

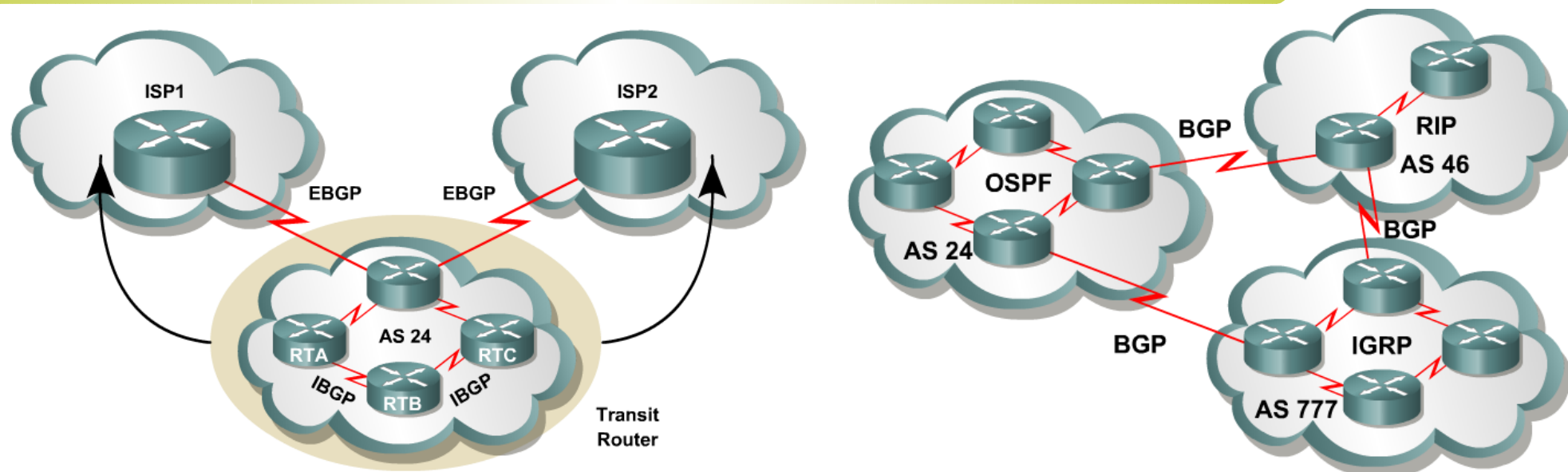
Interconnexion de Réseaux à Grande Échelle

Intro au protocole BGP

Quelques termes utilisés dans ce cours

- ▶ **IGP** (Interior Gateway Protocol)
 - ▶ RIP, IGRP, EIGRP, OSPF
 - ▶ Protocoles de routage utilisés à l'intérieur d'un système autonome (AS)
- ▶ **EGP** (Exterior Gateway Protocol)
 - ▶ BGP
 - ▶ Protocole de routage utilisé pour échanger des informations de routage entre différents systèmes autonomes (AS)
- ▶ **Autonomous System**
 - ▶ (RFC 1771) "Un ensemble de routeurs sous une seule administration, qui utilise un IGP et des métriques communes pour faire le routage des paquets à l'intérieur d'un AS and et un EGP pour le routage vers d'autres AS's."
- ▶ **BGP** un protocole de routage à vecteur de chemin (path vector)

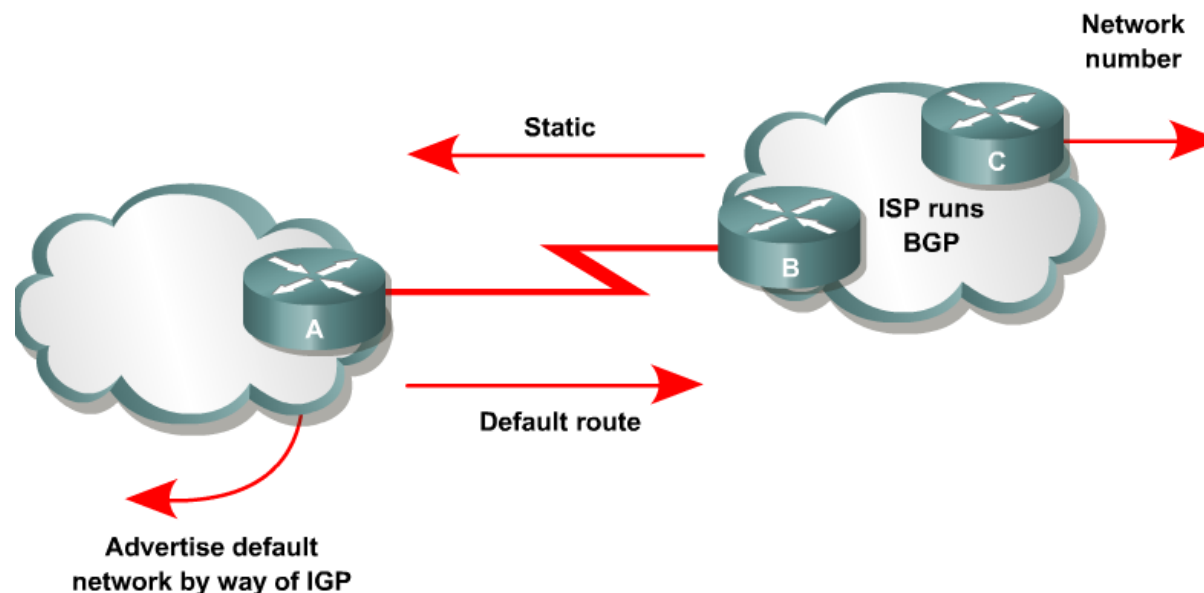
Est-il obligatoire d'utiliser BGP ?



L'utilisation de BGP dépend du niveau de connaissance technique et de la présence d'une de ces situations :

- ▶ Le AS permet le transit de paquets d'autres AS
- ▶ Le AS a plusieurs connexions à d'autres AS's (multihomed)
- ▶ Les flux entrants ou sortants de l'AS doivent être manipulés

BGP n'est pas nécessaire si



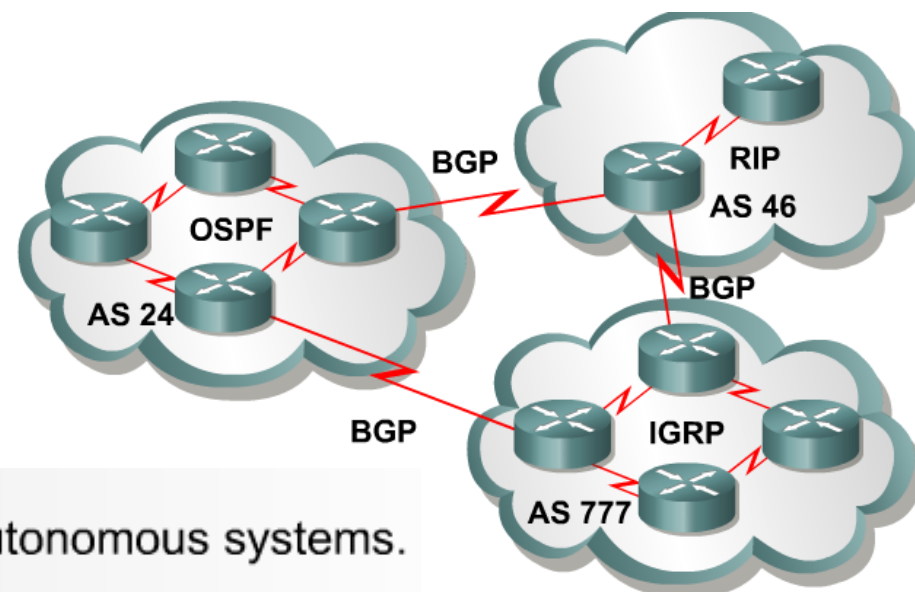
Ne pas utiliser BGP si :

- ▶ il n'y a qu'une seule connexion vers un autre AS ou FAI
- ▶ votre AS ne requiert des filtrages ou manipulations des flux
- ▶ les routeurs sont limités en mémoire ou capacité de calcul
- ▶ l'administrateur a une connaissance limitée du processus de filtrage et sélection des chemins BGP
- ▶ la liaison entre les AS a une faible bande passante

Qui réellement a besoin de BGP ?

- ▶ Pas beaucoup de monde
- ▶ “Vous devez utiliser BGP uniquement si des fortes raisons vous obligent, comme par exemple lorsque les IGP ne permettent le contrôle souhaité ou si la taille des tables de routage ne peut pas être contrôlée ”avec de l'agrégation”
- ▶ “Dans la plupart des cas, ce ne concerne que les FAIs et leurs liaisons”
- ▶ Jeff Dolye, Routing TCP/IP Vol. II

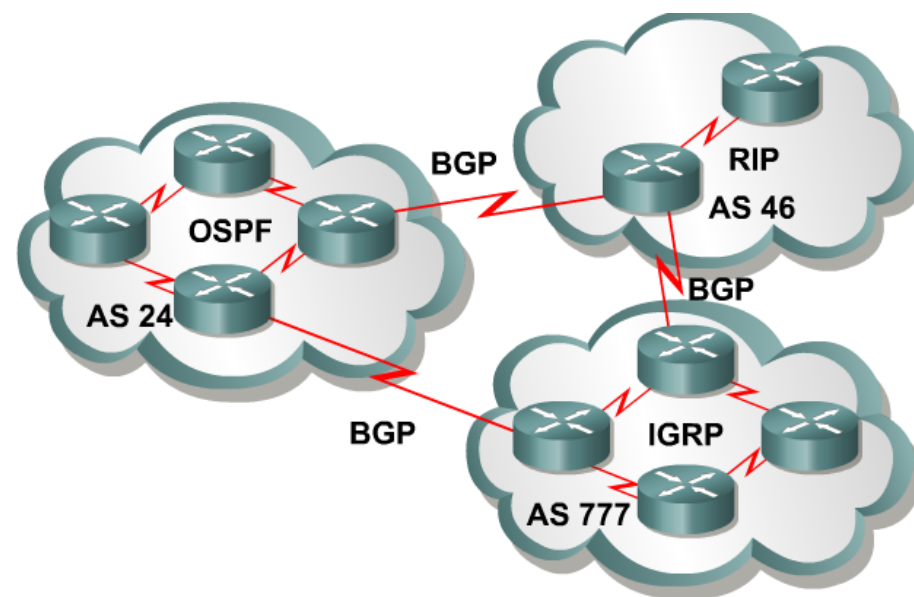
Définition des Systèmes Autonomes (AS)



EGPs, such as BGP, are used to interconnect autonomous systems.

