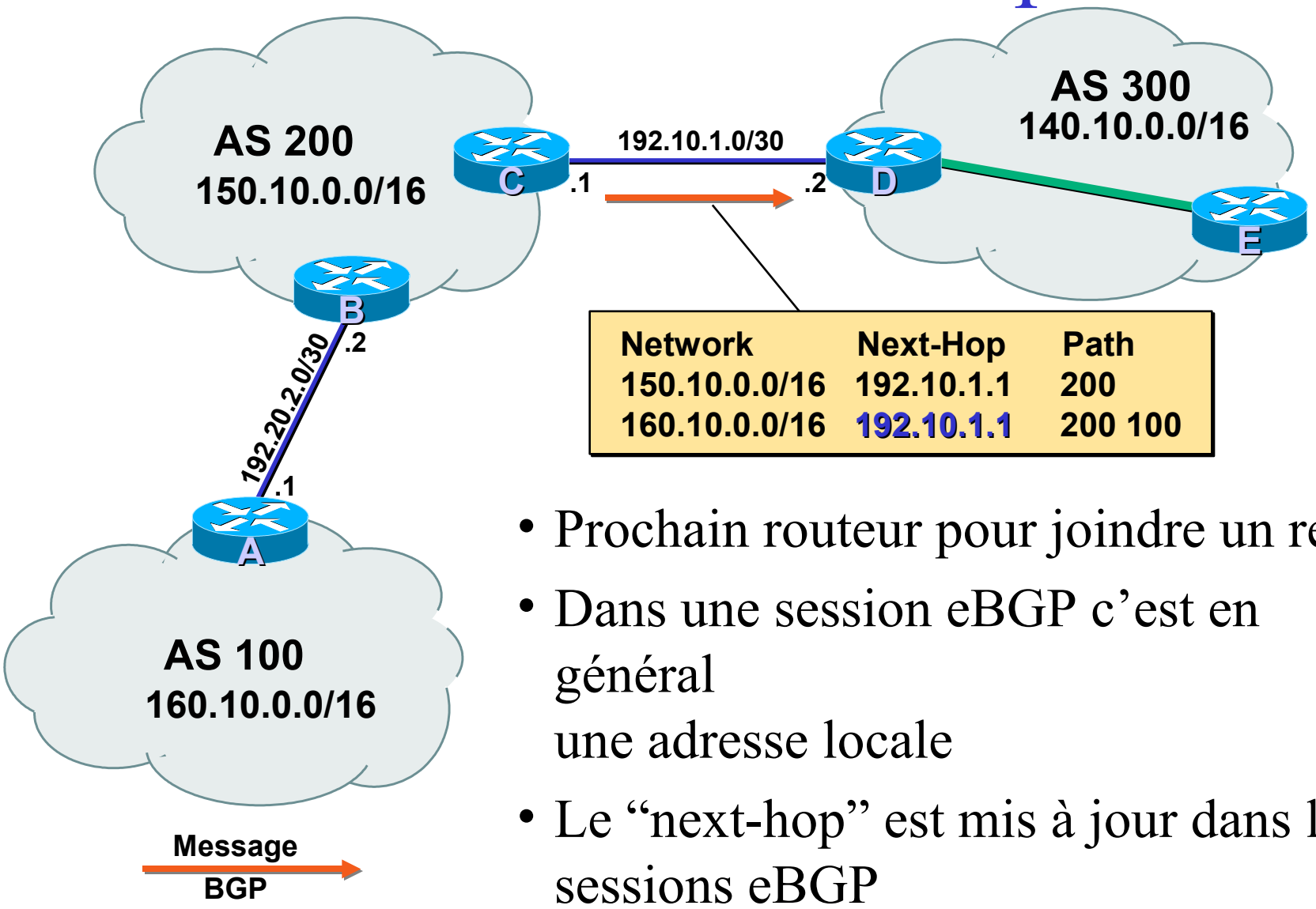


Attribut “Next-Hop”



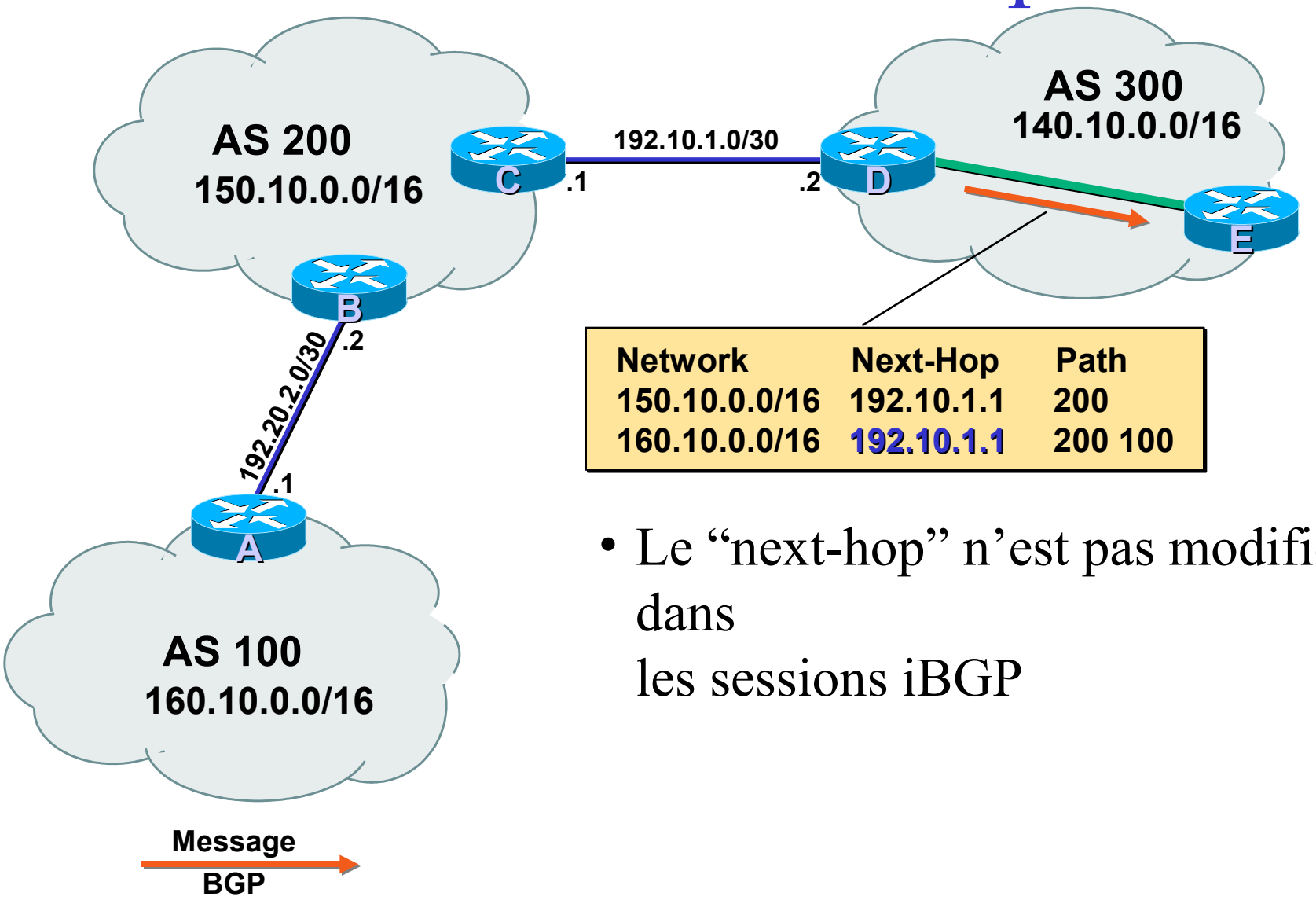
- Prochain routeur pour joindre un réseau
- Dans une session eBGP c’est en général une adresse locale

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”



- Le “next-hop” n’est pas modifié dans
  les sessions iBGP

Attribut “Next-Hop” (suite)

- Les adresses des “next-hops” doivent circuler dans l’IGP
- Recherche récursive des routes
- Permet de concevoir la topologie BGP indépendemment 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

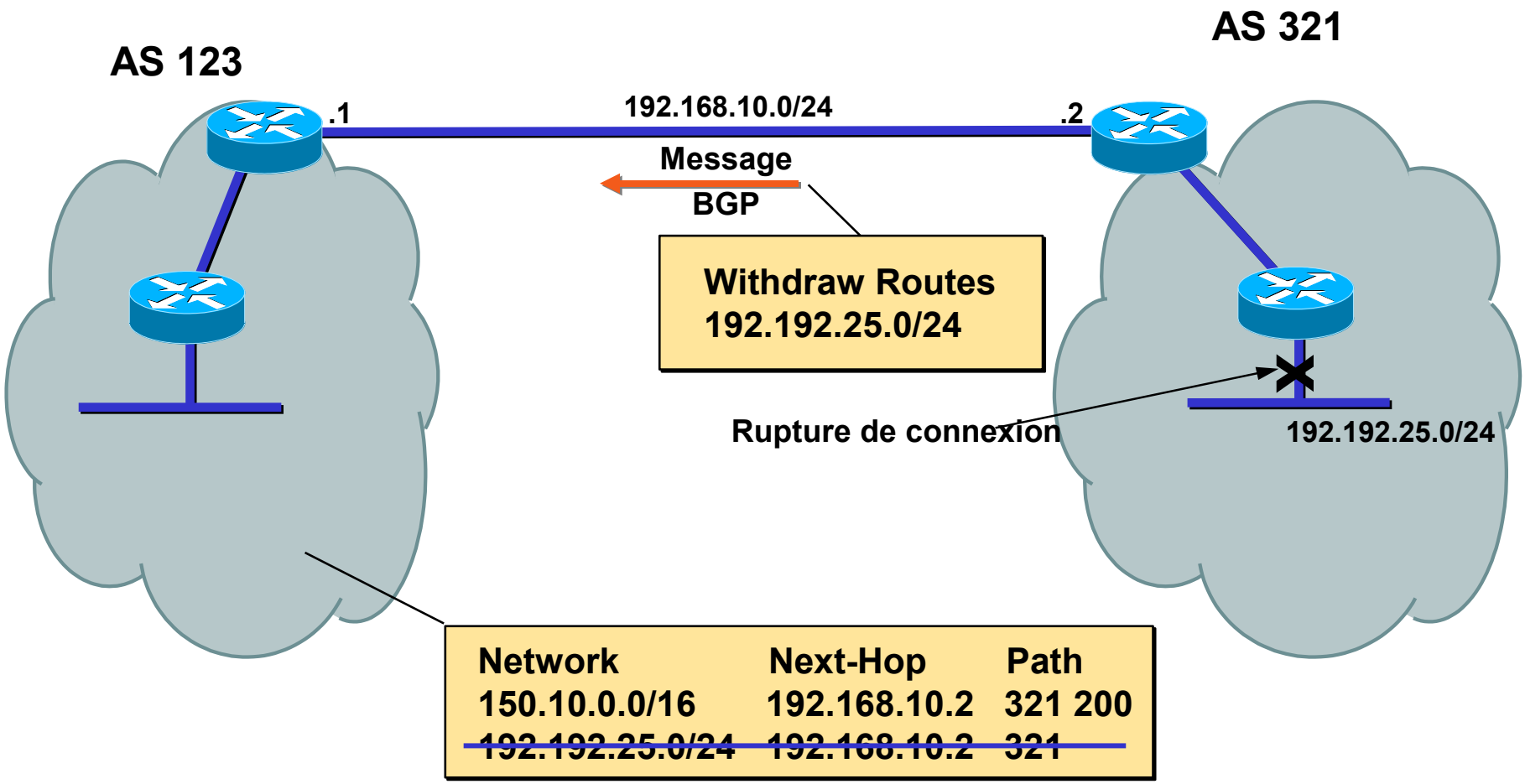
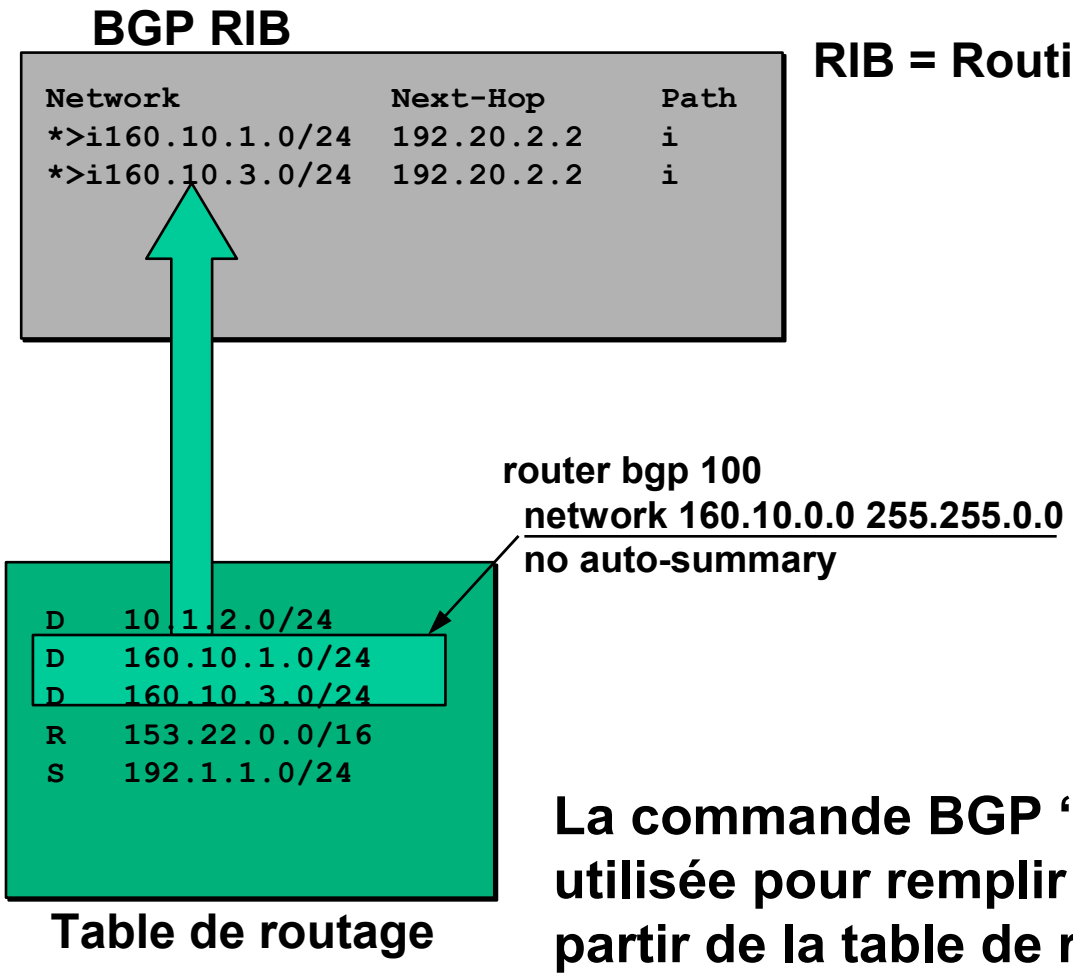
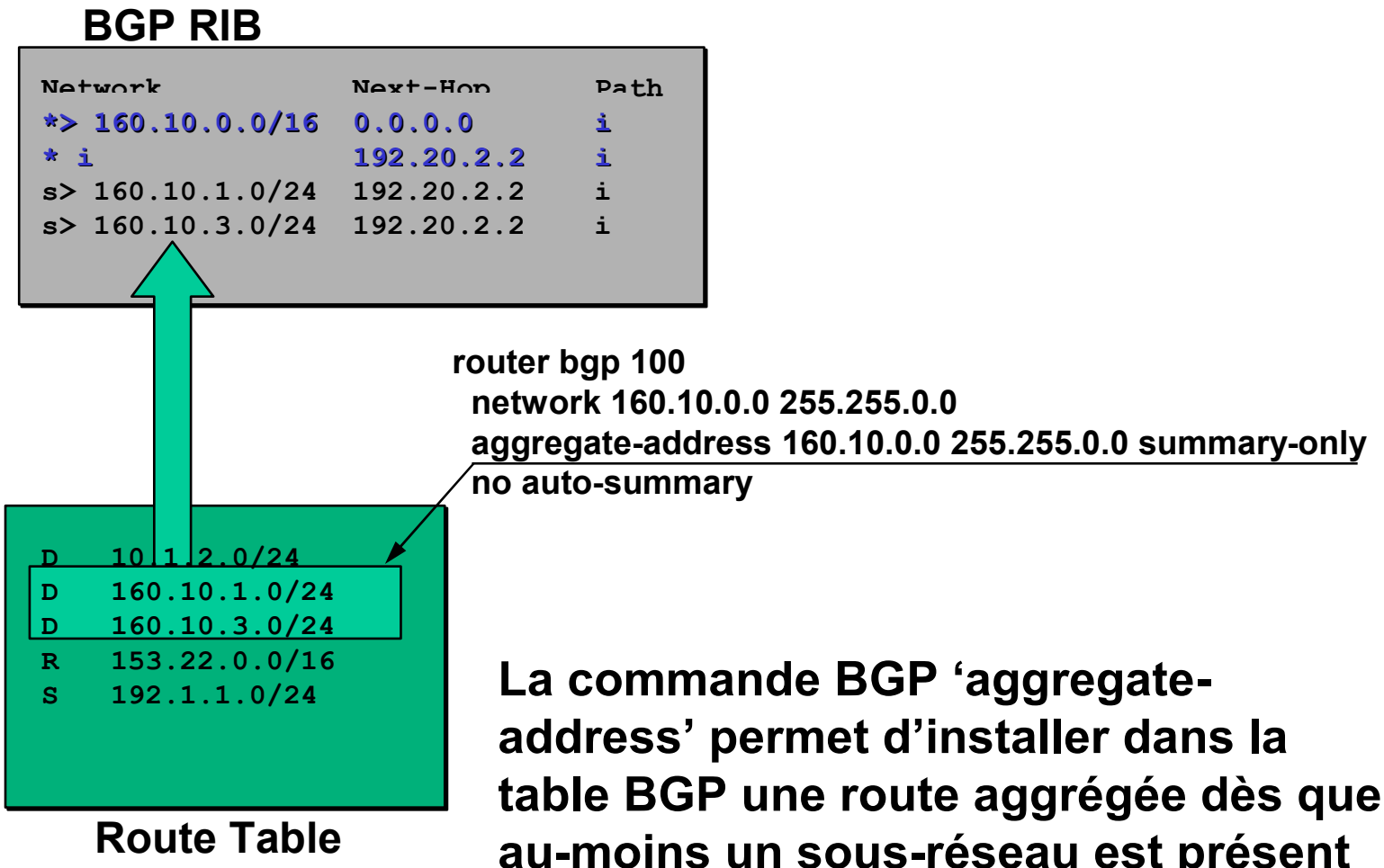