- ▶ **AS** - Un ensemble de routeurs qui partagent des politiques de routage similaires et qui sont gérés dans un même domaine administratif. De l'extérieur, un AS est vu comme une entité unique
- ▶ Un AS peut être composé :
 - ▶ d'un ensemble de routeurs qui tourne le même protocole IGP
 - ▶ d'un ensemble de routeurs qui tournent différents protocoles de routage (une grande entreprise ou un FAI, par exemple)

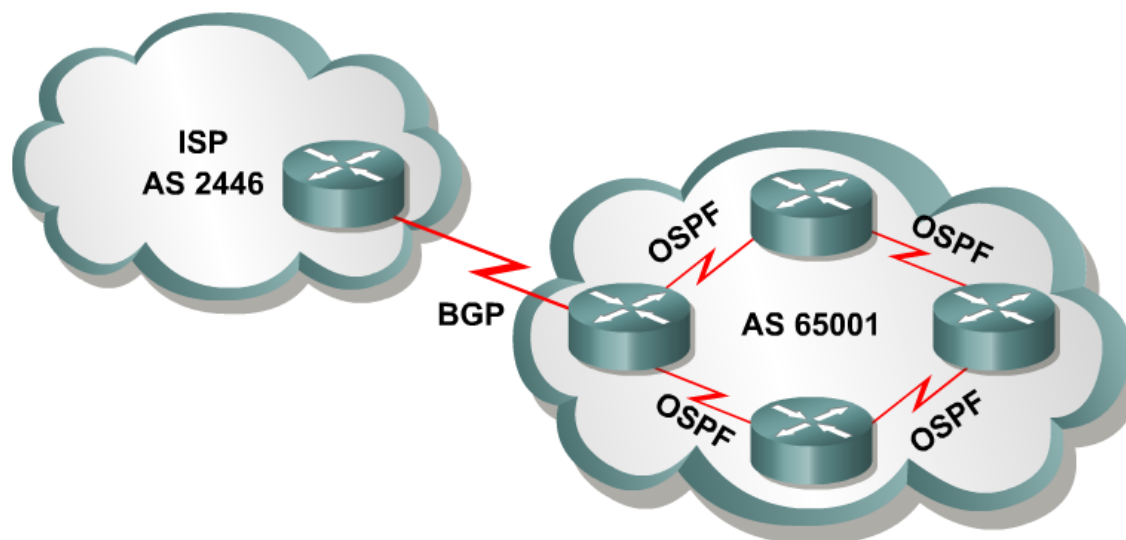
Identification des AS



Les numéros des AS

- ▶ Attribués par une entité
- ▶ Entre **1 et 65535**.
- ▶ Les adresse de **64512 à 65535** sont réservés pour un **usage privée**

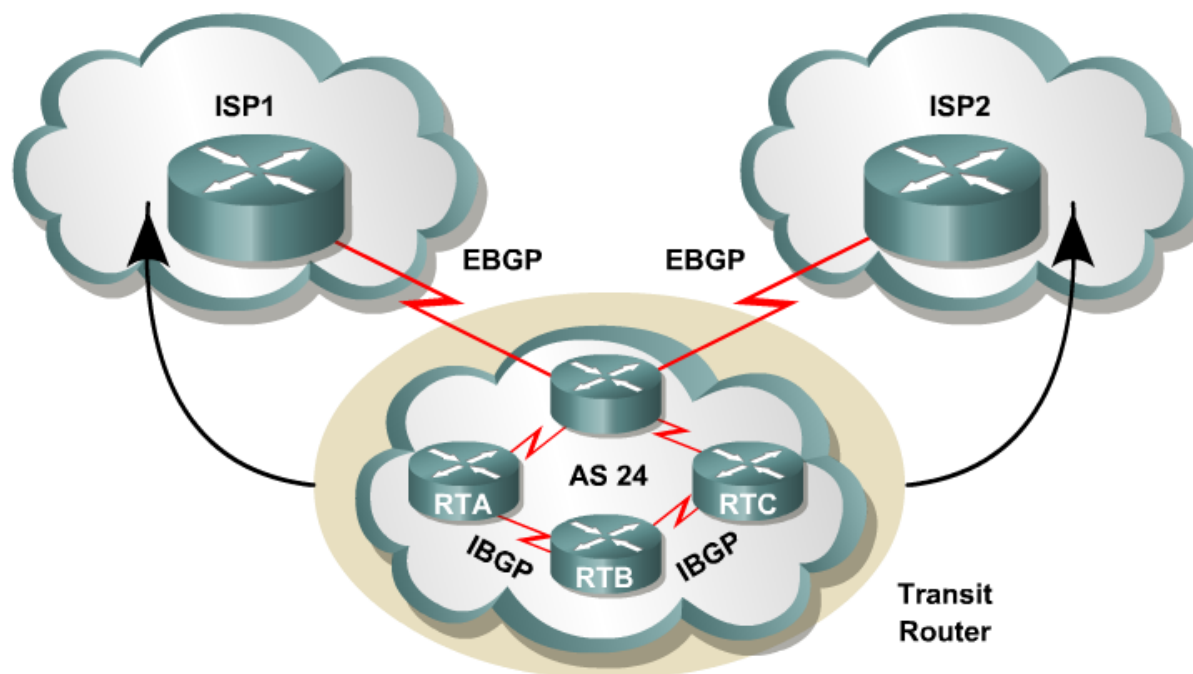
Attribution des AS



A provider may also choose to dynamically learn a customer's routes using BGP, which typically runs between the ISP router and the customer's boundary router.

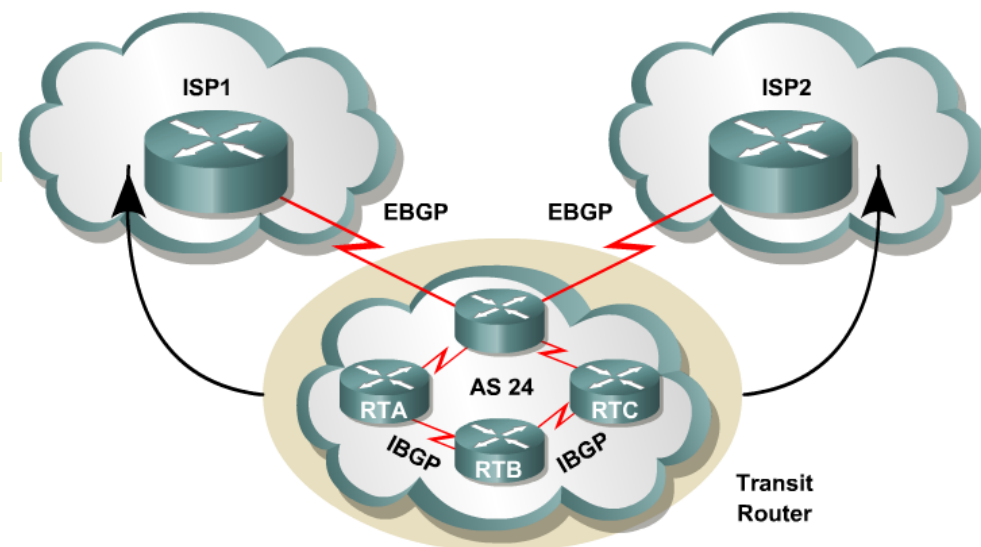
- ▶ À cause du nombre limité d'AS, il faut justifier le besoin d'un numéro AS
 - ▶ +/- 40000 déjà attribués
- ▶ Aujourd'hui, la IANA (Internet Assigned Numbers Authority) conseille fortement l'usage des AS 64512-65535 pour les AS qui sont connectés à un seul AS