Table du routeur BGP



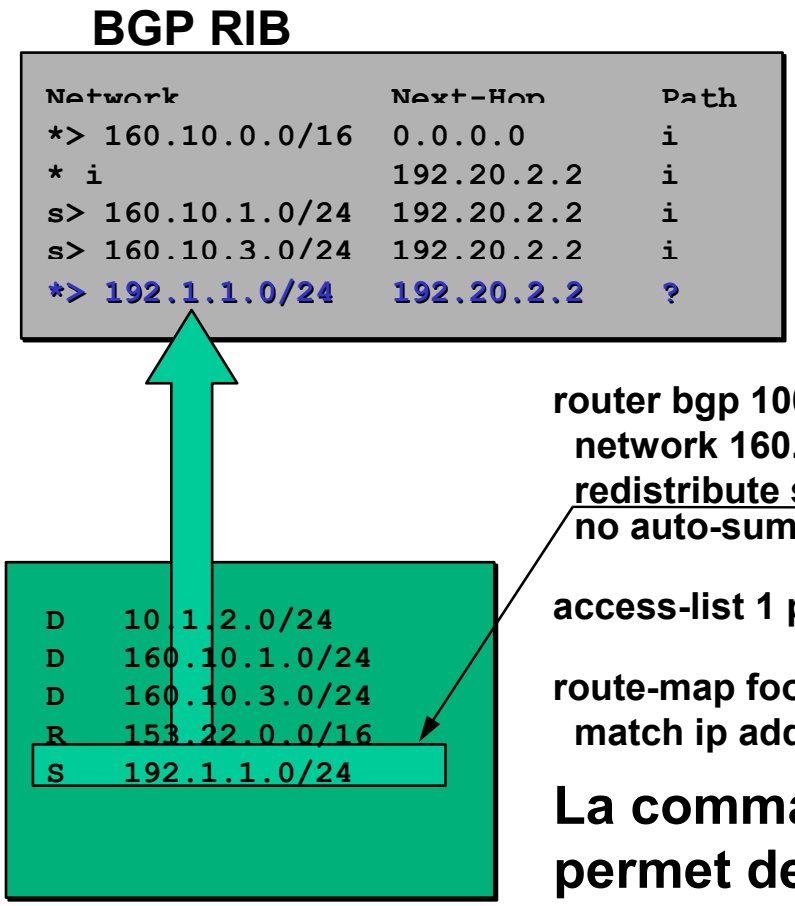
La commande BGP 'network' peut être utilisée pour remplir la table BGP à partir de la table de routage

Table du routeur BGP



La commande BGP 'aggregate-address' permet d'installer dans la table BGP une route aggrégée dès que au-moins un sous-réseau est présent

Table du routeur BGP



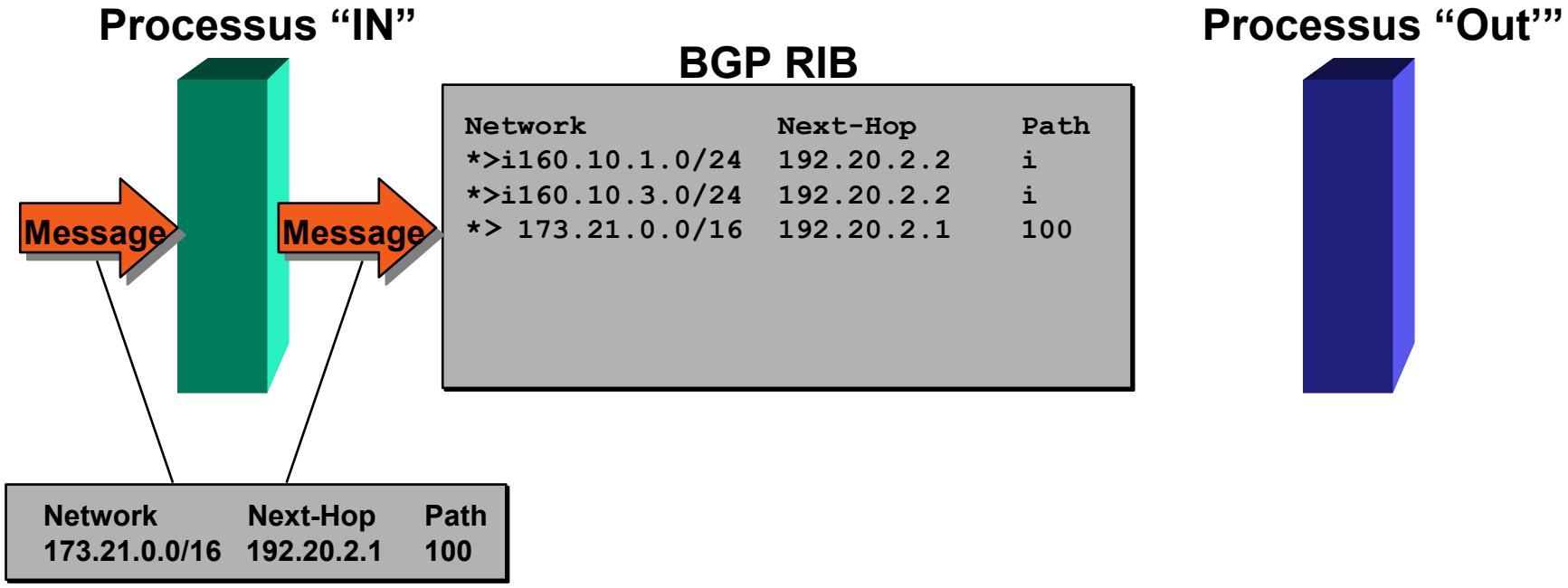

```
router bgp 100
  network 160.10.0.0 255.255.0.0
  redistribute static route-map foo
  no auto-summary

access-list 1 permit 192.1.0.0 0.0.255.255

route-map foo permit 10
  match ip address 1
```

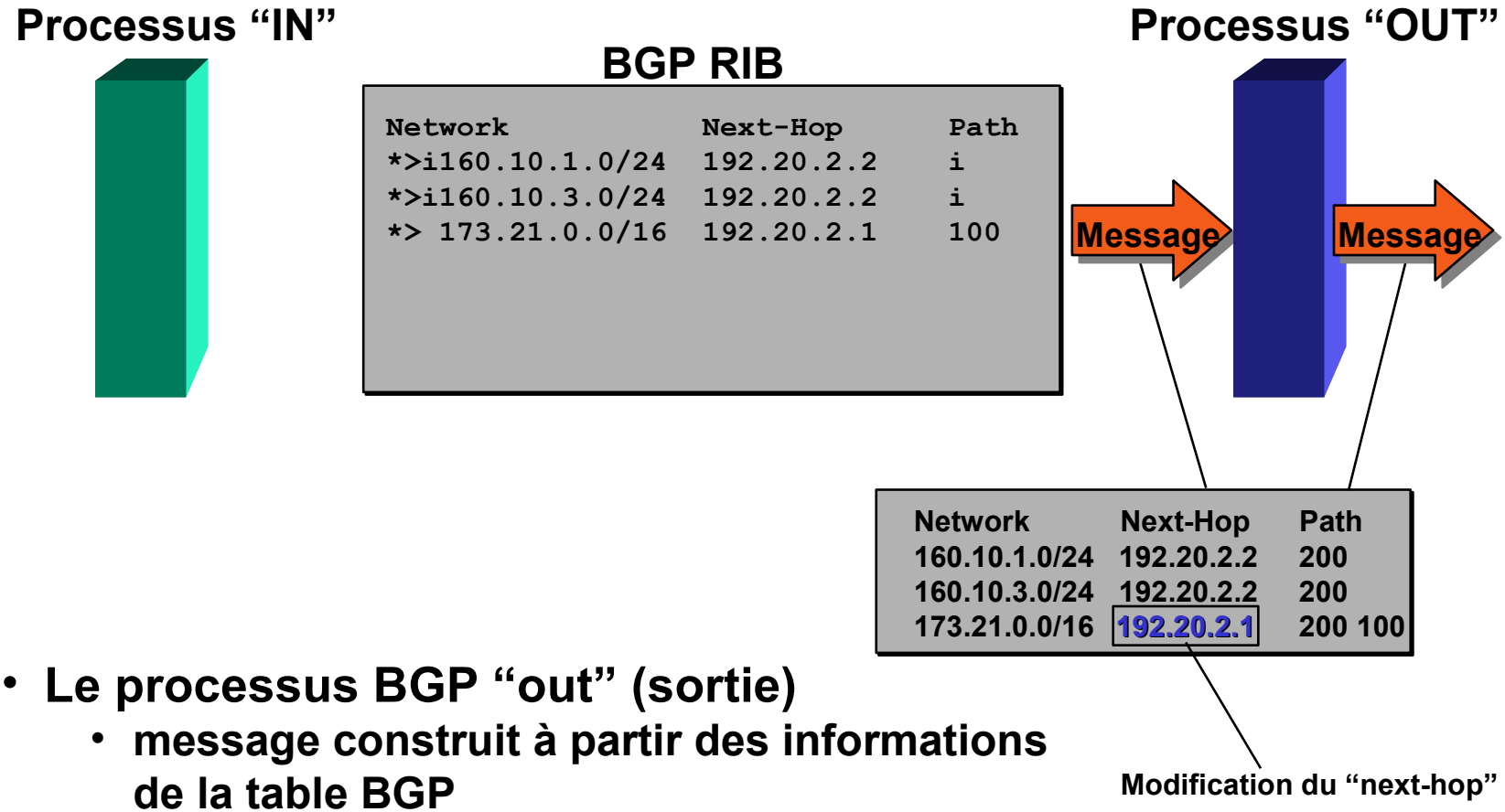
La commande BGP ‘redistribute’ permet de remplir la table BGP à partir de la table de routage en appliquant des règles spécifiques

Table du routeur BGP



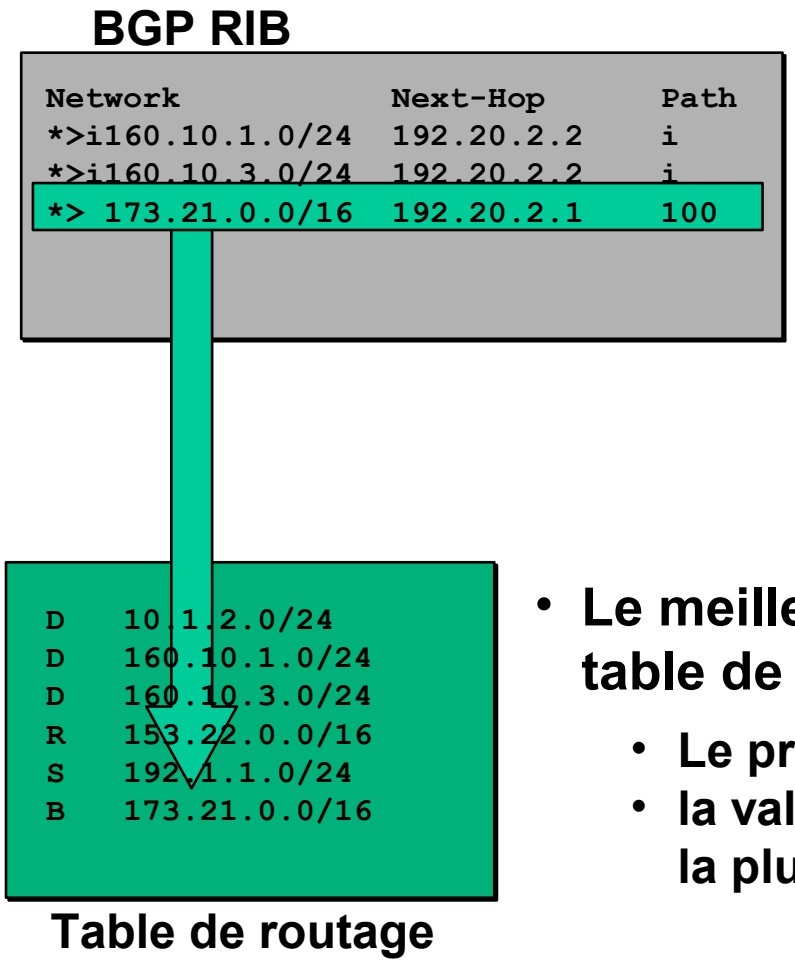
- **Le processus BGP “in” (entrée)**
  - **reçoit les messages des voisins**
  - **place le ou les chemins sélectionnés dans la table BGP**
  - **le meilleur chemin (best path) est indiqué avec le signe “>”**

Table du routeur BGP



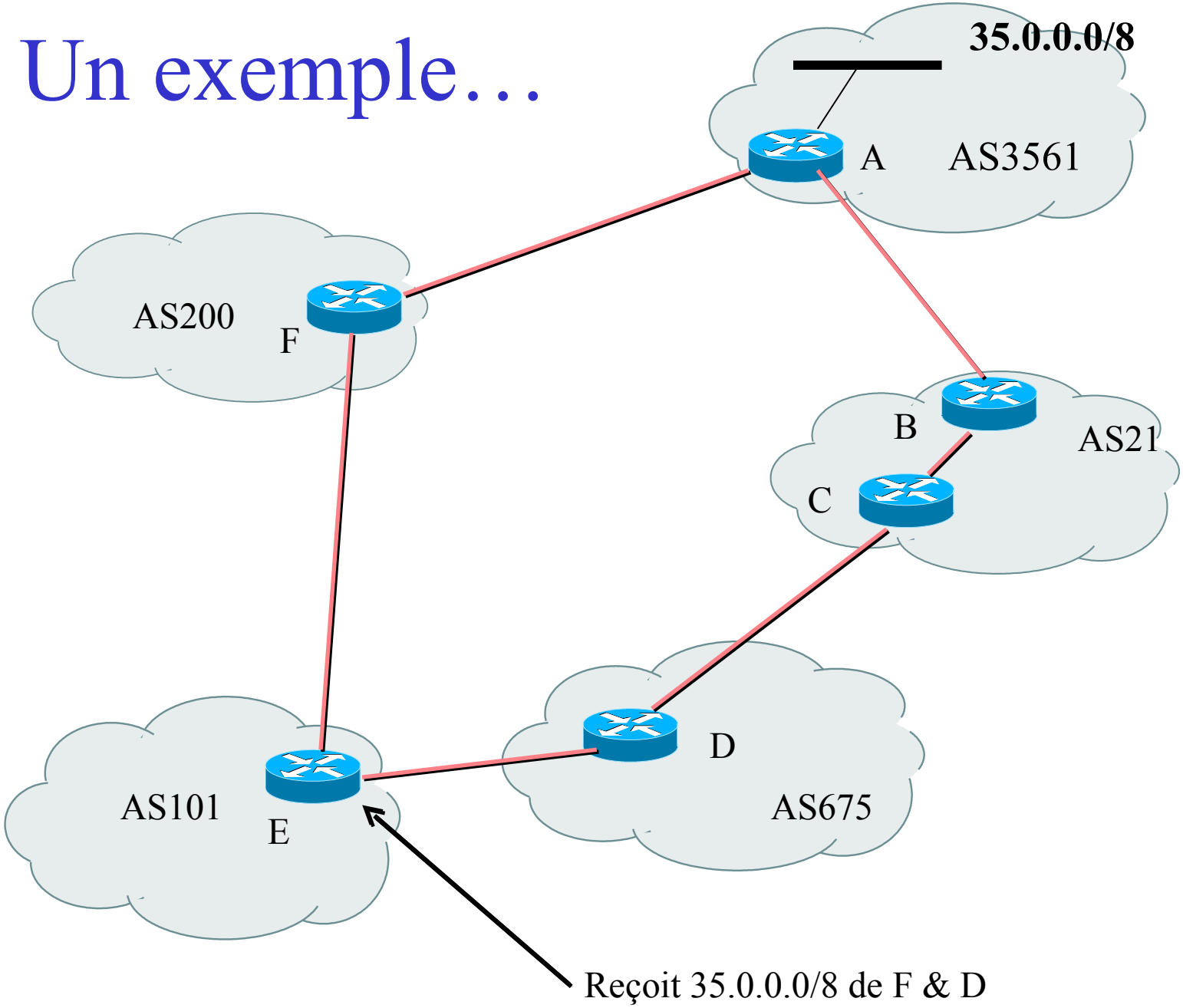
- **Le processus BGP “out” (sortie)**
  - **message construit à partir des informations de la table BGP**
  - **modification du message selon configuration**
  - **envoi du message aux voisins**

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…


35.0.0.0/8
A
AS3561
AS200
F
B
AS21
C
D
AS675
AS101
E
Reçoit 35.0.0.0/8 de F & D

Configuration de BGP

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`

Ajout de préfixes dans la table BGP

- Cela peut se faire de deux grandes manières
  - “redistribute static” (redistribuer les routes statiques)
  - utiliser la commande BGP “network”

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 controllée (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

Aggrégats et routes vers Null0

- Rappel : la route doit exister dans la table de routage pour être annoncée via BGP

```
router bgp 1
   network 198.10.0.0 mask 255.255.0.0
ip route 198.10.0.0 255.255.0.0 null0 250
```

- Une route vers “null0” est souvent utilisée pour faire de l’aggrégation
  - destination en dernier ressort pour le préfixe
  - distance de 250 pour être sûr d’être le dernier choix
- Très pratique pour la stabilité de la route
  - il ne peut y avoir de “flap” !

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

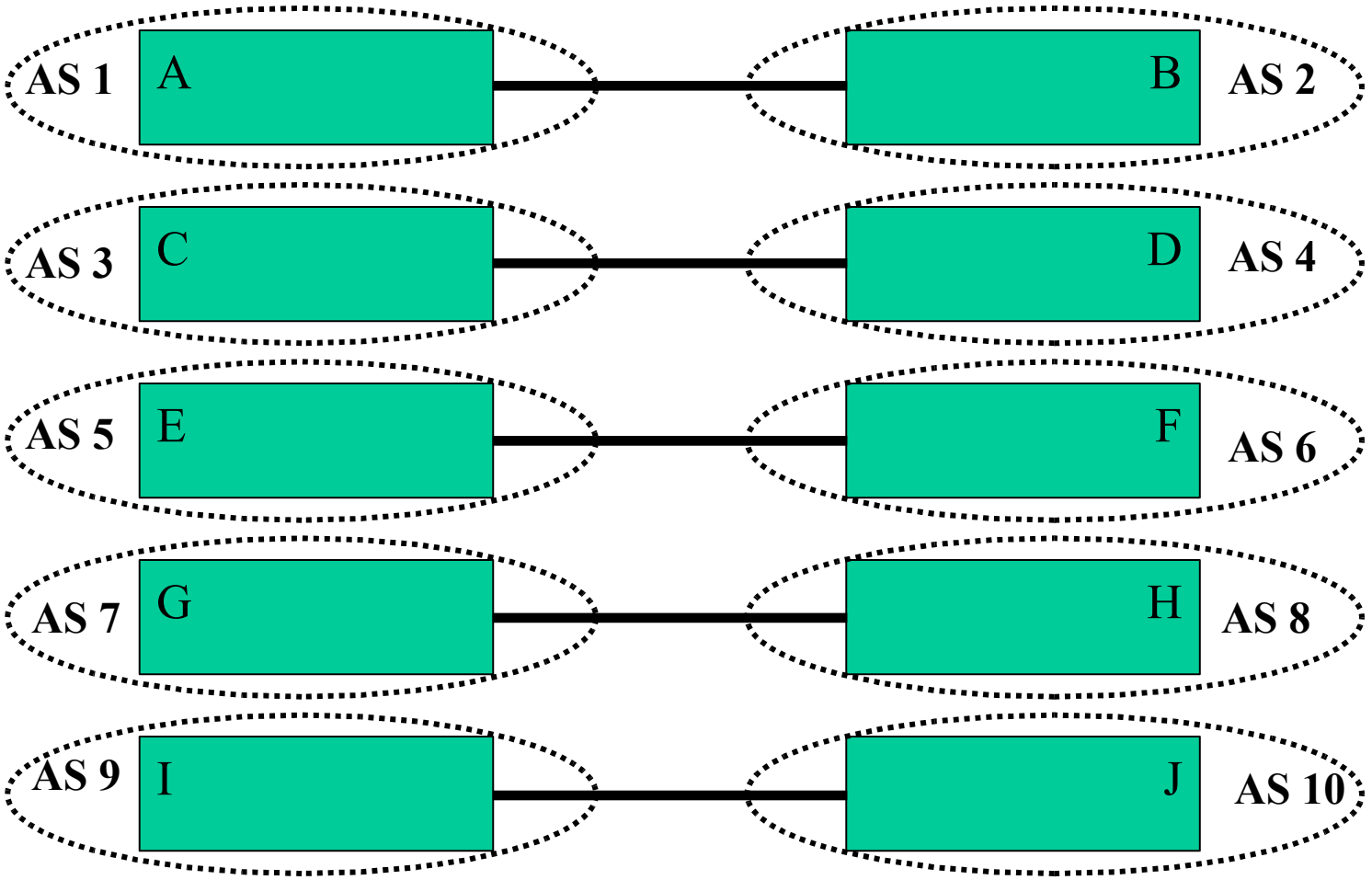
```
router  ospf <ID>
network <loopback-address> 0.0.0.0
```