BGP – concepts de base



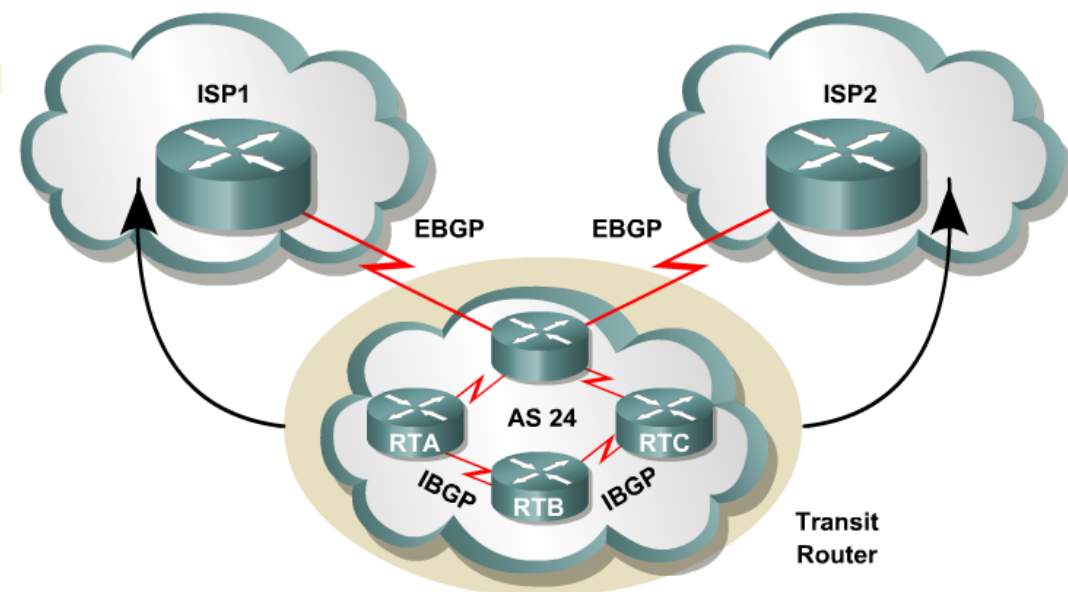
- ▶ BGP est un protocole de routage par vecteur de chemin
 - ▶ *path vector*
- ▶ RFC 1772
- ▶ Une route indique la liste d'ASs qu'il faut traverser pour arriver à la destination

BGP – concepts de base



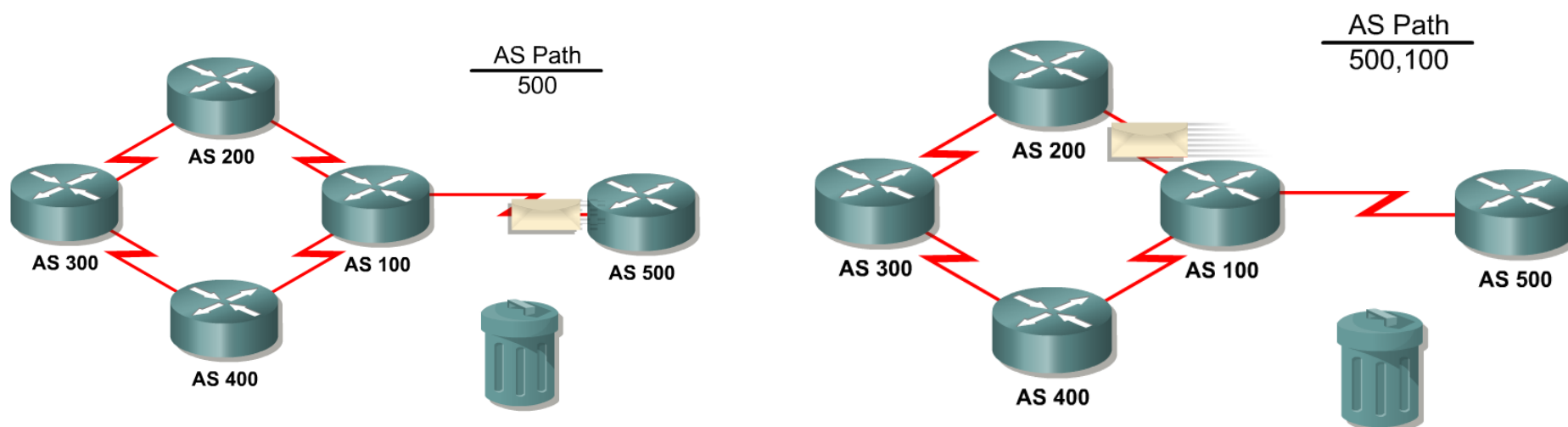
- ▶ BGP a por objectif :
 - ▶ l'échange d'informations de routage entre les ASs
 - ▶ garantir le choix d'une route sans boucle (loop-free)
- ▶ BGP4 est la première version de BGP qui supporte CIDR et agrégation de routes
- ▶ IGP comme RIP, OSPF et EIGRP utilisent des métriques
- ▶ BGP
 - ▶ n'utilise pas des métriques
 - ▶ fait les décisions selon des politiques de routage
 - ▶ n'indique pas les détails internes des ASs
 - ▶ ne représente qu'un arbre d'ASs
- ▶ **show ip bgp** – commande IOS pour afficher la table de routage BGP

BGP – concepts de base



- ▶ Les **mises à jour BGP** utilisent TCP sur le port 179
 - ▶ **RIP** utilise UDP sur le port 520
 - ▶ **EIGRP** utilise RTP
 - ▶ **OSPF** n'utilise pas un protocole de couche 4
- ▶ Du fait d'utiliser TCP, BGP a besoin de :
 - ▶ une connectivité IP entre les voisins BGP
 - ▶ ouvrir des connexions TCP avant d'échanger des données
- ▶ **BGP** bénéficie de la fiabilité de TCP

Chemin sans boucle (Loop-free)



- Pour éviter les boucles, BGP construit un graphe avec les données envoyées par les voisins
- La connexion entre deux AS représente un chemin
- **AS Path** – Le chemin pour atteindre un AS est représenté comme une séquence de numéros d'AS
- Grâce aux numéros d'AS, un système autonome peut détecter un chemin qui passe deux fois par le même AS