- 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
```

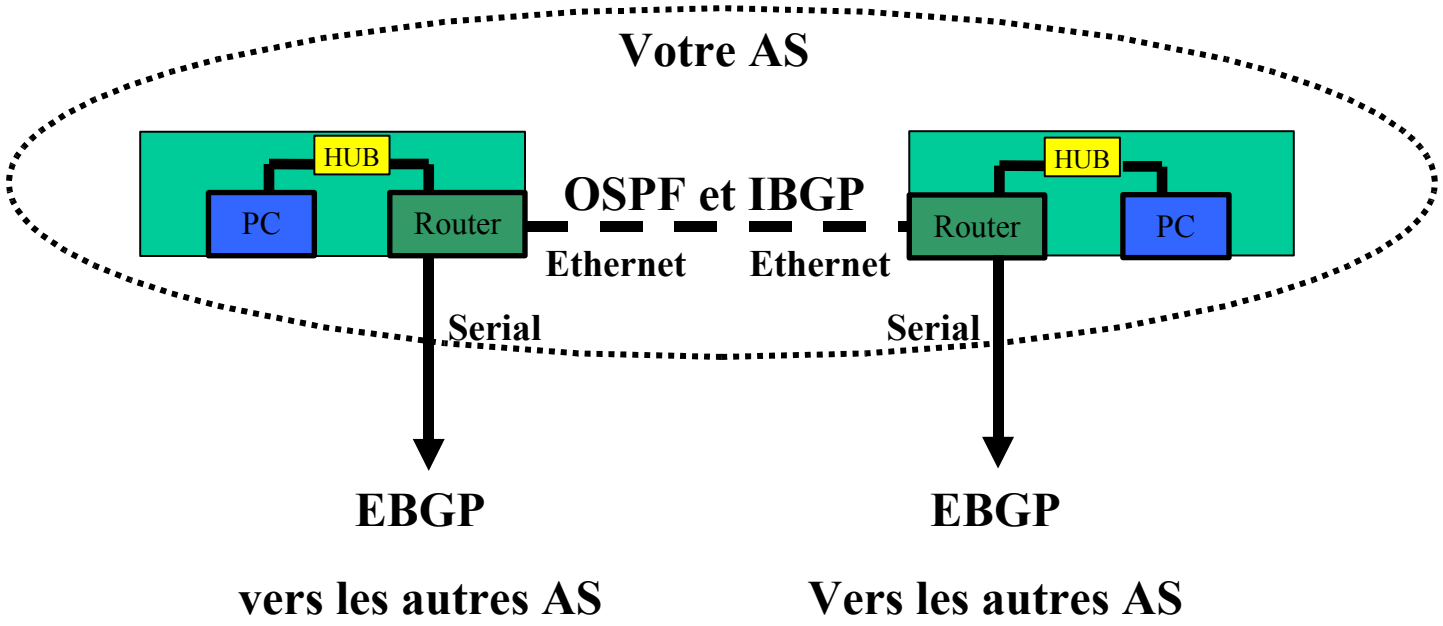
Exercice 1 - Configuration de BGP

Liste des sessions et numéros d'AS

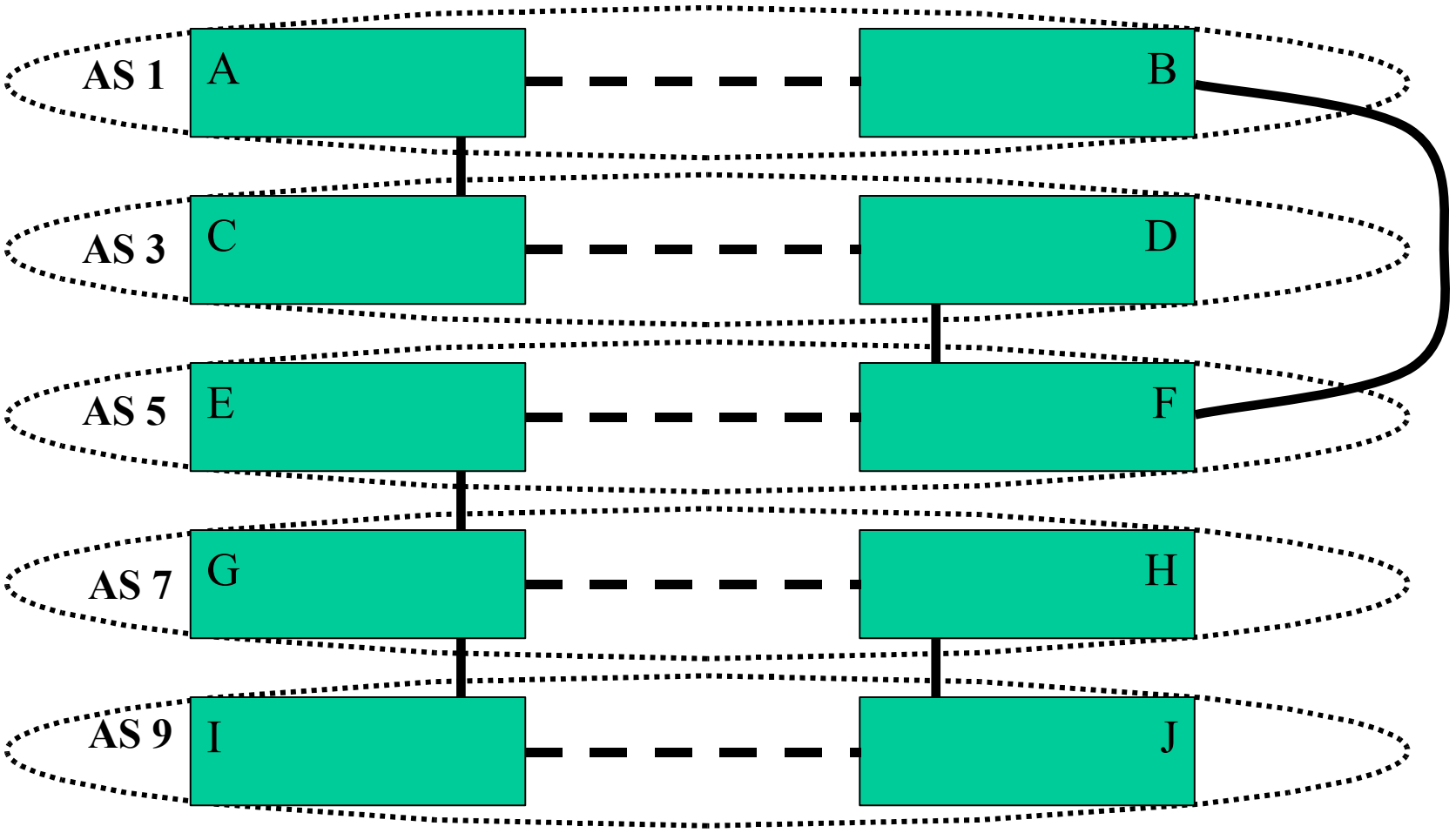


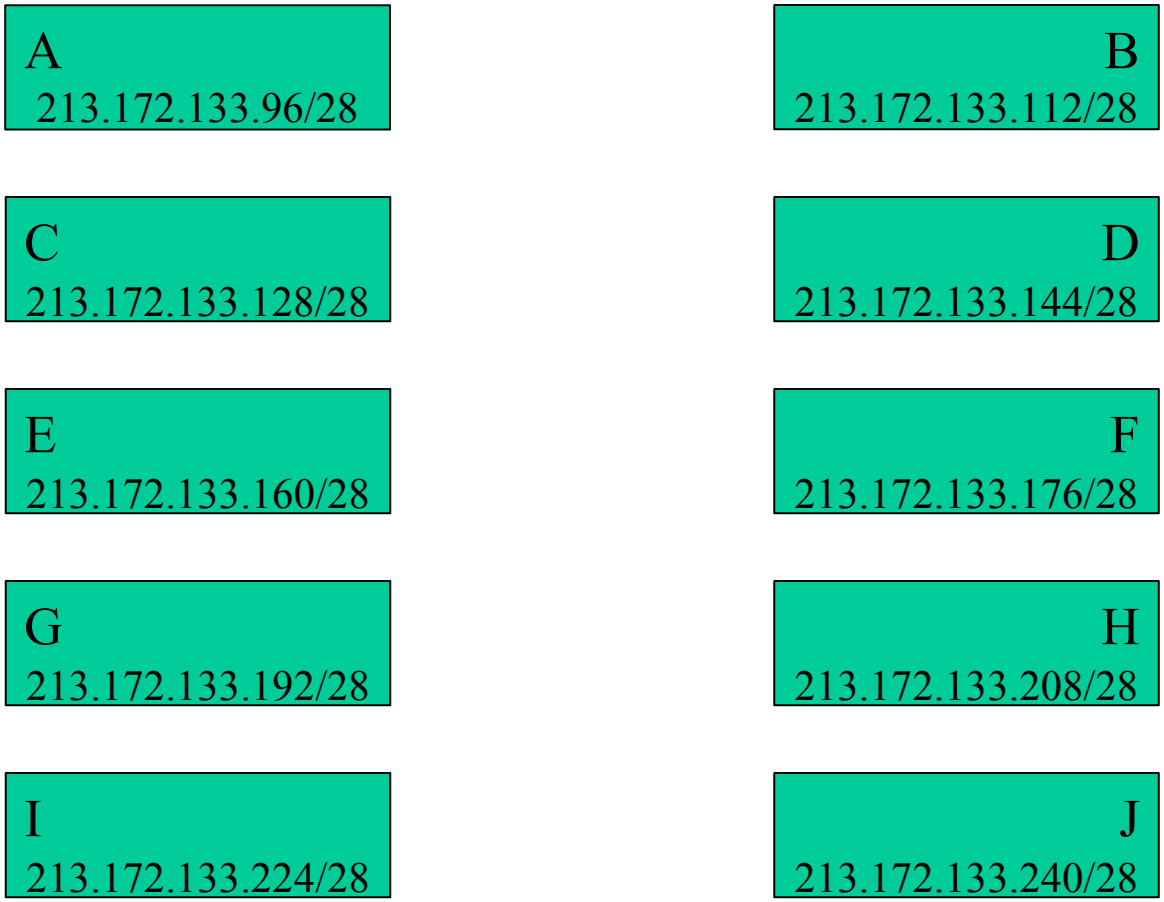
Session eBGP

Exercice 2 - Configuration de eBGP et iBGP


Votre AS
HUB
PC
Router
OSPF et IBGP
Ethernet
Ethernet
HUB
Router
PC
Serial
Serial
EBGP
EBGP
vers les autres AS
Vers les autres AS

Liste des sessions et numéros d'AS




A
213.172.133.96/28
B
213.172.133.112/28
C
213.172.133.128/28
D
213.172.133.144/28
E
213.172.133.160/28
F
213.172.133.176/28
G
213.172.133.192/28
H
213.172.133.208/28
I
213.172.133.224/28
J
213.172.133.240/28

BGP 4, suite...

Attributs de chemin BGP

- Encodés sous la forme d'un triplet Type, Longueur & Valeur (TLV)
- Attributs Transitifs ou non transitif
- Certains attributs sont obligatoires
- Ils sont utilisés pour choisir le meilleur chemin
- Ils permettent d'appliquer des règles d'ingéniérie du trafic (routage politique)

Liste des attributs de chemins BGP

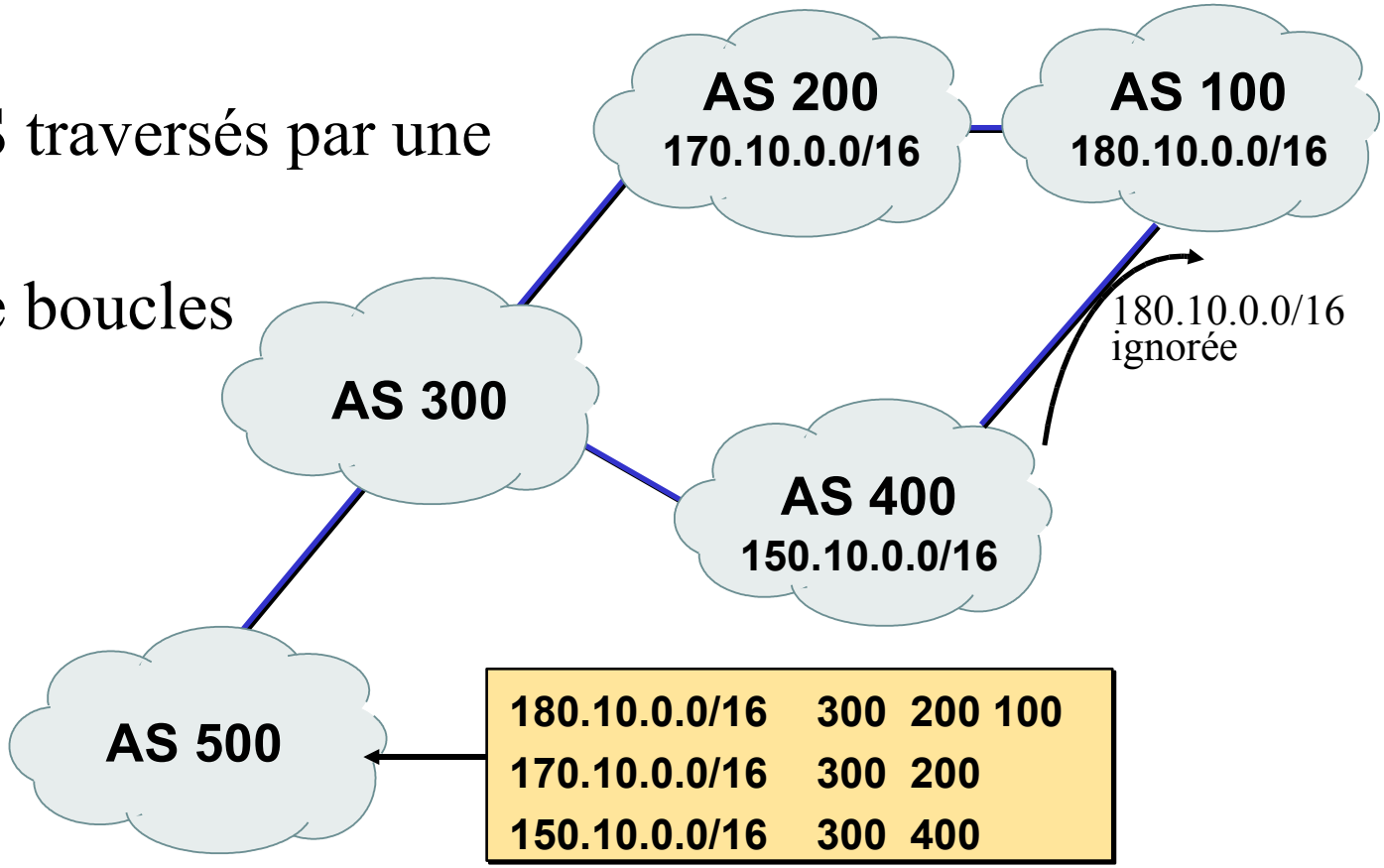
- Origine
- AS-path (chemin d'AS)
- Next-hop (prochain routeur)
- Multi-Exit Discriminator (MED)
- Local preference (préférence locale)
- BGP Community (communauté BGP)
- Autres...

AS-PATH (chemin d'AS)

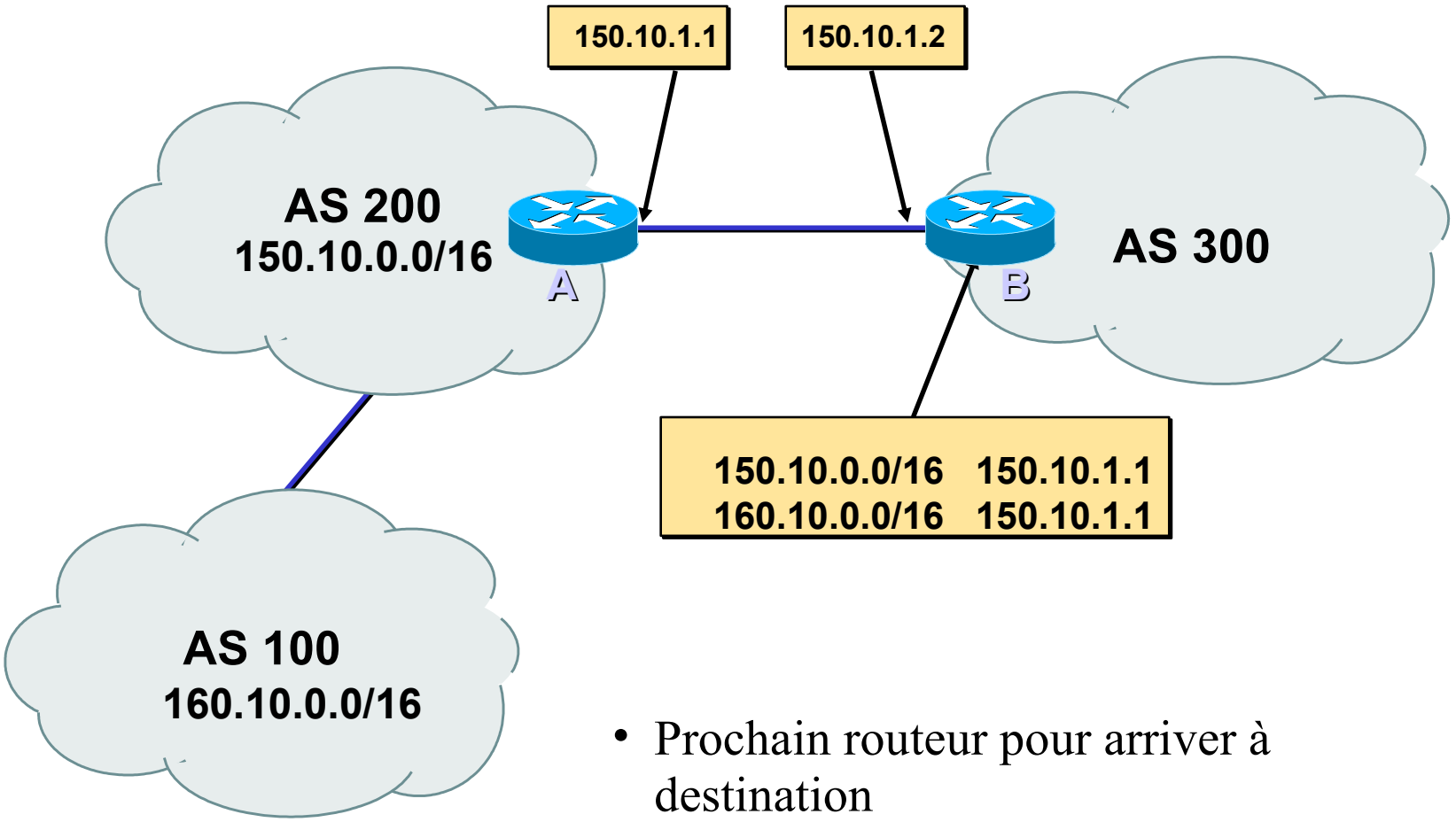
- Attribut mis à jour par le routeur envoyant un message BGP, en y ajoutant son propre numéro d'AS
- Contient la liste des AS traversés par le message
- Permet de détecter des boucles de routage
  - Une mise à jour reçue est ignorée si elle contient son propre numéro d'AS

AS-Path (chemin d'AS)

- Liste des AS traversés par une route
- Détection de boucles

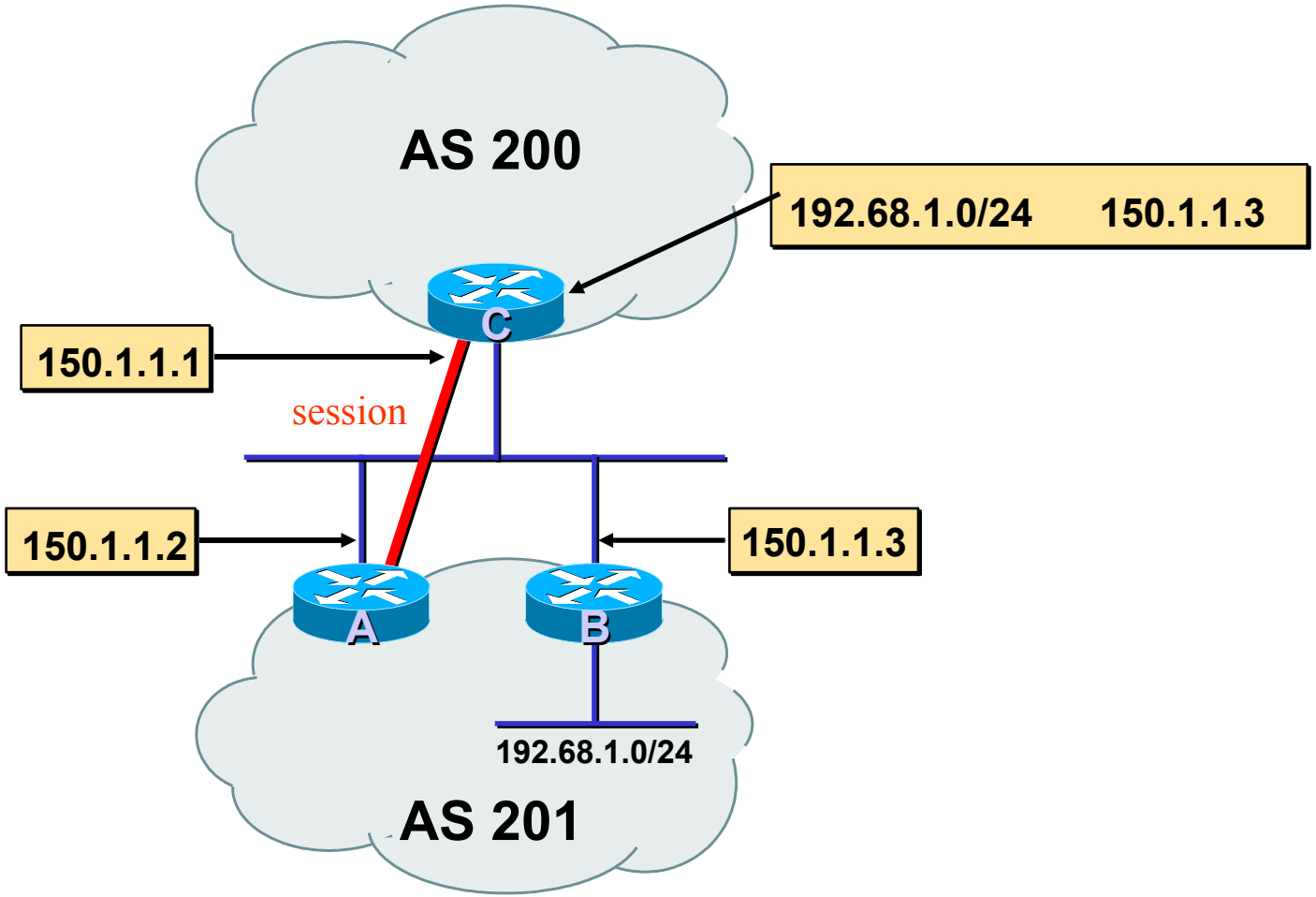


Next-Hop (prochain routeur)



- Prochain routeur pour arriver à destination
- Adresse de routeur ou de voisin en eBGP
- Non modifié en iBGP

Next-Hop sur un réseau tiers



- Serait plus efficace, mais c'est une mauvaise idée !

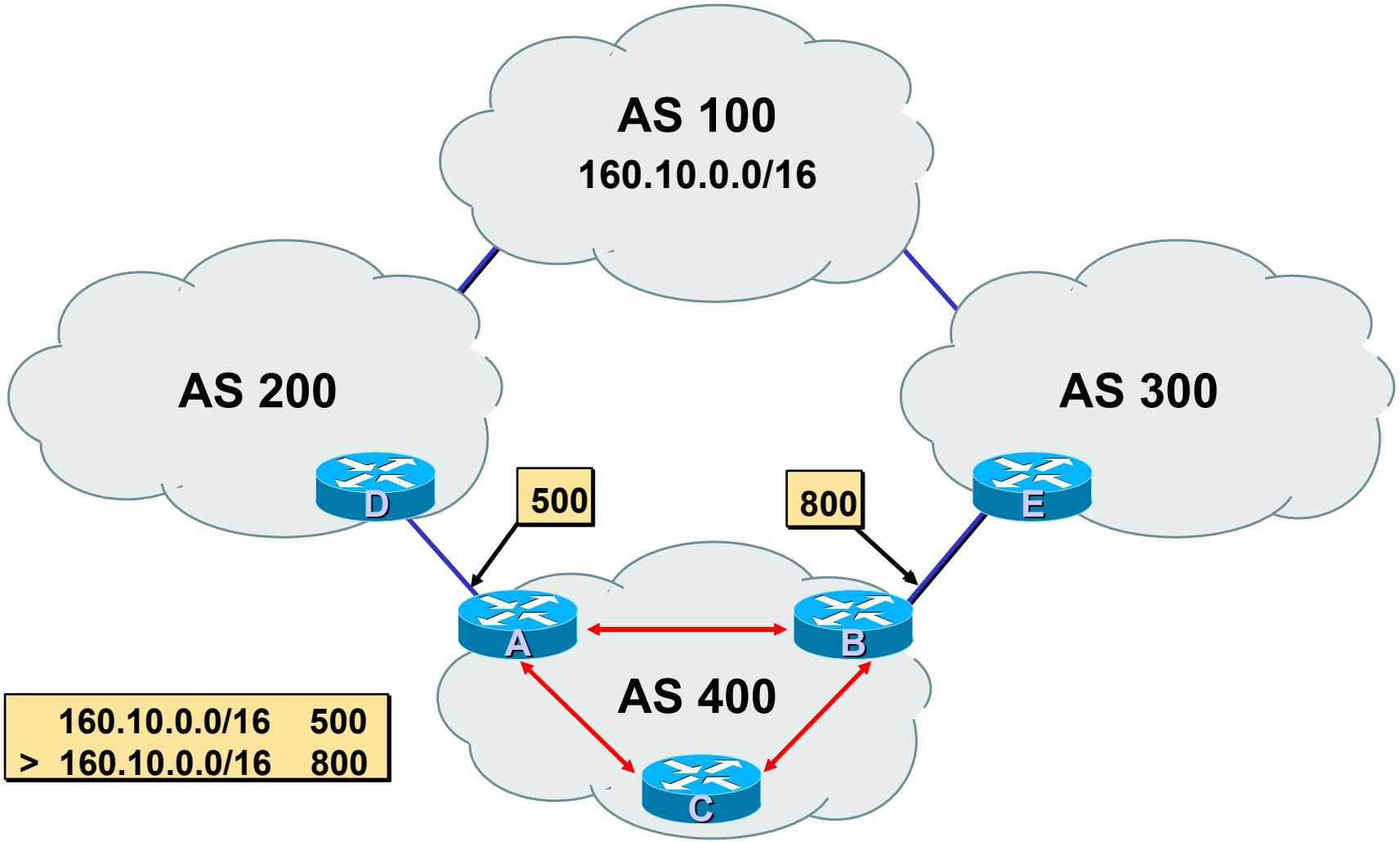
Next-Hop suite...

- Les routes vers l'ensemble des adresses de next-hop sont à transporter dans l'IGP
- Recherche de route récursive dans les tables
- BGP n'est plus lié à la topologie du réseau
- Les bonnes décisions de routage sont prises par le protocole IGP

Local Preference (préfence locale)

- Obligatoire pour iBGP, non utilisé dans eBGP
- Valeur par défaut chez Cisco : 100
- Paramètre local à un AS
- Permet de préférer une sortie à une autre
- Le chemin avec la préférence locale la plus élevée est sélectionné

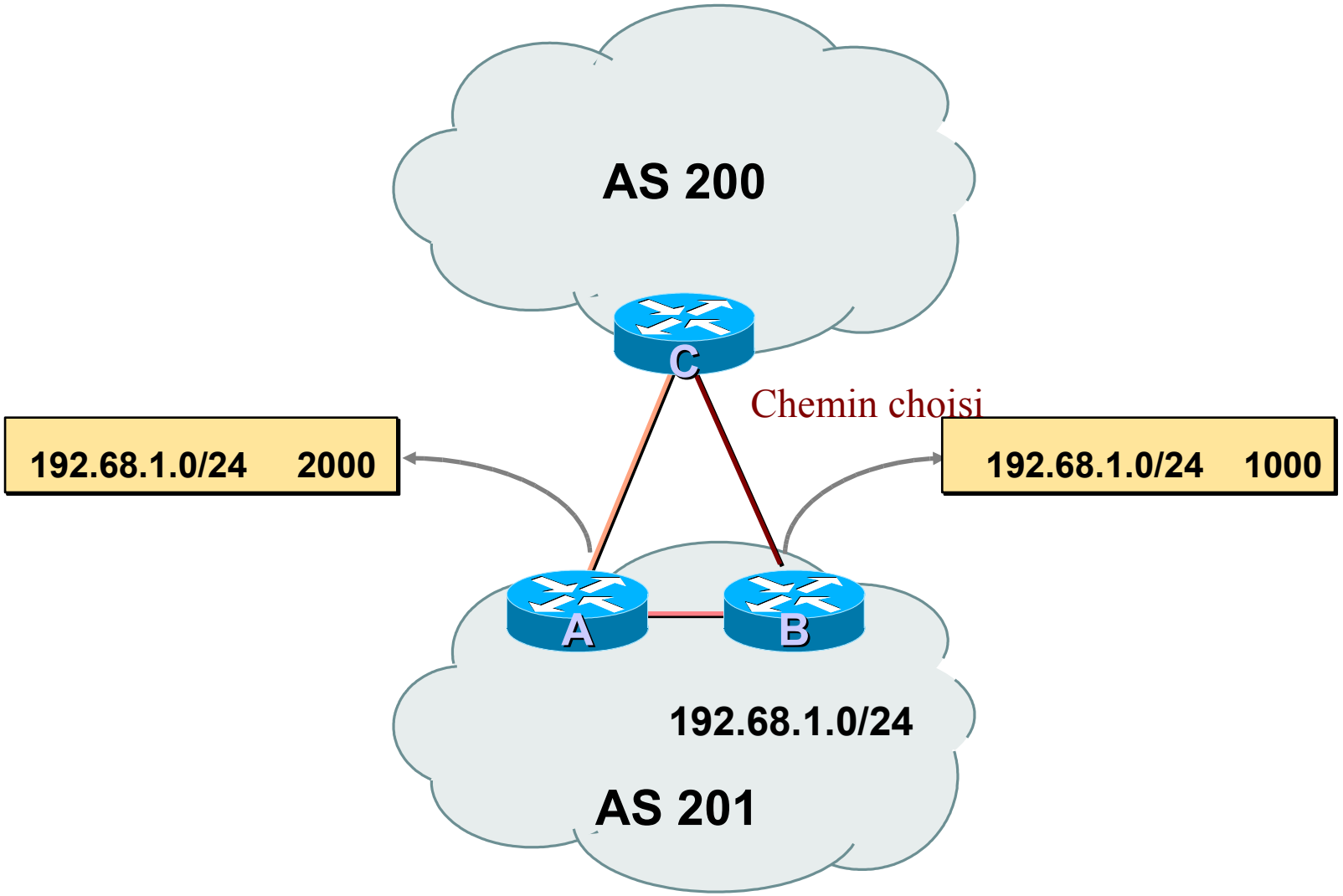
Local Preference (préférence locale)



Multi-Exit Discriminator

- Attribut non transitif
- Valeur numérique (0-0xffffffff)
- Permet de transporter des préférences relatives entre points de sortie
- Si les chemins viennent du même AS le MED peut être utilisé pour comparer les routes
- Le chemin avec le plus petit MED est sélectionné
- Le métrique IGP peut être choisi comme MED

Multi-Exit Discriminator (MED)



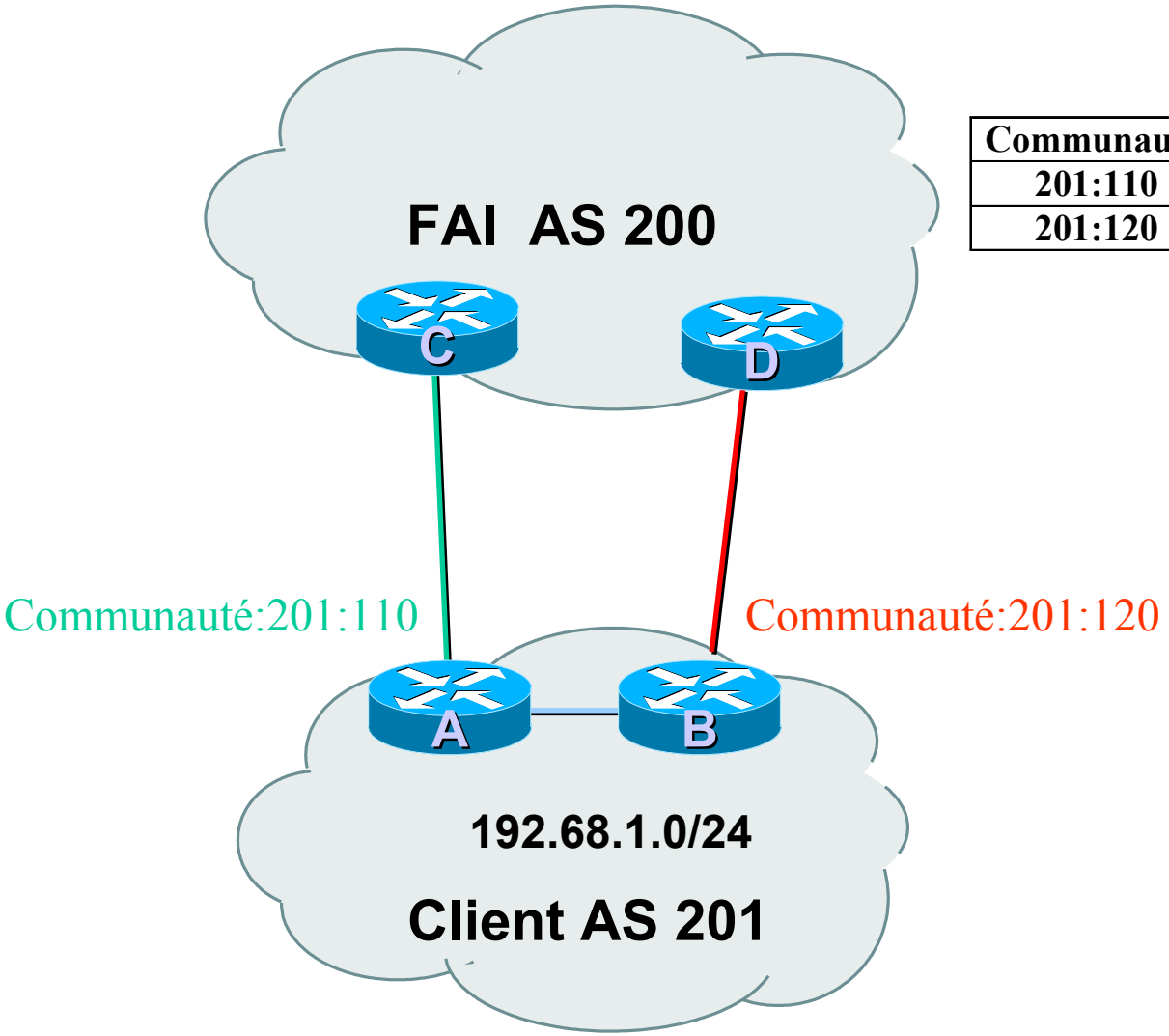
Origin (Origine de la route)

- Indique l'origine du préfixe
- Trois valeurs
  - IGP - préfixe obtenu avec une clause "network"
    - exemple : *network 35.0.0.0*
  - EGP - Redistribué par un EGP
  - Incomplete - Redistribué par un IGP
    - exemple : *redistribute ospf*
- IGP < EGP < INCOMPLETE

Communautés BGP

- Transitives, attribut facultatif
- Valeur numérique (0-0xffffffff)
- Permettent de créer des groupes de destinations
- Chaque destination peut appartenir à plusieurs communautés
- Attribut très flexible, car il permet de faire des choix avec des critères inter ou intra-AS

Communautés BGP



Communauté	Local Preference
201:110	110
201:120	120

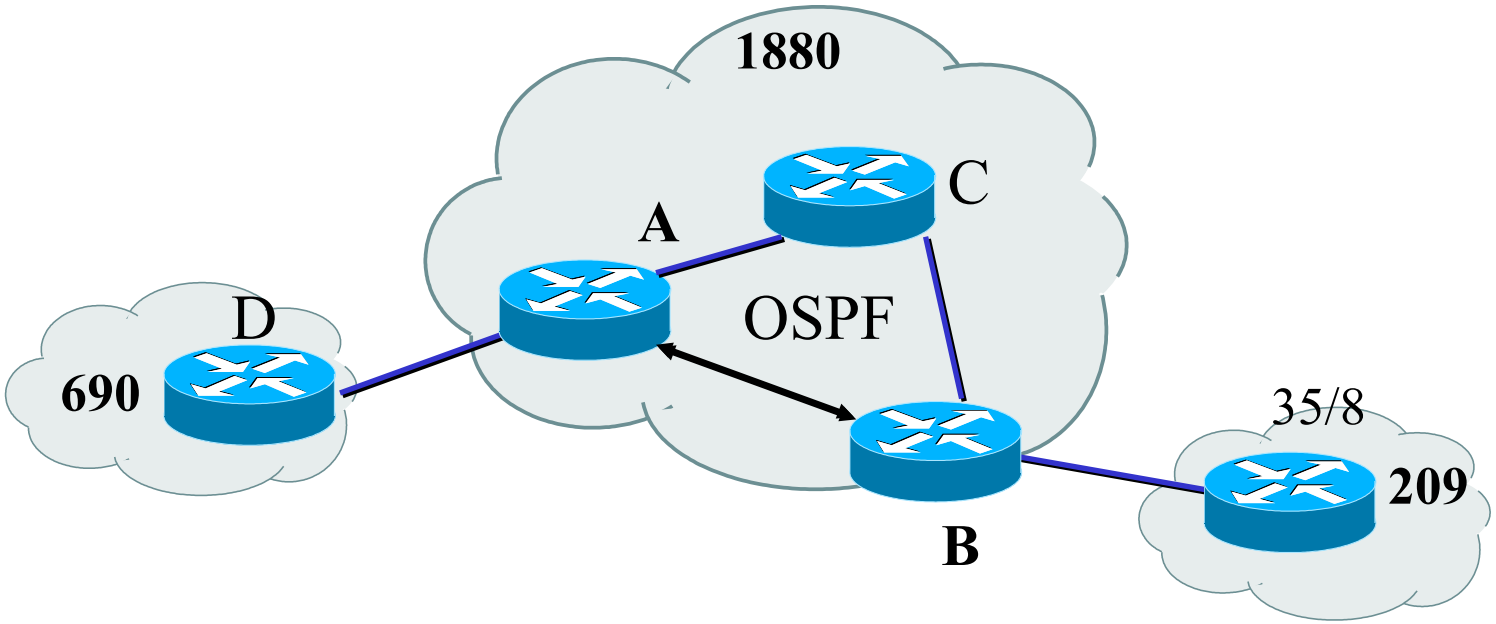
Poids (Weight)

- Attribut spécifique Cisco utilisé lorsqu'il y a plus d'une route vers la même destination
- Attribut local à un routeur (non propagé ailleurs)
- Valeur par défaut 32768 pour les chemins dont l'origine est le routeur et 0 pour les autres
- Lorsqu'il y a plusieurs choix, on préferra la route dont le poids est le plus élevé.

Distance administrative

- Les routes peuvent être apprises par plusieurs protocoles de routage