BGP Updates

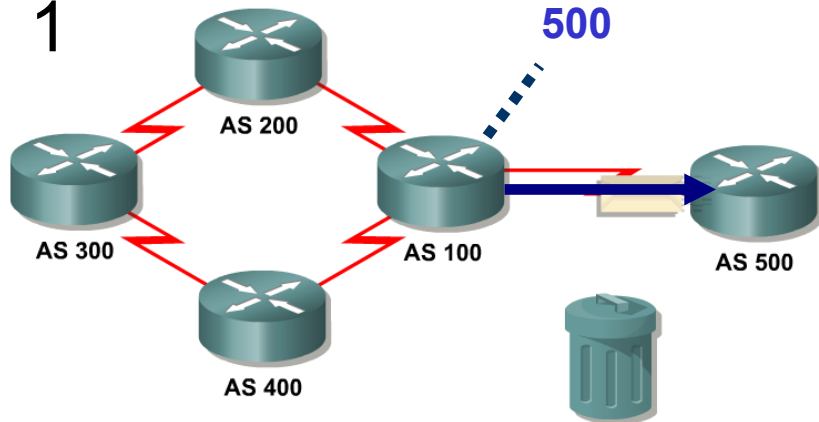
UNIVERSITÉ
DE CHAMPAGNE

Académie Cisco

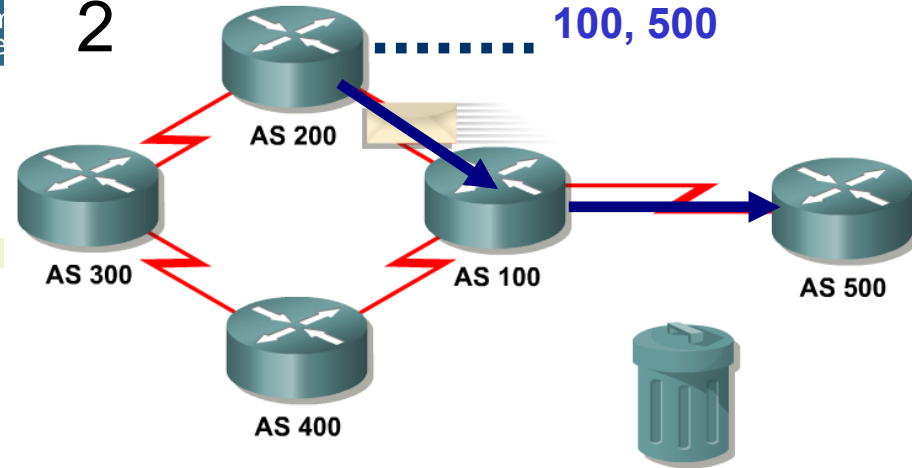
Reims
Champagne

Faculté des Sciences
Reims

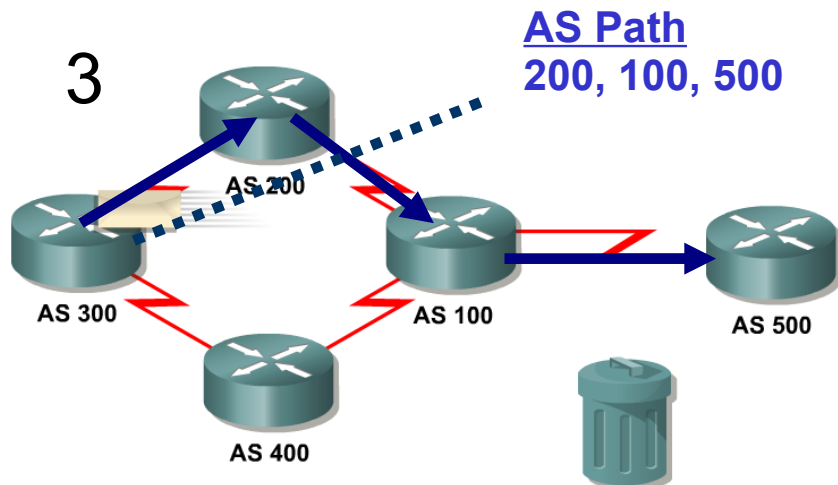
1



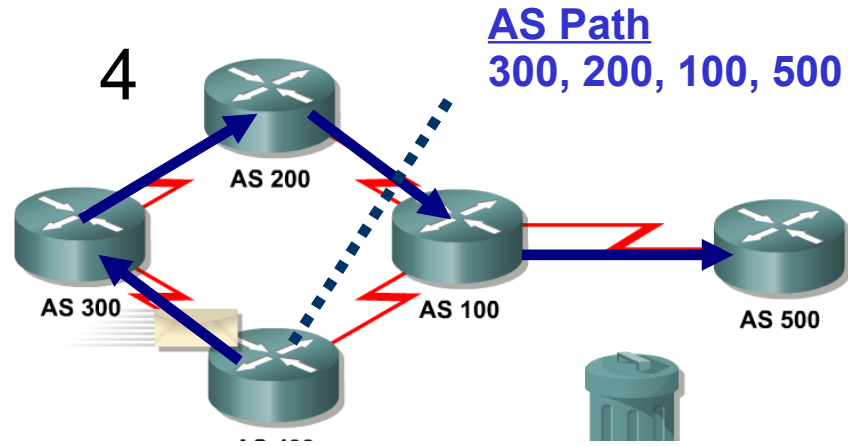
2



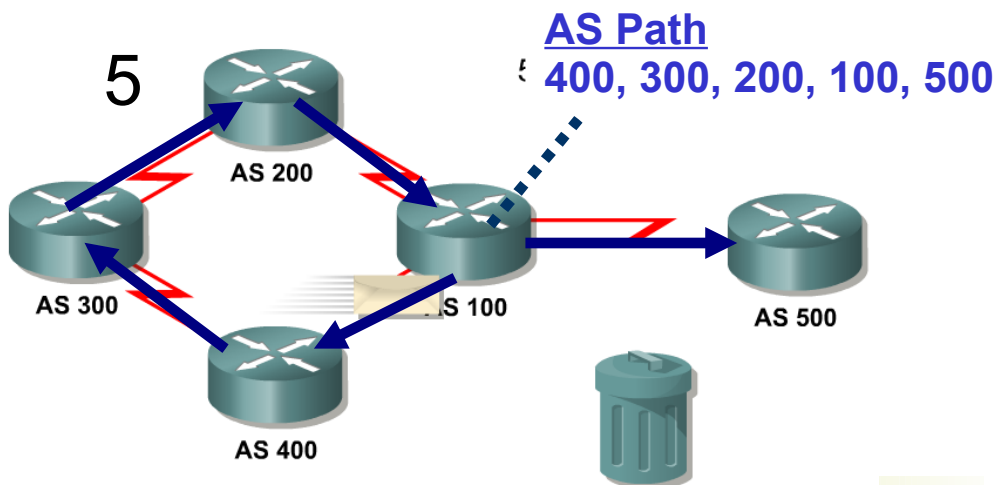
3



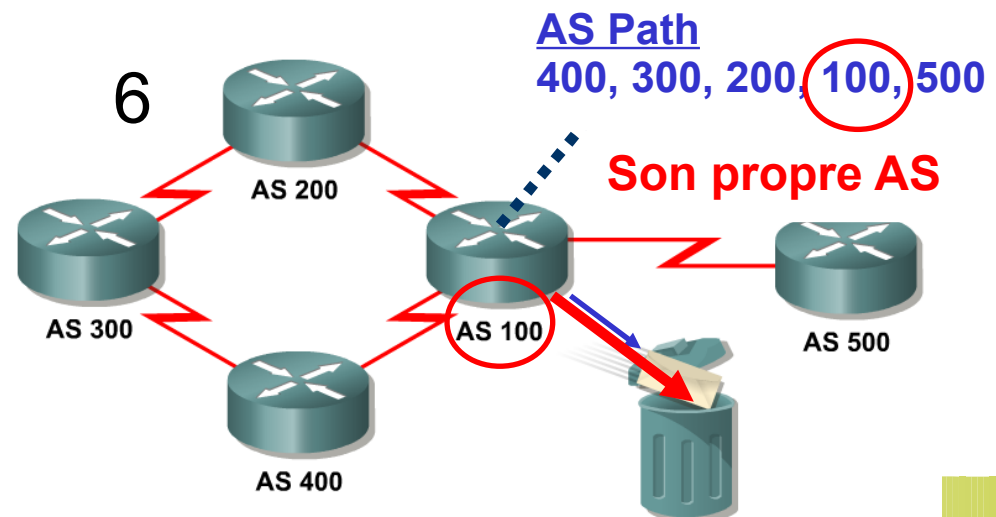
4



5



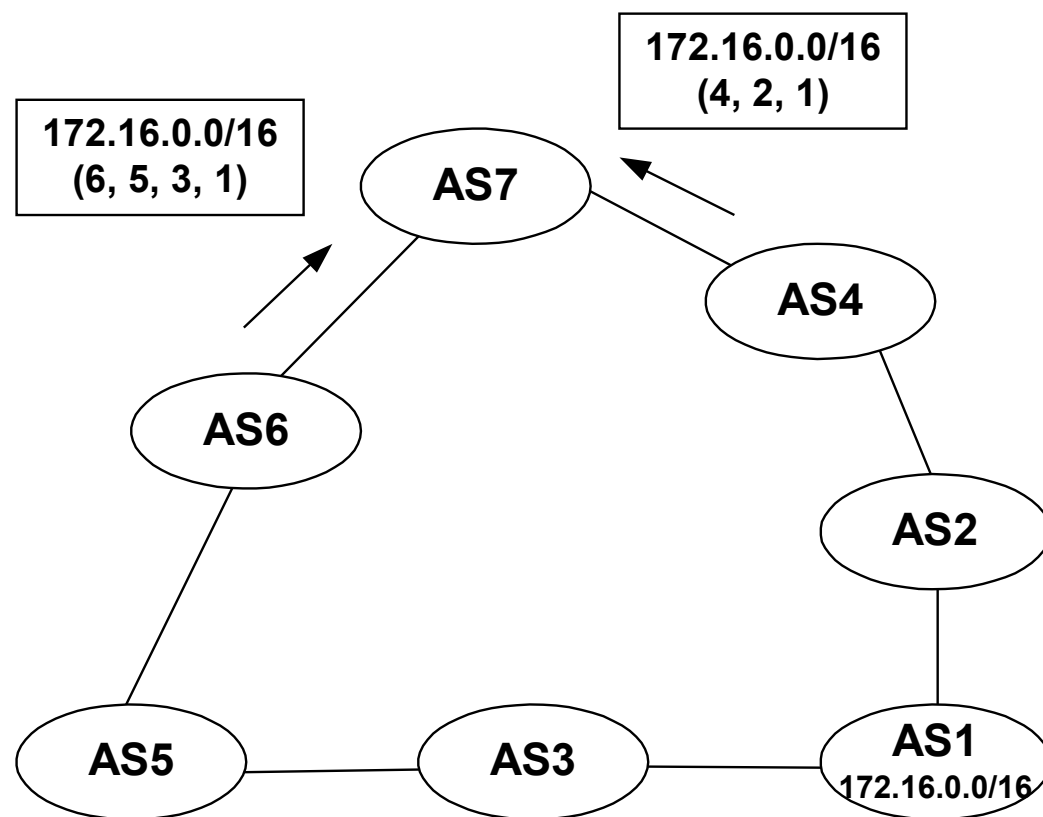
6



Chemin sans boucles

► AS_PATH

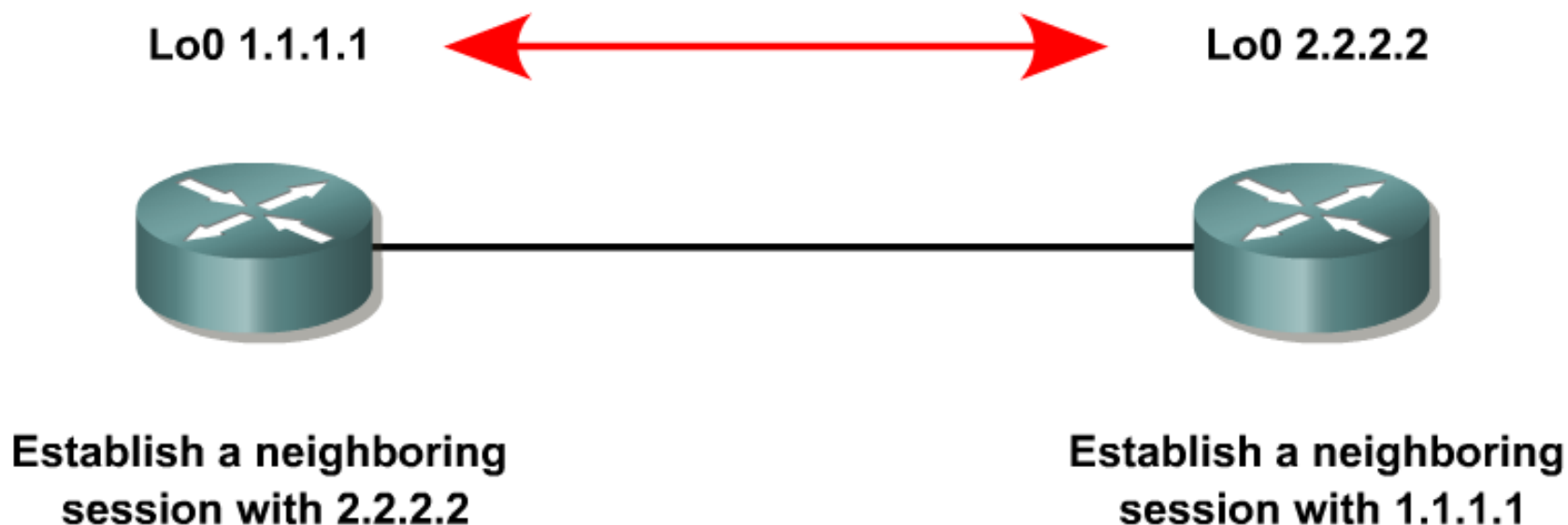
- Liste des AS associés à une route
- L'un des attributs associés aux routes (à voir plus tard)
- La stratégie la plus simple consiste à choisir le chemin le plus court
- Pour deux annonces de réseaux identiques (mêmes prefixes), BGP préfère les routes avec le AS_PATH le plus petit
- Dans cet exemple, AS7 choisit le chemin le plus court (4, 2, 1)



Les types de message BGP

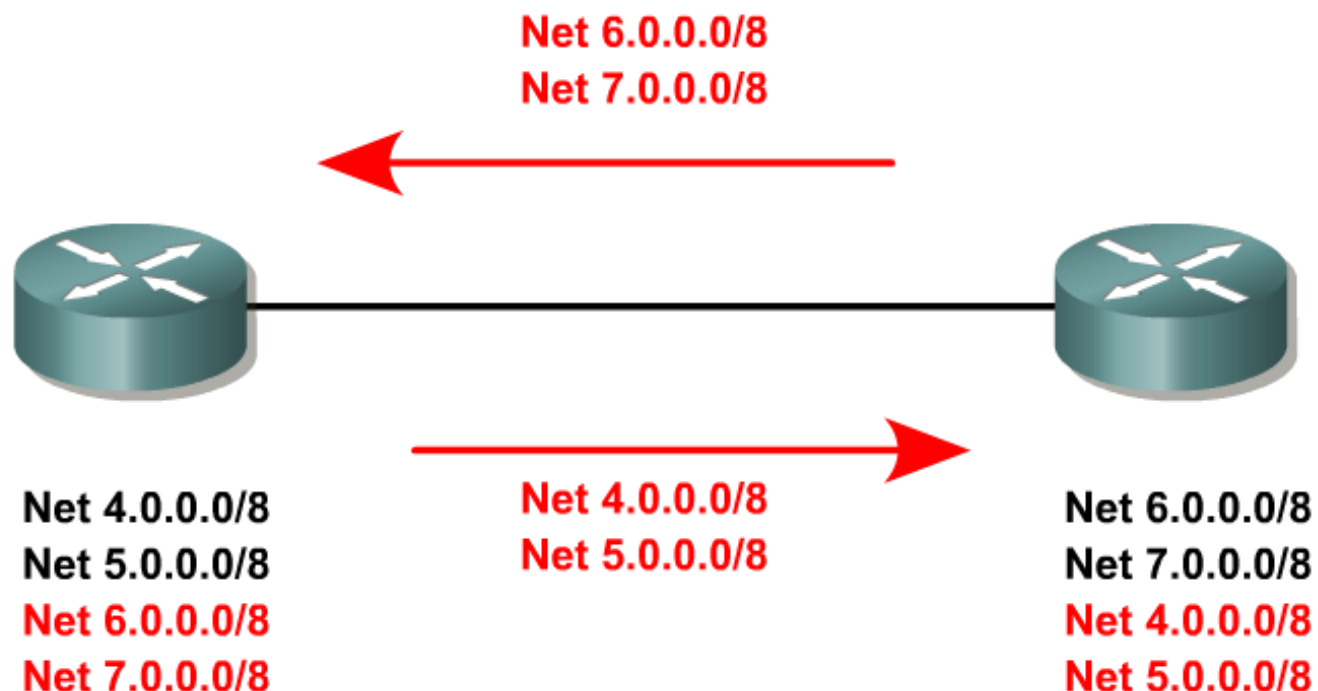
- ▶ Avant toute communication entre deux peers il est nécessaire de démarrer une connexion TCP sur le port 179.
- ▶ Une fois ouverte la connexion, BGP échange plusieurs messages avec les paramètres de la connexion et les informations de routage.
- ▶ Tous les messages BGP sont unicast vers un seul partenaire.
- ▶ Il y a 4 types de message BGP :
 - ▶ Type 1: OPEN – identification initiale entre les peers
 - ▶ Type 2: KEEPALIVE
 - ▶ Type 3: UPDATE – mises à jour
 - ▶ Type 4: NOTIFICATION - erreurs

Fonctionnement de BGP



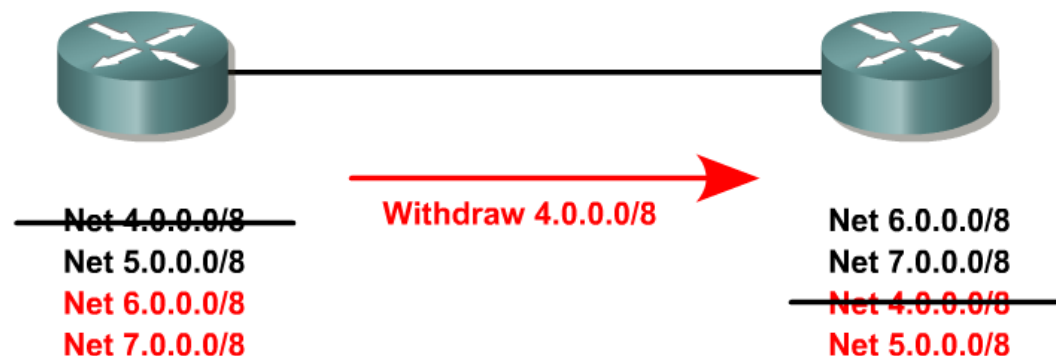
- ▶ **Voisins** ou **pairs** – Deux routeurs qui établissent une connexion TCP
- ▶ Chaque routeur doit exécuter le protocole BGP

Échange Initial



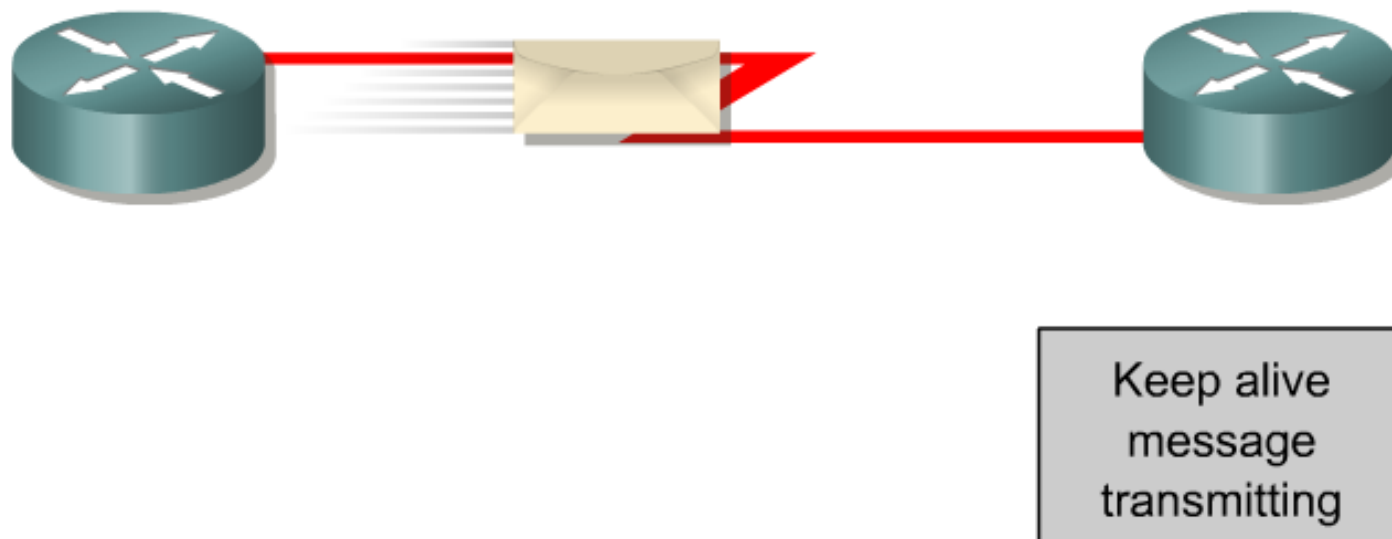
- ▶ Lorsqu'une connexion est établie, les routeurs échangent leurs routes candidates
- ▶ Après l'échange initial, des mises à jour incrémentales sont envoyées entre les routeurs

Suppression de Routes



- ▶ Les informations sur les routes peuvent être aussi
 - ▶ une route qui devient inaccessible
 - ▶ un meilleur chemin disponible
- ▶ Les routeurs BGP gardent un tableau de numéros de version
 - ▶ Liste des numéros de version des mises à jour reçues de chaque routeur
 - ▶ Incrementé à chaque modification des tables de routage BGP
 - ▶ *Une incrementation rapide indique normalement un réseau instable ou mal configuré*

BGP Keepalives



- ▶ Les pairs échangent des messages **keepalive** pour garder active la connexion
- ▶ Le défaut Cisco est de **60 secondes** (le RFC 1771 n'espécifie pas une valeur standard)
- ▶ Si l'intervalle est dépassé trois fois (**180 seconds**) un pair est considéré **down**
- ▶ Les intervalles peuvent être configurées