  - il faut les classifier pour faire un choix
- La route issue du protocole avec la plus faible distance est installée dans la table de routage
- Distances par défaut en BGP:
  - local (routes provenant du routeur) : 200
  - eBGP : 20, iBGP : 200
- Cela n'a pas d'impact dans l'algorithme de choix des chemins BGP, mais il y a un impact quand à installer ou pas une route BGP dans la table de routage IP

Synchronization (synchronisation)



- C ne tourne pas BGP (non-pervasive BGP)
- A n'annoncera pas 35/8 à D tant qu'il ne l'aura pas appris par l'IGP
- Il faut désactiver la synchronisation pour éviter ce problème

  router bgp 1880

  no sync

Synchronization (synchronisation)

- Spécifique IOS Cisco : BGP n'annoncera pas une route avant que l'ensemble des routeurs de l'AS ne l'ait apprise par un IGP
- Désactiver la synchronisation si :
  - Votre AS ne sert pas d'AS de transit, ou
  - Tous les routeurs de transit tournent BGP, or
  - iBGP est utilisé sur le cœur de réseau (backbone)

Sélection d'une route BGP (bestpath)

Il ne peut y avoir qu'un seul meilleur chemin ! (sauf multipath)

- La route doit être synchronisée

  C'est à dire être dans la table de routage

- Le "Next-hop" doit être joignable

  Il se trouve dans la table de routage

- Prendre la valeur la plus élevée pour le poids (weight)

  Critère spécifique Cisco et local au routeur

- Choisir la préférence locale la plus élevée

  Appliqué pour l'ensemble des routeurs de l'AS

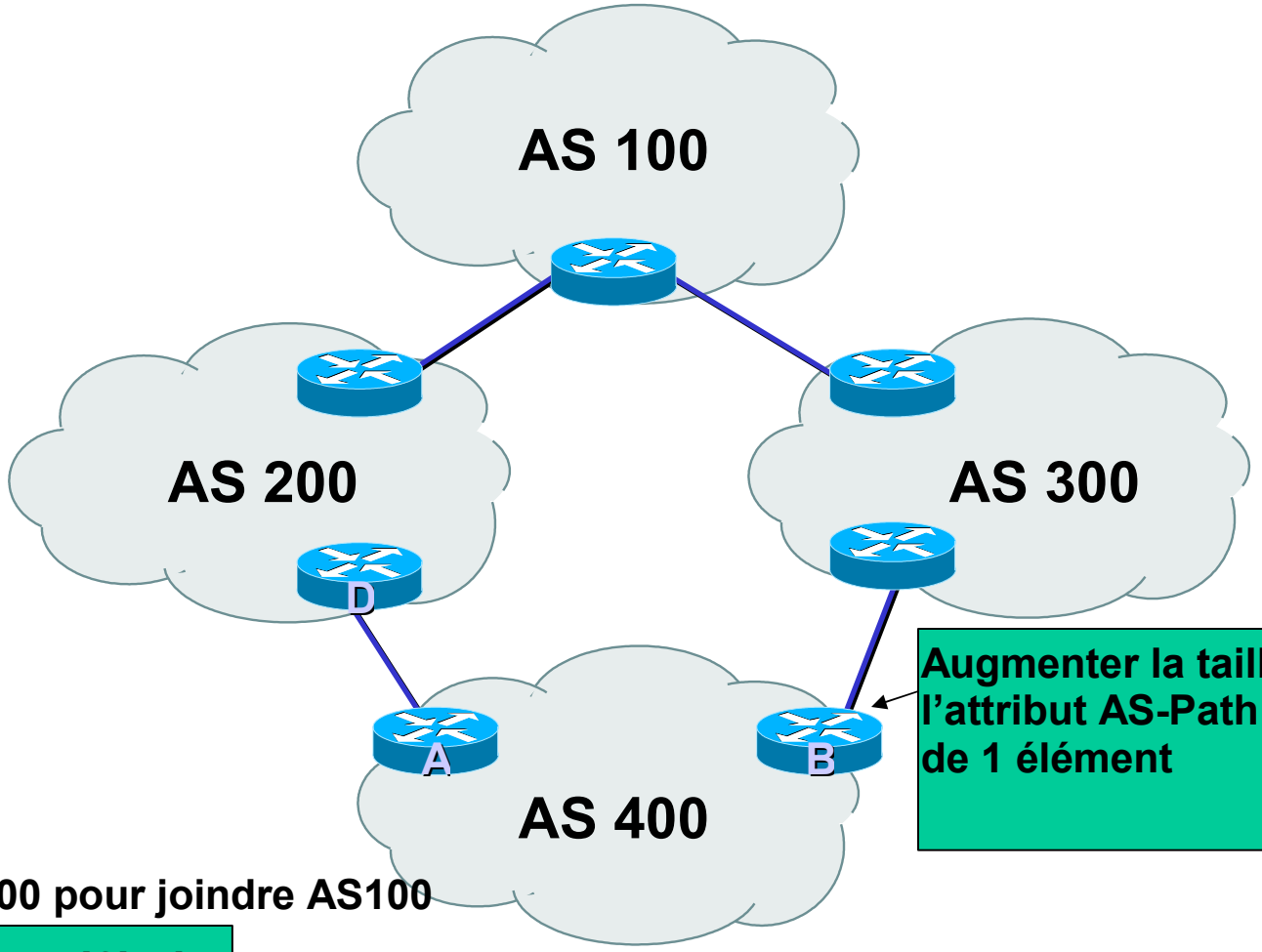
- La route est d'origine locale

  Via une commande BGP "redistribute" ou "network"

Sélection d'une route BGP ...

- Choisir le plus court chemin d'AS
  - en comptant le nombre d'AS dans l'attribut AS-Path
- Prendre l'origine de valeur la plus faible
  - IGP < EGP < INCOMPLETE
- Choisir le plus petit MED
  - pour des chemins en provenance d'un même AS
- Préférer une route Externe sur une route Interne
  - prendre la sortie la plus proche
- Choisir le "next-hop" le plus proche
  - Plus faible métrique IGP, donc plus proche de la sortie de l'AS
- Plus petit "Router-ID"
- Adresse IP du voisin la plus petite

Sélection d'une route BGP...



Augmenter la taille de l'attribut AS-Path au-moins de 1 élément

Politique AS 400 pour joindre AS100

AS 200 est préférré

AS 300 en secours

Politique de routage - Liste de préfixes, Route Maps et Listes de distribution (distribute lists)

Politique de routage

- Pourquoi ?
  - Pour envoyer le trafic vers des routes choisies
  - Filtrage de préfixes en entrée et sortie
  - Pour forcer le respect des accords Client-ISP
- Comment ?
  - Filtrage basé sur les AS - filter list
  - Filtrage basé sur les préfixes - distribute list
  - Modification d'attributs BGP - route maps

Politique - Liste de préfixes

- Filtrage par voisin
  - c'est une configuration incrémentielle
- Access-list utilisées très performantes
- Fonctionne en entrée comme en sortie
- Basé sur les numéros de réseaux (adressage IPv4 réseau/maque)
- Un "deny" est implicite à la fin de la liste

Liste de préfixes - Exemples

- Ne pas accepter la route par défaut
  - `ip prefix-list Exemple deny 0.0.0.0/0`
- Autoriser le préfixe 35.0.0.0/8
  - `ip prefix-list Exemple permit 35.0.0.0/8`
- Interdire le préfixe 172.16.0.0/12
  - `ip prefix-list Exemple deny 172.16.0.0/12`
- Dans 192/8 autoriser jusqu'au /24
  - `ip prefix-list Exemple permit 192.0.0.0/8 le 24`
  - Ceci autorisera toute route dans 192.0.0.0/8, sauf les /25, /26, /27, /28, /29, /30, /31 and /32

Listes de préfixes - Exemples 2

- Dans 192/8 interdire /25 et au-delà
  - `ip prefix-list Exemple deny 192.0.0.0/8 ge 25`
  - Ceci interdit les préfixes de taille /25, /26, /27, /28, /29, /30, /31 and /32 dans le bloc 192.0.0.0/8
  - Très ressemblant au précédent exemple
- Dans 192/8 autoriser les préfixes entre /12 et /20
  - `ip prefix-list Exemple permit 192.0.0.0/8 ge 12 le 20`
  - Ceci interdit les préfixes de taille /8, /9, /10, /11, /21, /22 et au-delà dans le bloc 192.0.0.0/8
- Autoriser tous les préfixes
  - `ip prefix-list Exemple 0.0.0.0/0 le 32`

Utilisation des listes de préfixes

- Exemple de configuration