Attributs d'un Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

- ▶ **Chaque route** a son propre ensemble d'attributs, qui peut inclure
 - ▶ des informations sur le chemin
 - ▶ préférences sur les routes
 - ▶ next-hop
 - ▶ informations d'agrégation
- ▶ Les admins utilisent ces informations pour composer des politiques de routage
- ▶ À partir des valeurs des attributs, BGP peut être configuré pour
 - ▶ filtrer des informations sur certaines routes
 - ▶ choisir des chemins préférentiels

Attributs de Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

- ▶ Certains attributs ne sont pas reconnus par tous les fabricants
- ▶ Les attributs de chemin ont quatre types
 - ▶ **Well-known mandatory**
 - ▶ **Well-known discretionary**
 - ▶ **Optional transitive**
 - ▶ **Optional non-transitive**

Attributs de Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

Well-known mandatory

- ▶ **Un attribut qui doit être présent dans une mise à jour BGP**
- ▶ Reconnu par toute implémentation BGP
- ▶ Si un attribut well-known est manquant, une notification d'erreur sera créée
 - ▶ établit un standard minimum des attributs

Exemple : l'attribut AS_PATH

Attributs de Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

Well-known discretionary

- ▶ Attribut qui doit aussi être reconnu par toute implémentation BGP
- ▶ Peut ne pas être envoyé dans un message BGP UPDATE

Exemple : LOCAL_PREF

Attributs de Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

Optional transitive

- ▶ Un attribut qui peut ne pas être reconnu par une implémentation BGP (donc **optionnel**)
- ▶ Cet attribut est **transitif**
 - ▶ BGP doit **accepter et annoncer** l'attribut même s'il ne le reconnaît pas

Exemple : COMMUNITY

Attributs de Chemin

Attribute Code	Type
1 — ORIGIN	Well-known mandatory
3 — NEXT_HOP	Well-known mandatory
2 — AS_PATH	Well-known mandatory
4 — MULTI_EXIT_DISC	Optional nontransitive
5 — LOCAL_PREF	Well-known discretionary
6 — ATOMIC_AGGREGATE	Well-known discretionary
7 — AGGREGATOR	Well-known discretionary
8 — COMMUNITY	Optional transitive (Cisco)
9 — ORIGINATOR_ID	Optional nontransitive (Cisco)
10 — Cluster List	Optional nontransitive (Cisco)
11 — Destination Preference	(MCI)
12 — Advertiser	(Baynet)
13 — rcid_path	(Baynet)
255 — Reserved	—

Optional non-transitive

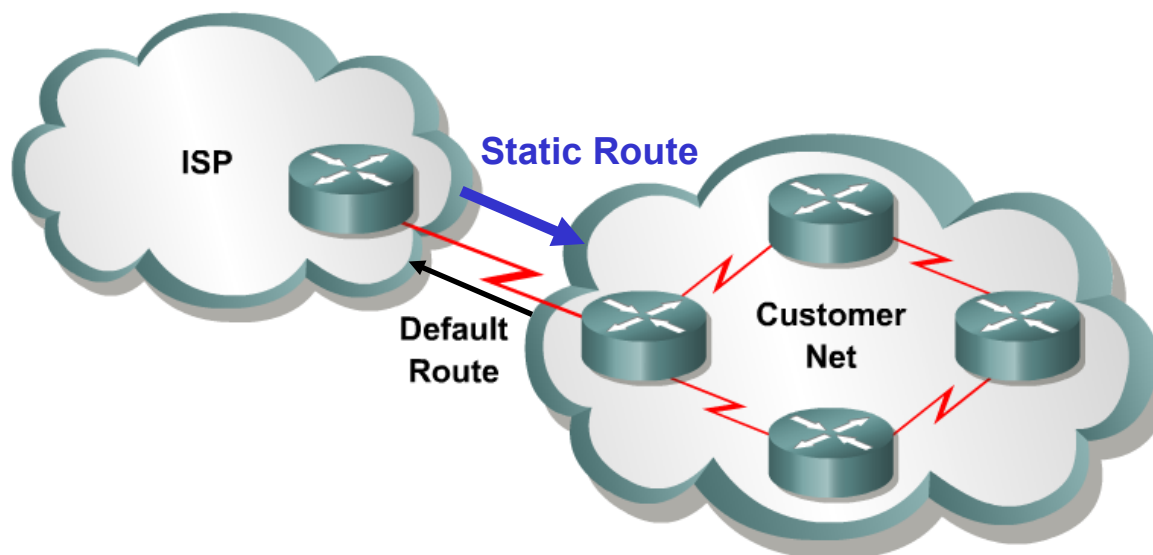
- ▶ Un attribut qui peut ne pas être reconnu par une implémentation BGP
- ▶ Même si un router reconnaît l'attribut, il est **non-transitif**
 - ▶ **Ne doit PAS être repassé aux autres pairs BGP**

Exemple : ORIGINATOR_ID

Résumé de la procédure du choix des routes

- ▶ BGP choisit un seul chemin (pas d'équilibrage de charge)
- ▶ Quand un chemin est choisi, il est placé dans la table de routage et annoncé aux autres routeurs
- ▶ L'ordre de choix :
 1. Si le prochain saut est inaccessible, ignore le chemin
 2. Préférence aux chemins avec **le plus grand poids**
 3. Préférence aux chemins avec **la plus grande préférence locale**
 4. Préférence aux chemins indiqués par le processus BGP
 5. Préférence aux chemins avec **le plus court AS_PATH**
 6. Préférence aux chemins avec **le plus petit type d'origine**
IGP < EGP < Incomplete
 7. Préférence aux chemins avec **le plus petit attribut MED**
 8. Préférence aux chemins extérieurs
 9. Préférence aux chemins avec **le plus proche voisin IGP**
 10. Préférence aux chemins avec **la plus petite adresse IP**, comme indiqué par **BGP router ID**.

Quoi faire dans le cas des AS «single-homed» ?

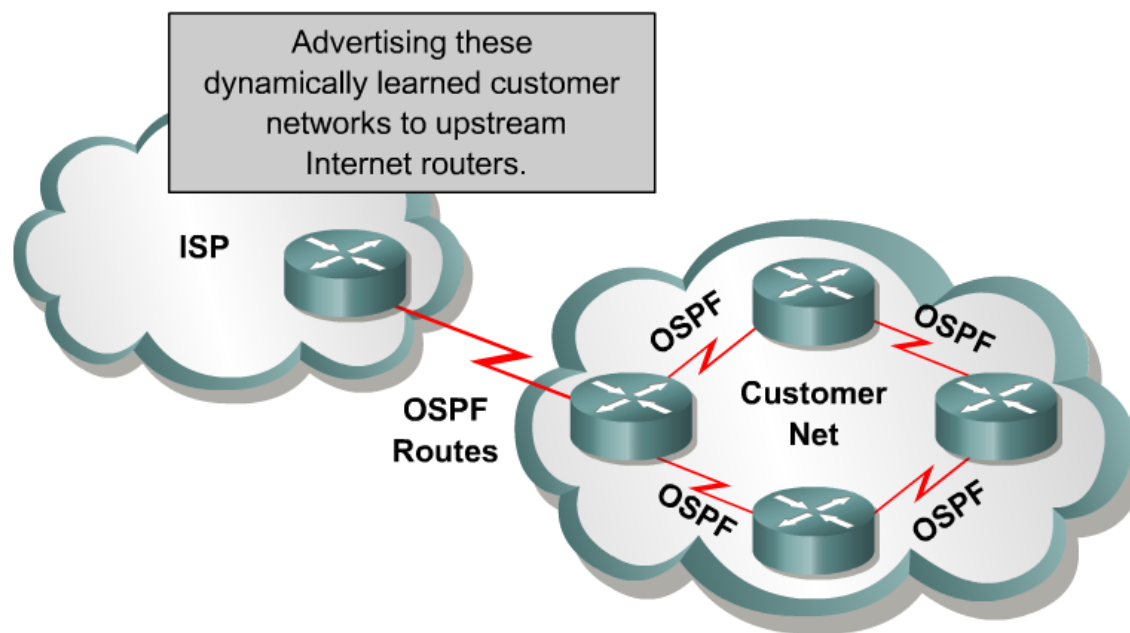


A single-homed AS can be configured with a default route to reach outside networks.

Première alternative : routes statiques

- ▶ **Un AS Single-homed system** n'a qu'un seul point de sortie
 - ▶ aussi connus comme des réseaux stub
 - ▶ généralement utilisent une route défaut pour le trafic destiné à l'extérieur de l'AS
- ▶ Dans ce cas, BGP **n'est pas** nécessaire

Quoi faire dans le cas des AS «single-homed» ?