```
router bgp 200
  network 215.7.0.0
  neighbor 220.200.1.1 remote-as 210
  neighbor 220.200.1.1 prefix-list PEER-IN in
  neighbor 220.200.1.1 prefix-list PEER-OUT out
!
ip prefix-list PEER-IN deny 218.10.0.0/16
ip prefix-list PEER-IN permit 0.0.0.0/0 le 32
ip prefix-list PEER-OUT permit 215.7.0.0/16
ip prefix-list PEER-OUT deny 0.0.0.0/0 le 32
```

Tout accepter du voisin, sauf nos réseaux

Envoyer uniquement nos réseaux au voisin

Distribute list - avec des ACL IP

```
access-list 1 deny 10.0.0.0
access-list 1 permit any
access-list 2 permit 20.0.0.0
```

… il faut créer des ACL avec l'ajout de nouveaux préfixes ...

```
router bgp 100
  neighbor 171.69.233.33 remote-as 33
  neighbor 171.69.233.33 distribute-list 1 in
  neighbor 171.69.233.33 distribute-list 2 out
```

Filtrage avec des expression régulières

- L'expression régulière décrit la forme que doit avoir l'argument
- Est utilisé pour comparer l'attribut AS-Path
- Exemple : ^3561.*100.*1$
- Grande fléxibilité qui permet de générer des expression complexes

Filtrage avec des expressions régulières

```
ip as-path access-list 1 permit 3561
ip as-path access-list 2 deny 35
ip as-path access-list 2 permit .*

router bgp 100
   neighbor 171.69.233.33 remote-as 33
   neighbor 171.69.233.33 filter-list 1 in
   neighbor 171.69.233.33 filter-list 2 out
```

Accepter les routes d'origine AS 3561. Tout le reste est rejeté en entrée ("deny" implicite).

Ne pas annoncer les routes de l'AS 35, mais tout le reste est envoyé (en sortie).

Route Maps

```
router bgp 300
neighbor 2.2.2.2 remote-as 100
neighbor 2.2.2.2 route-map SETCOMMUNITY out
!
route-map SETCOMMUNITY permit 10
match ip address 1
match community 1
set community 300:100
!
access-list 1 permit 35.0.0.0
ip community-list 1 permit 100:200
```

Route-map : clauses match & set

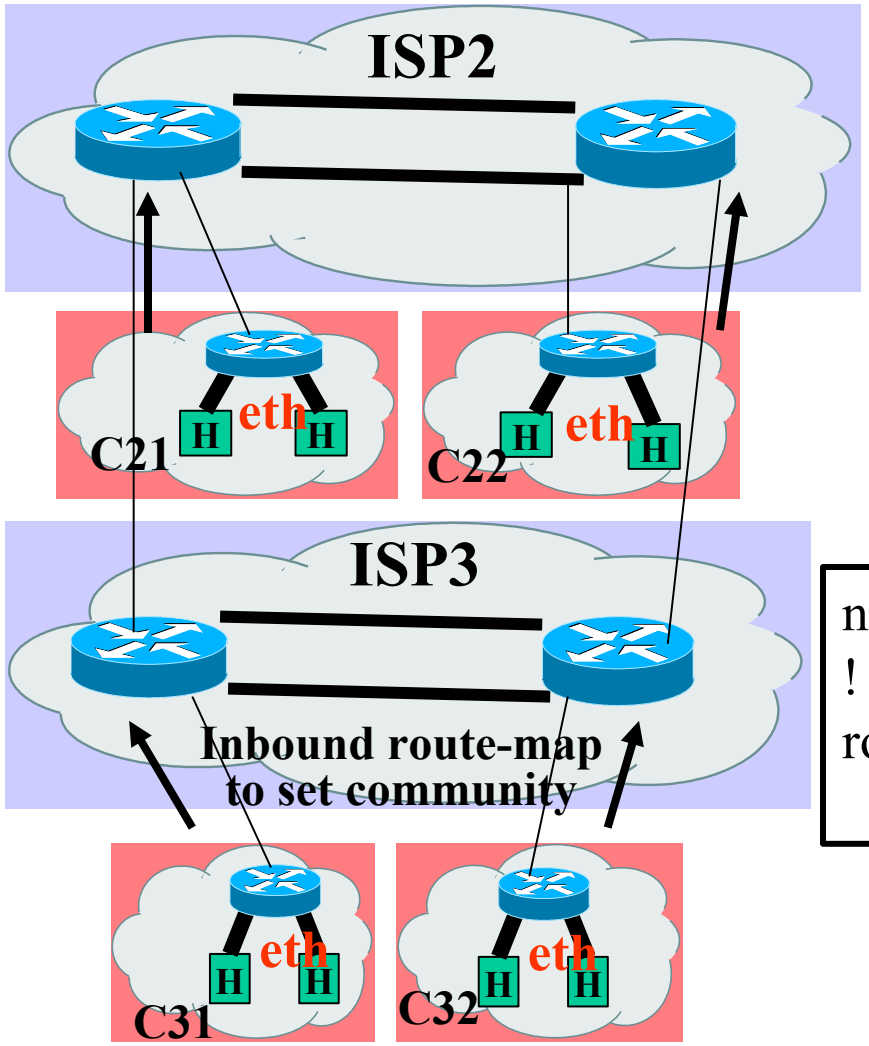
Match Clauses

- AS-path
- Community
- IP address

Set Clauses

- AS-path prepend
- Community
- Local-Preference
- MED
- Origin
- Weight
- Autres...

Exemple de configuration avecRoute-map




```
neighbor <y.y.y.y> route-map AS200_IN in
!
route-map AS200_IN permit 10
            match community 1
            set local-preference 200
!
ip community-list 1 permit 100:200
```

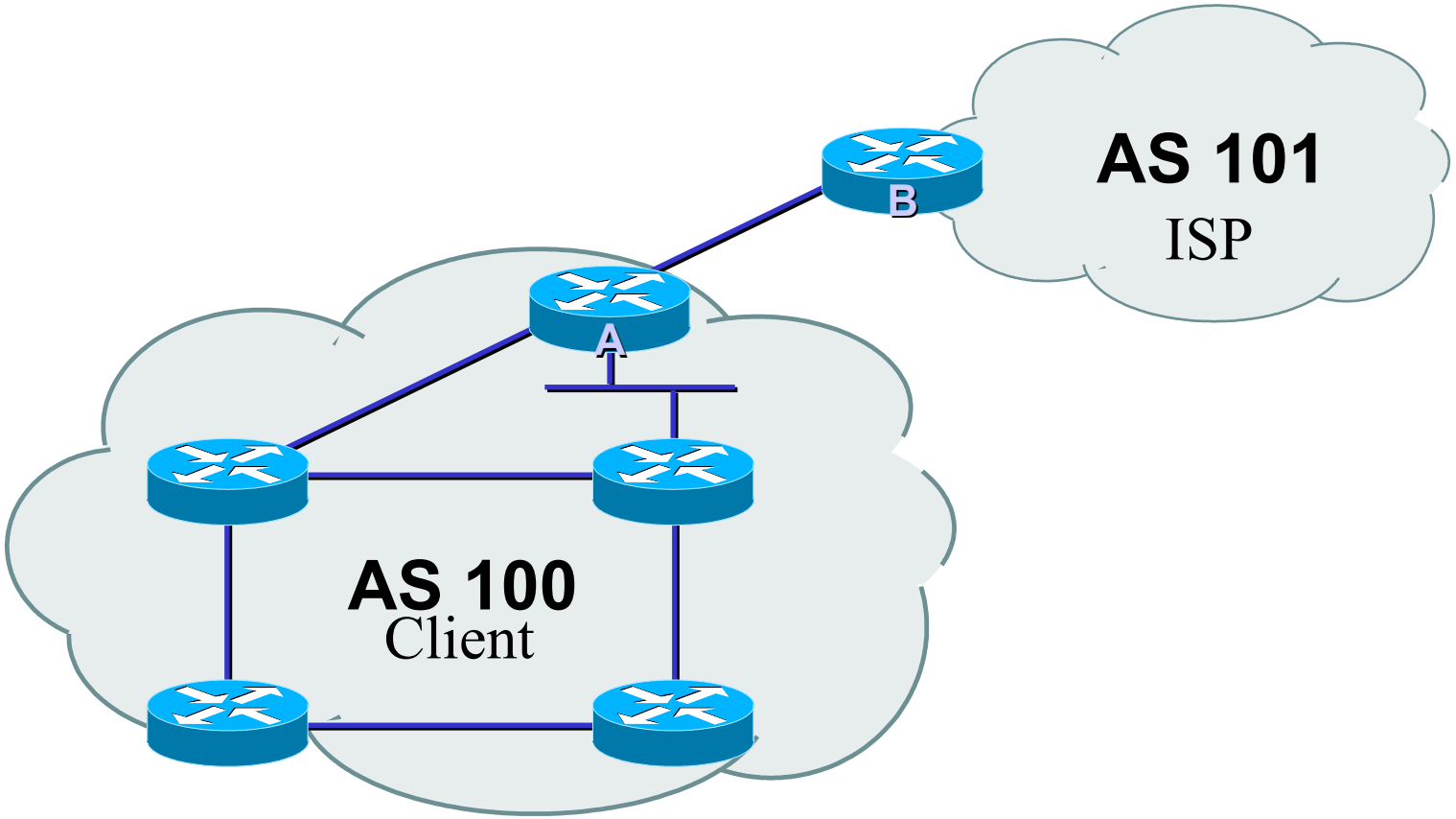
```
neighbor <x.x.x.x> route-map AS100_IN in
!
route-map AS100_IN permit 10
            set community 100:200
```

BGP et architecture de réseaux

AS “feuille” (stub AS)

- Situation ne nécessitant pas de BGP
- Route par défaut chez le FAI
- Le FAI annonce vos réseaux dans on AS
- La politique de routage de votre FAI est également la vôtre

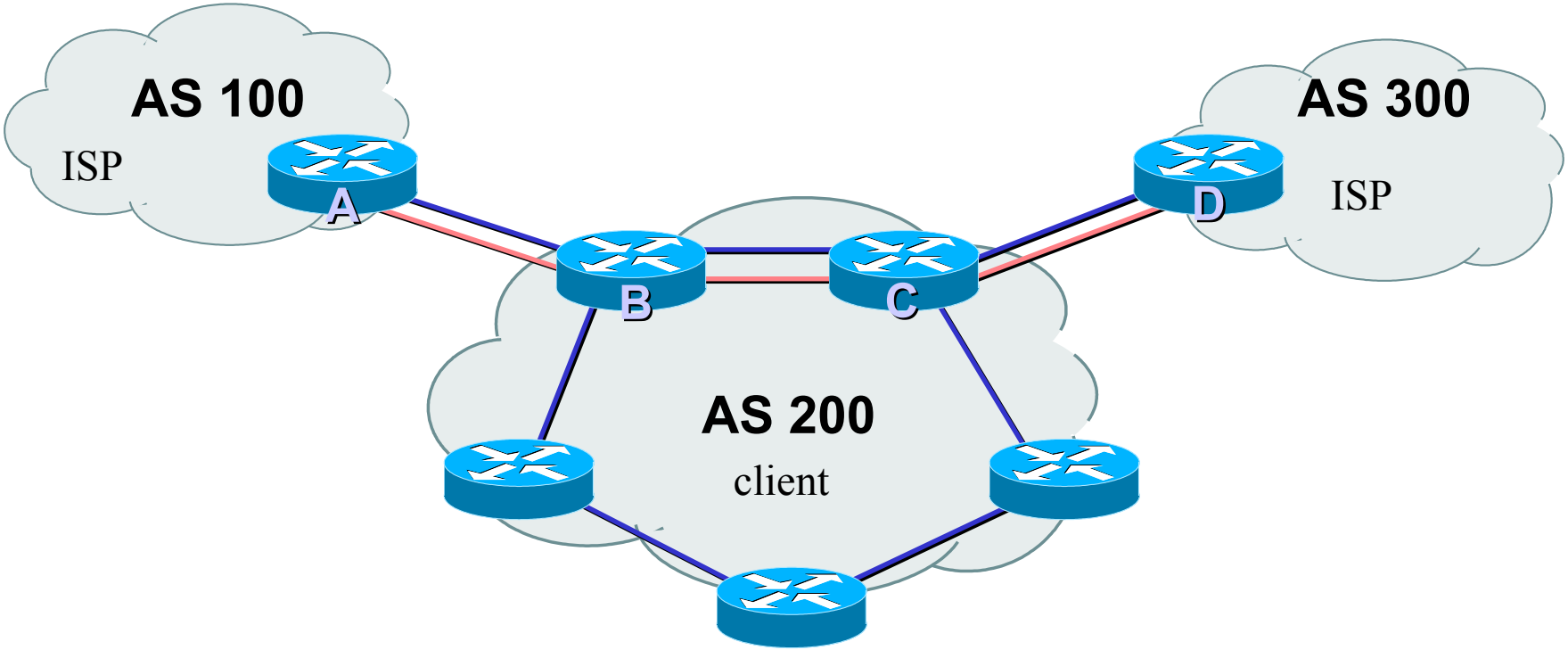
AS feuille



AS multi-raccordé (multi-homed)

- Les routeurs d'extrêmité font du BGP
- Sessions IBGP entre ces routeurs
- Il faut redistribuer les routes apprises avec prudence dans l'IGP, ou bien utiliser une route par défaut

AS multi-homé

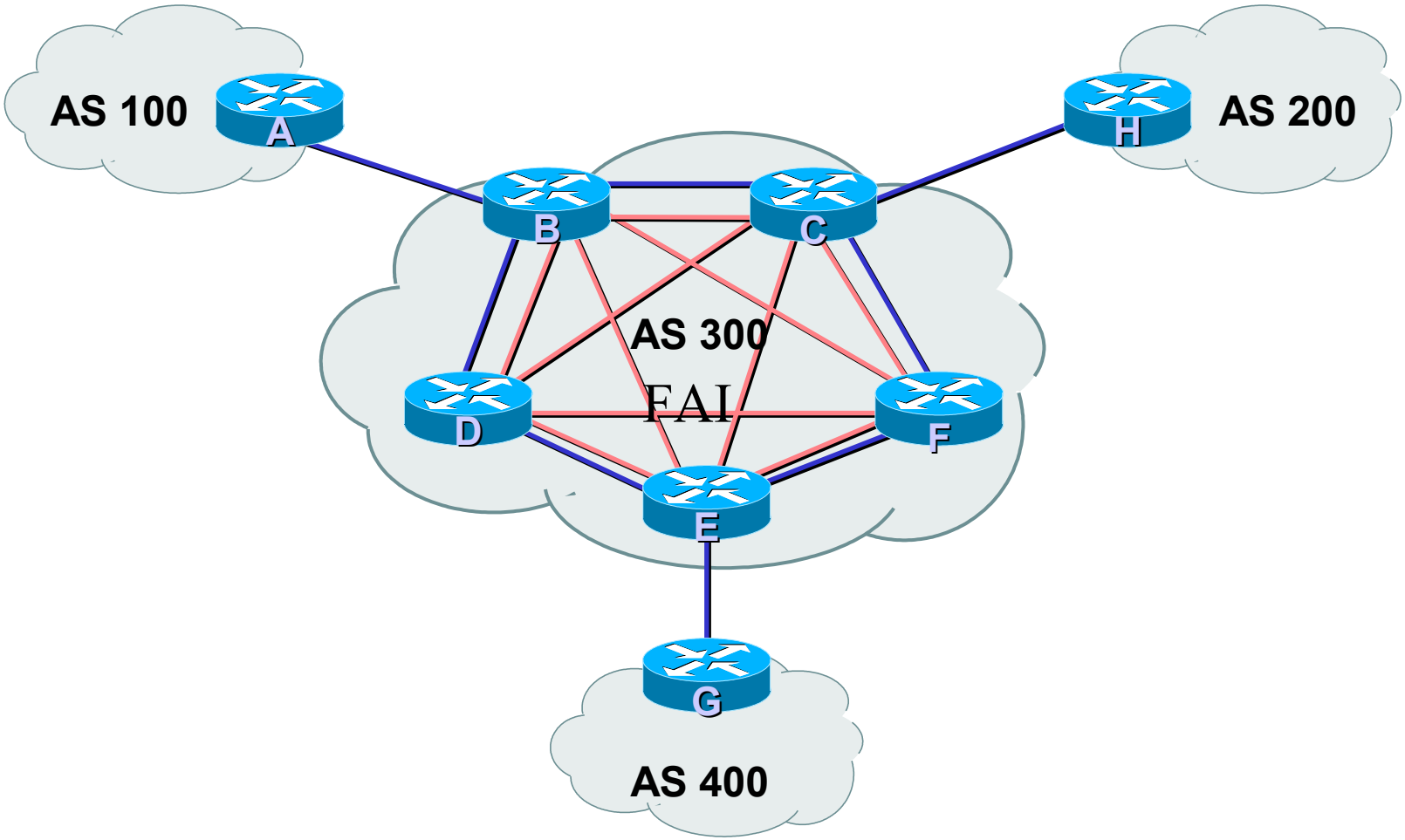


- Plus d'informations plus loin...

Réseau d'un FAI

- IBGP permet de transporter les routes extérieures à l'AS
- Un IGP permet de gérer la topologie du réseau
- Un maillage complet iBGP est requis

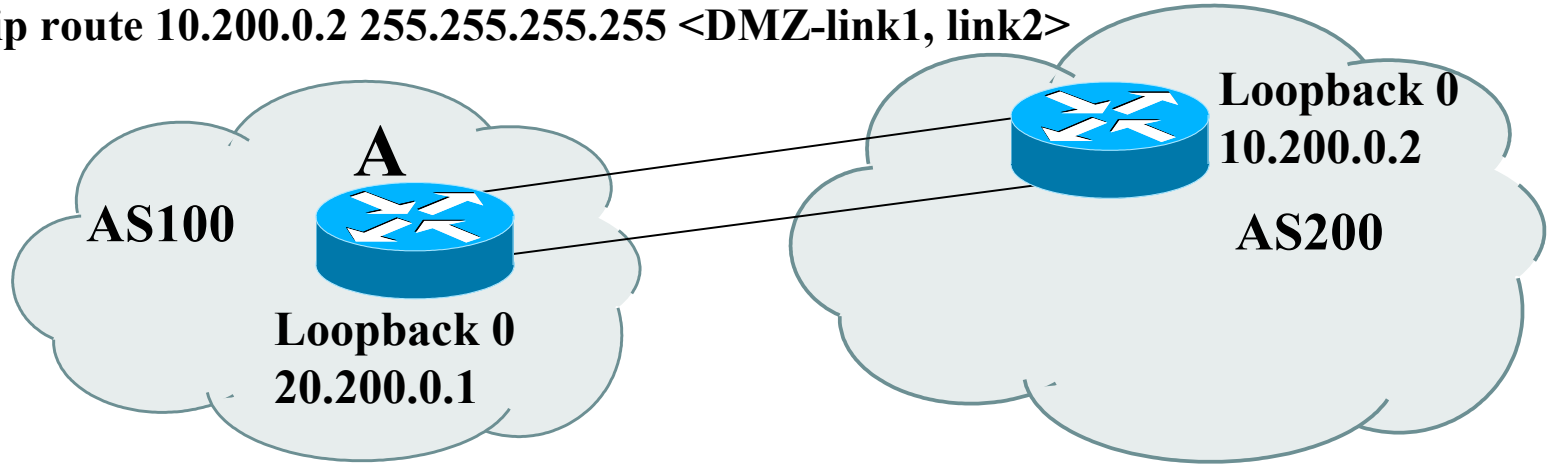
Réseau typique d'un FAI



Partage de charge - 1 chemin

Routeur A:

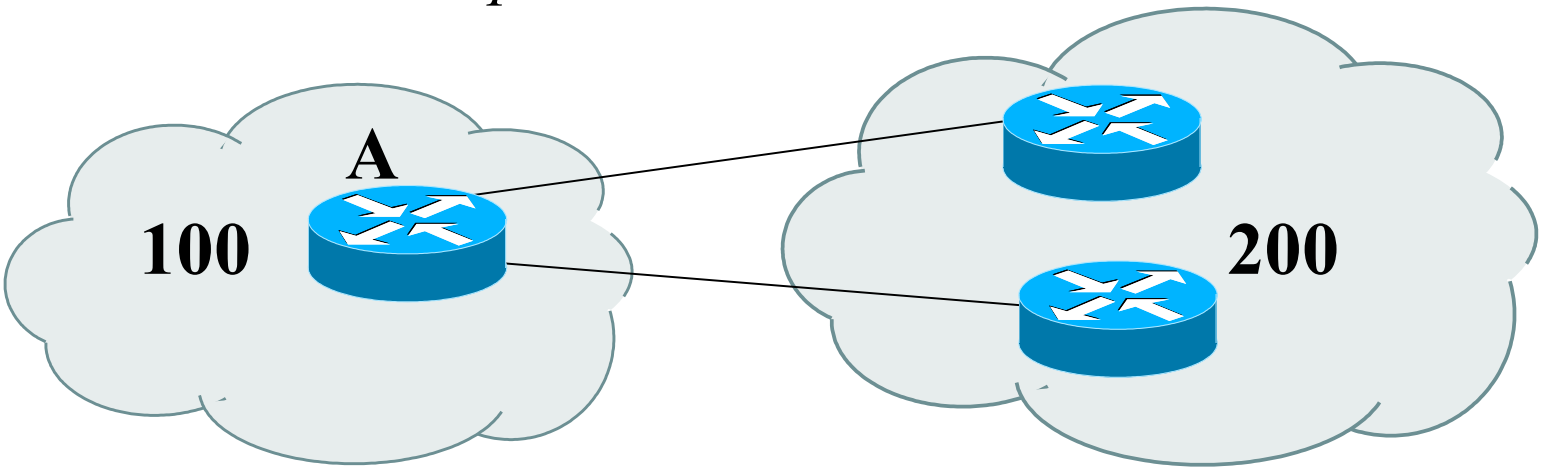
```
interface loopback 0
ip address 20.200.0.1 255.255.255.255
!
router bgp 100
neighbor 10.200.0.2 remote-as 200
neighbor 10.200.0.2 update-source loopback0
neighbor 10.200.0.2 ebgp-multi-hop 2
!
ip route 10.200.0.2 255.255.255.255 <DMZ-link1, link2>
```



Partage de charge - Plusieurs chemins disponibles

Routeur A:

router bgp 100
neighbor 10.200.0.1 remote-as 200
neighbor 10.300.0.1 remote-as 200
maximum-paths 2



Note : A n'annoncera que 1 seul "bestpath" à ses voisins iBGP

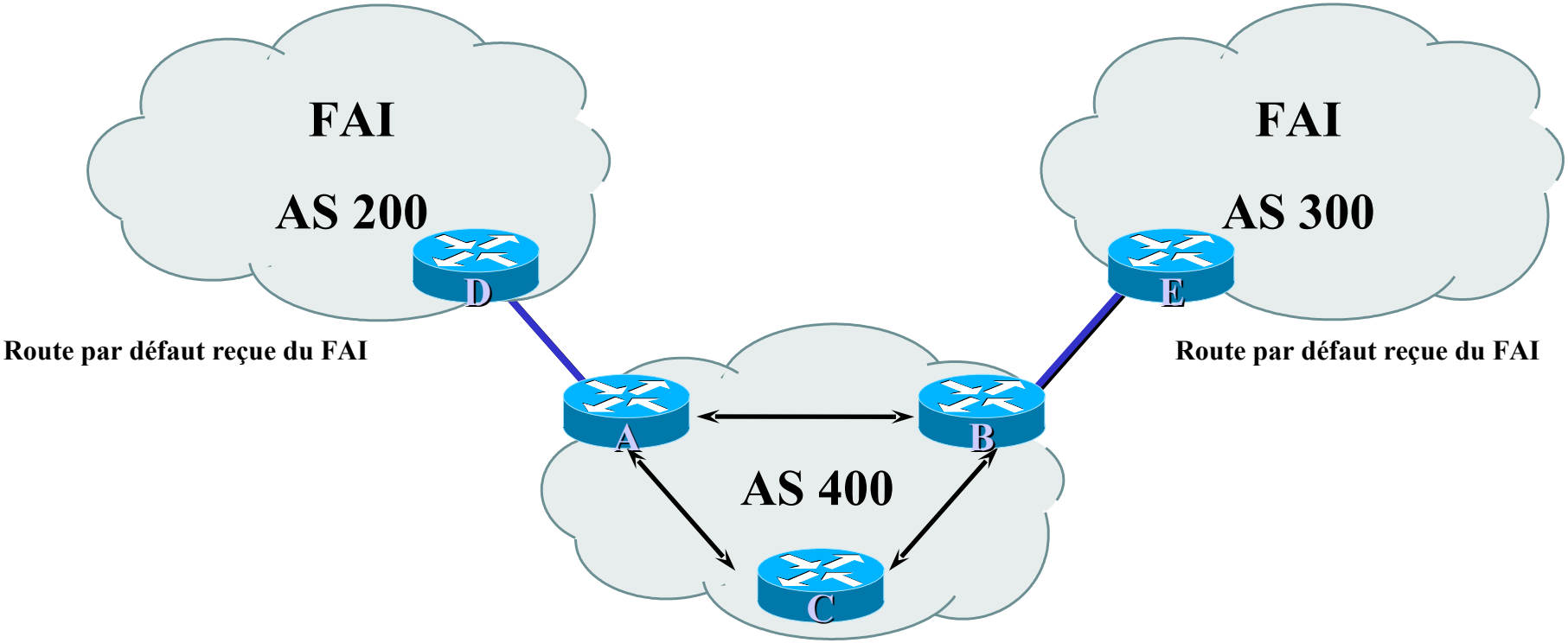
Redondance - Multi-homing

- Etre connecté de manière fiable à l'Internet
- 3 situations courantes en multi-homing
  - accepter la route par défaut des prestataires
  - clients + route par défaut chez les prestataires
  - recevoir toutes les routes de tous les voisins
- Adressage IP
  - fourni par les prestataires "upstream", ou
  - obtenu directement auprès d'un registre IP

Route par défaut des FAI

- Permet d'économiser la mémoire et la puissance de calcul
- Le FAI envoie une route par défaut BGP
  - le métrique IGP permet de choisir le FAI