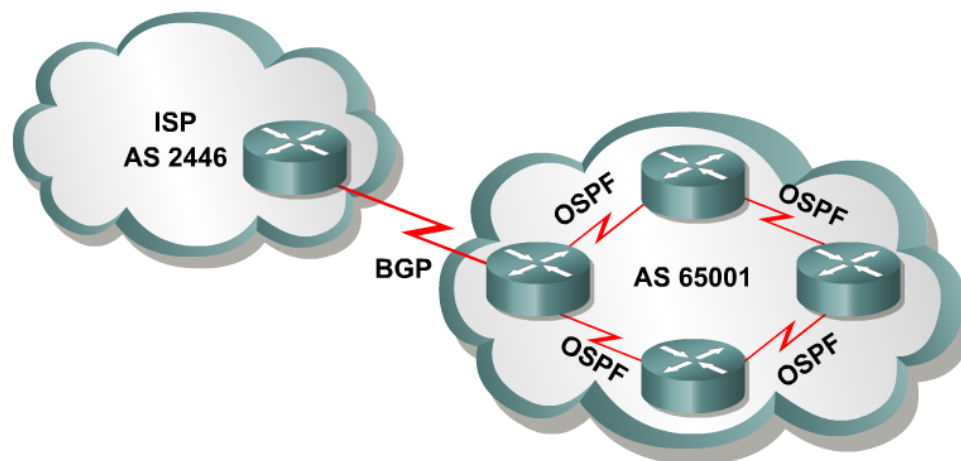


A provider may choose to dynamically learn customer routes using an IGP, such as OSPF.

Utilisation d'un protocole IGP

- ▶ Le FAI et le client peuvent utiliser une protocole IGP pour partager les routes du client
- ▶ A l'avantage de faire le ISP bénéficier du routage dynamique
- ▶ Dans ce cas, BGP **n'est pas** nécessaire

Quoi faire dans le cas des AS «single-homed» ?

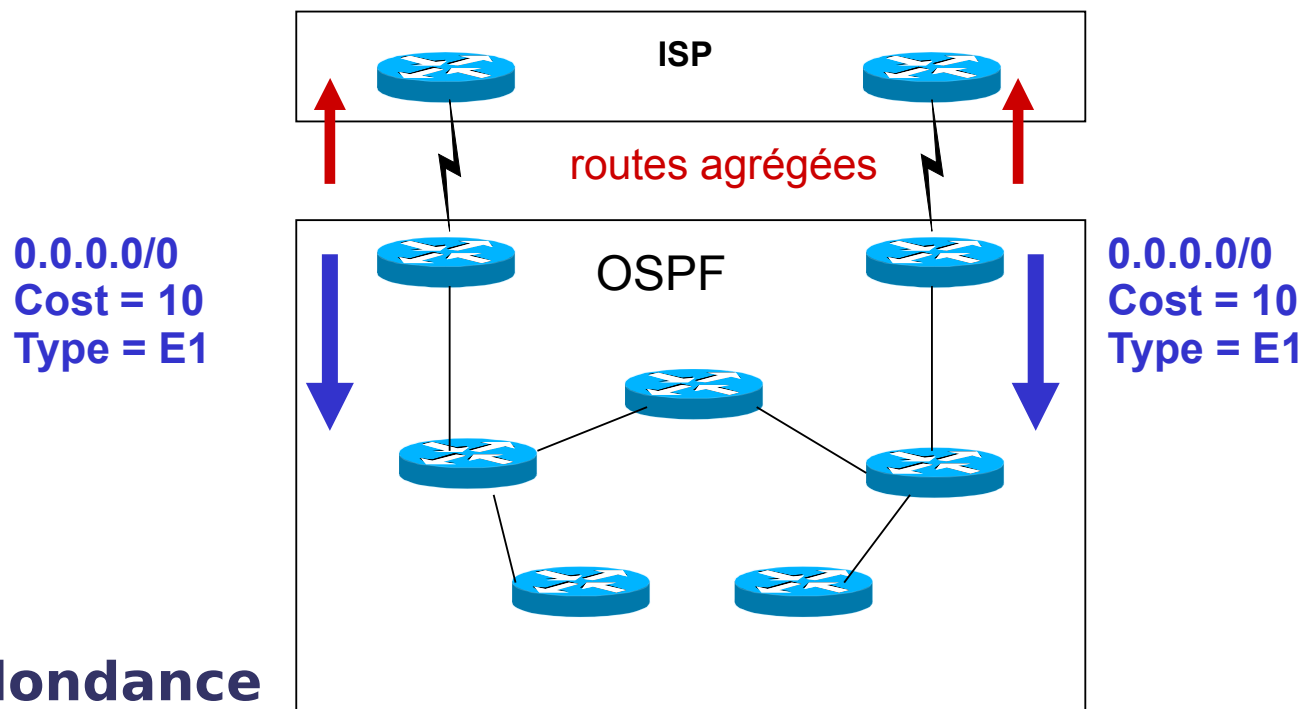


A provider may also choose to dynamically learn a customer's routes using BGP, which typically runs between the ISP router and the customer's boundary router.

Utiliser BGP

- ▶ le réseau single-homed peut utiliser BGP pour incrémenter les politiques de routage
 - ▶ difficilement on obtient un AS, il faut utiliser un numéro privé
 - ▶ Le FAI remplace le numéro privé pour l'annoncer au coeur de l'Internet

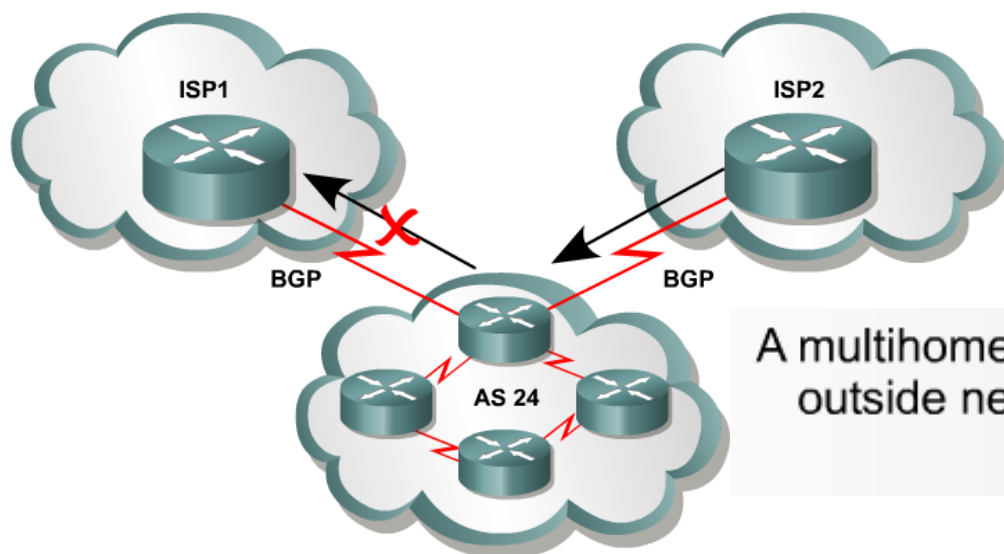
Des AS Multi-homed



AS Multi-homed

- ▶ On cherche de la **redondance**
- ▶ Options
 - ▶ Liens **primaire** et **backup**
 - ▶ Possibilité d'utiliser les deux chemins
- ▶ BGP n'est pas encore obligatoire
- ▶ Risque de devenir un AS de passage lors d'une défaillance

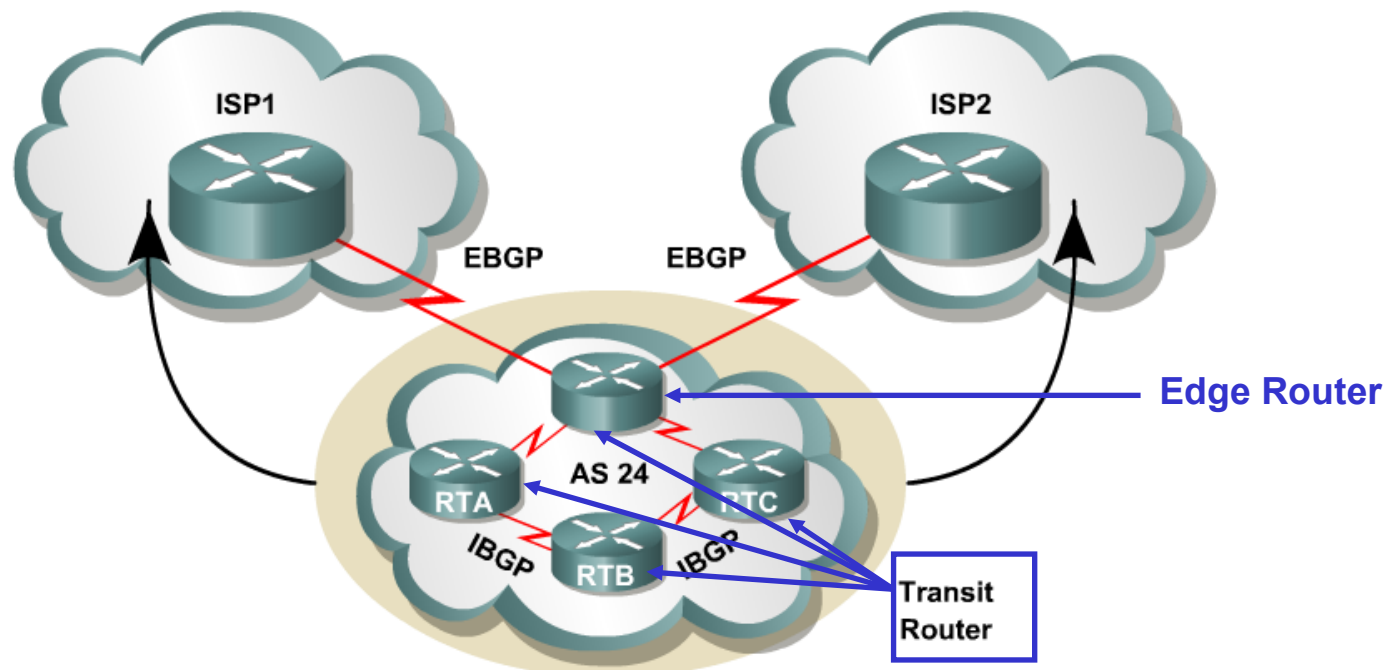
AS Multi-homed : avec ou sans transit de messages



A multihomed nontransit AS features more than one exit point to outside networks, but does not allow traffic to pass from one outside connection to another.