- La politique des FAI détermine votre politique de trafic entrant
  - Il est cependant possible d'influencer cela en utilisant une politique de sortie, par exemple: AS-path prepend

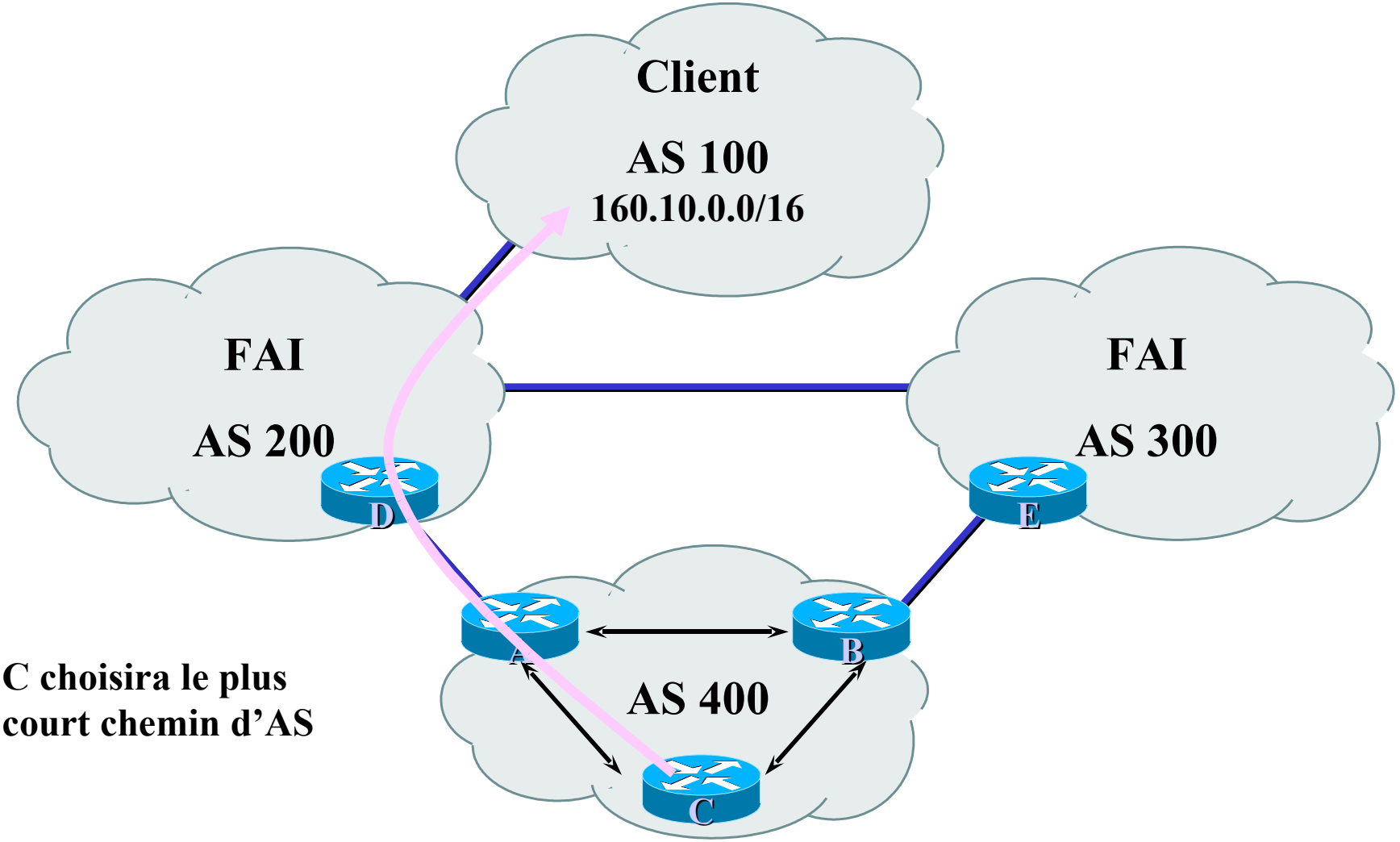
Route par défaut des FAI



Clients + route par défaut des FAI

- Consommation modérée de mémoire et CPU
- Gestion individuelle des routes des clients et route par défaut pour le reste
  - il est nécessaire de connaître les routes du client !
- Politique de routage entrant laissée aux FAI choisis
  - mais il est possible d'influencer ces choix (exemple : as-path prepend)

Les ISP annoncent les routes de leurs clients

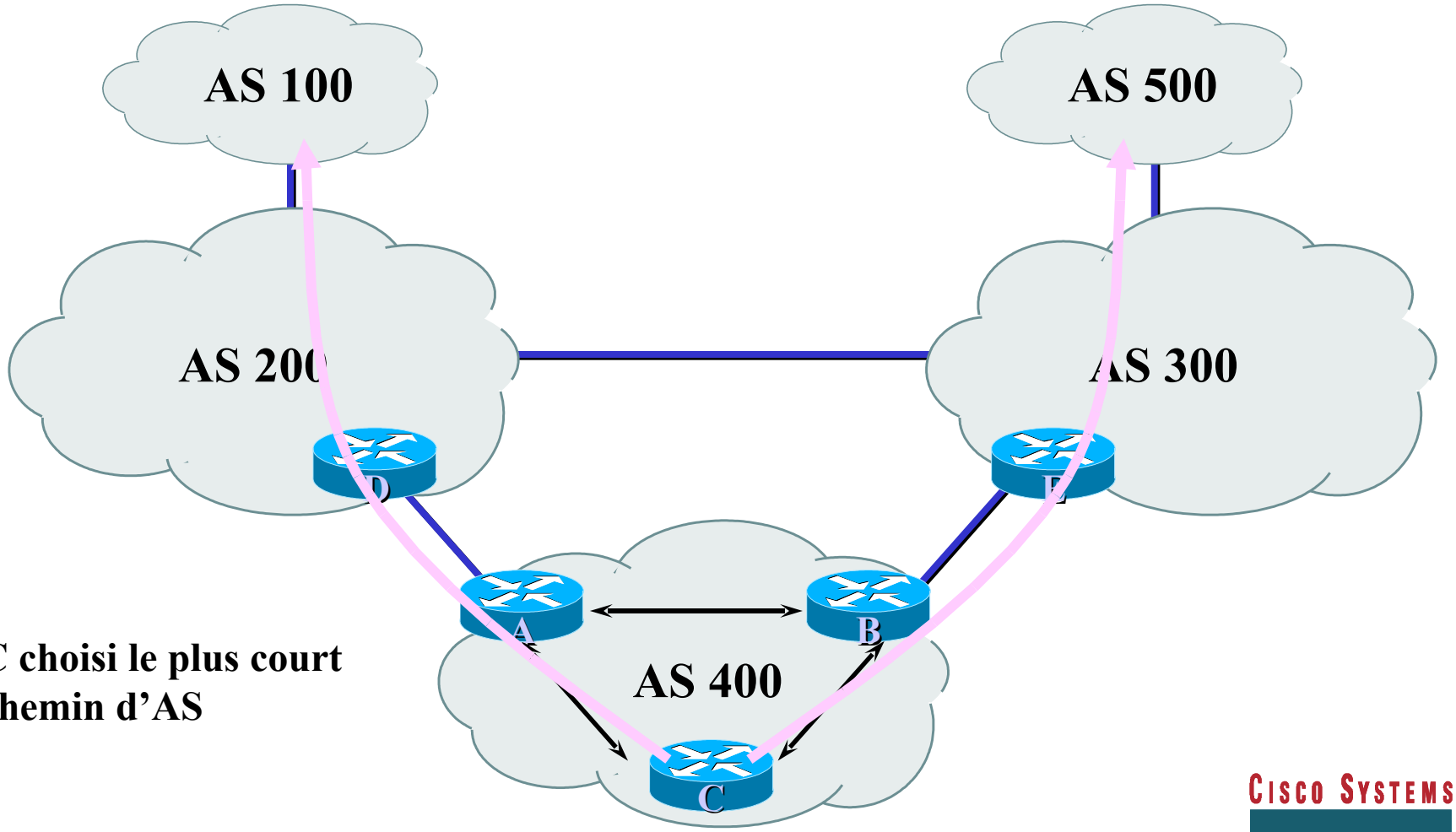


C choisira le plus court chemin d'AS

Gérer toutes les routes “full routing”

- Plus de consommation mémoire et CPU
- Contrôle plus poussé sur la politique de routage
- Les AS de transit gèrent généralement toutes les routes
- BGP est généralement le principal protocole de routage

Tous les prestataires envoient toutes les routes



C choisi le plus court chemin d'AS

Etat de l'art
Choix de l'IGP dans le Backbone

- L'IGP assure la gestion de la topologie de votre infrastructure - pas des réseaux de vos clients
- L'IGP doit converger rapidement
- L'IGP doit transporter les routes et masques - OSPF, IS-IS, EIGRP

Etat de l'art...
Raccorder un client

- Routes statiques
  - Vous les contrôlez directement
  - pas de "flaps"
- Protocole de routage dynamique
  - Vous devez filtrer ce que votre client annonce
  - Risque de "flaps"
- Utiliser BGP pour les clients "multi-homés"

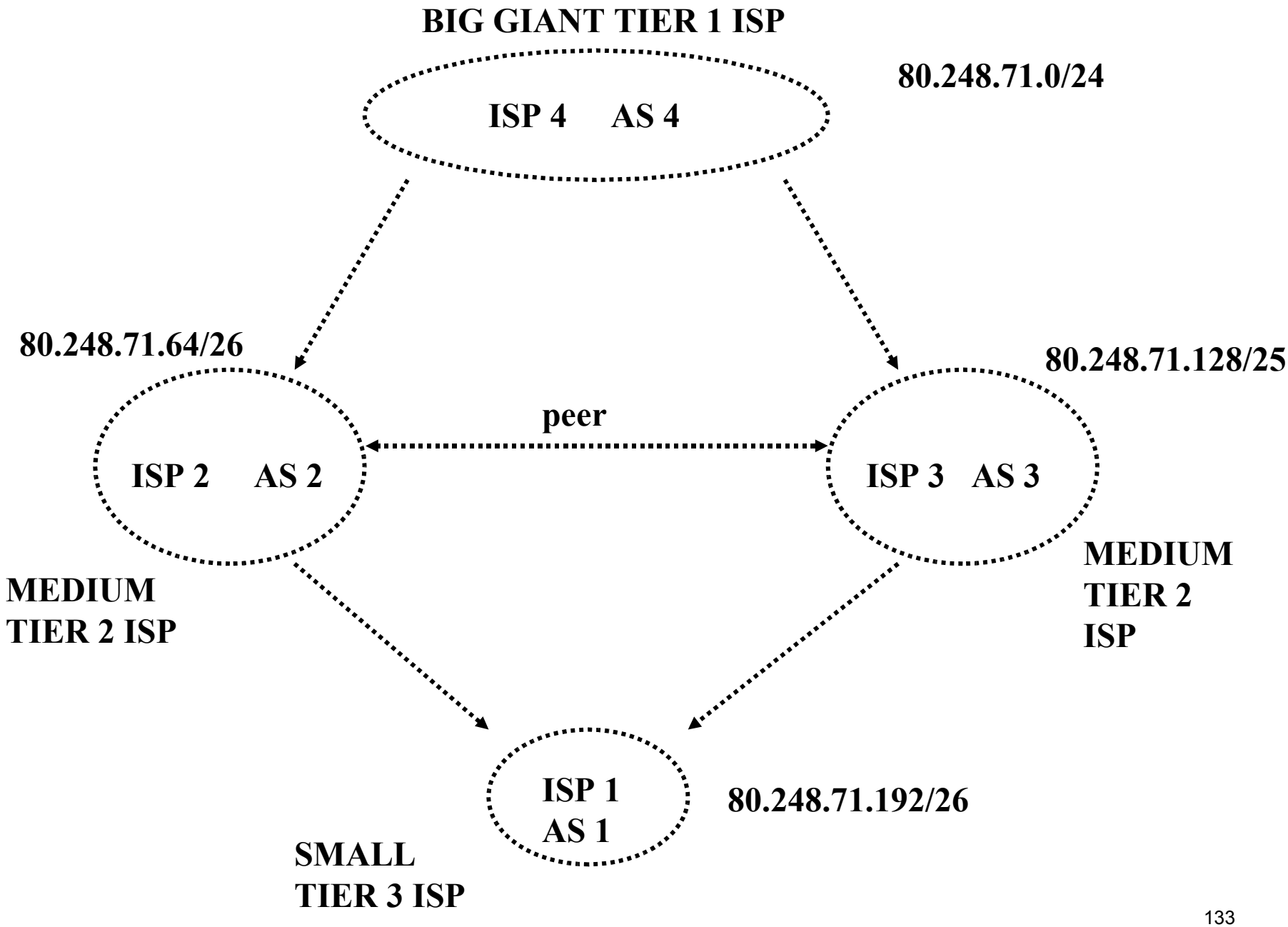
Etat de l'art...
Se raccorder à d'autres FAI

- Annoncez uniquement vos réseaux
- Acceptez le minimum nécessaire
- Prendre le plus court chemin vers la sortie
- Agrégez les routes !!!
- FILTREZ ! FILTREZ! FILTEZ!

Etat de l'art...
Les points d'échange

- Les raccordements longue distance sont chers
- Ils permettent de profiter d'un point unique pour se raccorder à plusieurs partenaires

Exercice 3 - Se connecter à un FAI


BIG GIANT TIER 1 ISP
80.248.71.0/24
ISP 4 AS 4
80.248.71.64/26
80.248.71.128/25
peer
ISP 2 AS 2
ISP 3 AS 3
MEDIUM
TIER 2 ISP
MEDIUM
TIER 2
ISP
ISP 1
AS 1
80.248.71.192/26
SMALL
TIER 3 ISP

Exercice 4 - Changement de politique de routage BGP

Questions & réponses