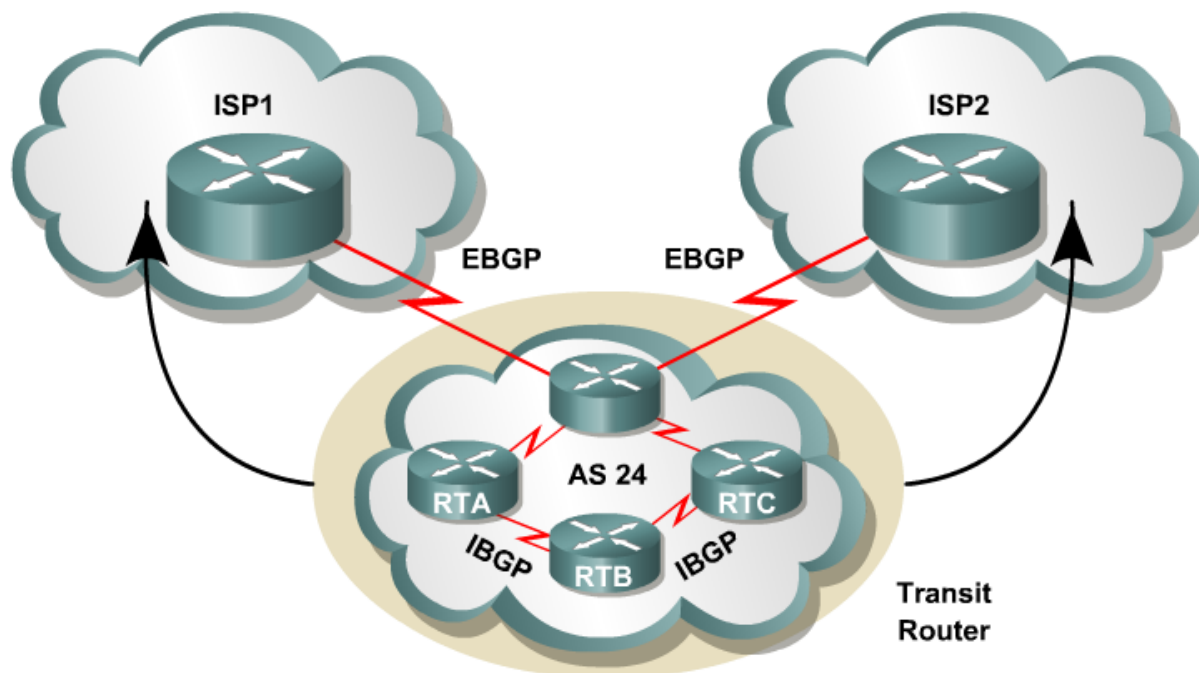
- ▶ **Transit AS** - permet le passage du trafic entre deux autres AS
- ▶ **Non-transit AS** - empêche ce passage
 - ▶ Annonce uniquement SES routes aux deux autres AS
 - ▶ Dans l'exemple, ISP1 n'utilise pas AS 24 pour atteindre ISP2 ou vice-versa
 - ▶ Pas obligatoire d'utiliser BGP, mais fortement recommandé
- ▶ BGP offre plusieurs avantages dont :
 - ▶ Contrôle accru de la propagation des routes
 - ▶ Filtrage des routes

Multi-homed Transit Autonomous Systems



- ▶ **Internal BGP (IBGP)** - lorsque BGP est utilisé à l'intérieur d'un AS
- ▶ **External BGP (EBGP)** - lorsque BGP est utilisé entre deux AS
- ▶ **Routeur de passage** - un routeur BGP utilisé pour le passage du trafic IBGP
- ▶ **Routeur frontalier** - routeurs à la limite des AS utilisés pour échanger des routes EBGP

Les risques d'utilisation de BGP



Voisinage BGP

- ▶ Créer un voisinage BGP requiert l'accord entre les administrateurs
 - ▶ vous faites confiance à ce que l'autre admin fait dans son réseau
 - ▶ au même temps, vous faites tout le possible pour prévenir les erreurs commis par les autres admins
- ▶ "Paranoia is your friend."

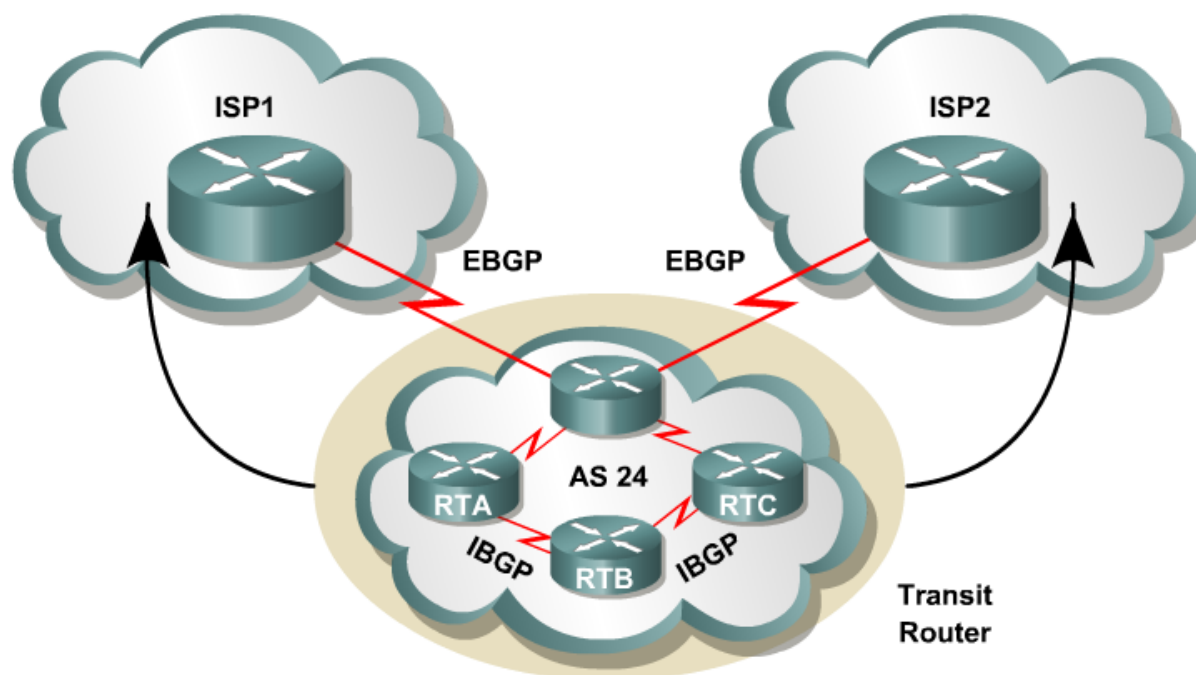
Les risques de BGP – appropriation de routes

- ▶ Le cas du FAI pakistanais et de YouTube
 - ▶ Le dimanche 24 février 2008, Pakistan Telecom (AS17557) a annoncé sans autorisation le préfixe 208.65.153.0/24.
 - ▶ Selon l'ambassade du Pakistan, c'était une mesure gouvernementale destinée à bloquer l'accès de YouTube en Pakistan
- ▶ L'un des FAI de Pakistan Telecom (PCCW Global - AS3491) a transmis cet annonce sur l'Internet, ce qui a résulté dans le détournement global du trafic destiné à YouTube
- ▶ **Le début**
 - ▶ AS36561 (**YouTube**) annonce **208.65.152.0/22**.
- ▶ **Dimanche 24 février 2008, 18:47 (UTC) :**
 - ▶ AS17557 (**Pakistan Telecom**) commence à annoncer **208.65.153.0/24**
 - ▶ AS3491 (PCCW Global – Hong Kong) retransmet l'annonce
 - ▶ Les routeurs du monde entier reçoivent l'annonce et le trafic de YouTube est redirigé vers le Pakistan

Suite et fin

- ▶ **Dimanche 24 février 2008, 20:07 (UTC):**
 - ▶ AS36561 (YouTube) annonce **208.65.153.0/24**.
 - ▶ Avec deux préfixes identiques, des politiques de routage BGP telles que le chemin le plus court font leur choix
 - ▶ AS17557 (Pakistan Telecom) continue à recevoir une partie du trafic destiné à YouTube.
- ▶ **Dimanche 24 février 2008, 20:18 (UTC):**
 - ▶ AS36561 (**YouTube**) annonce **208.65.153.128/25 et 208.65.153.0/25** (et aussi /26)
 - ▶ À cause du préfixe plus long, chaque routeur qui reçoit cet annonce choisit de s'adresser à YouTube
- ▶ **Dimanche 24 février 2008, 20:51 (UTC):**
 - ▶ Pakistan Telecom via AS3491 (PCCW Global) annonce les préfixes /25 et /26
 - ▶ La plupart des routeurs continu à s'adresser à YouTube à cause de la longueur du chemin
- ▶ **Dimanche 24 février 2008, 21:01 (UTC):**
 - ▶ **AS3491 (PCCW Global) retire tous les préfixes annoncés par AS17557** (Pakistan Telecom), finissant le détournement des routes.

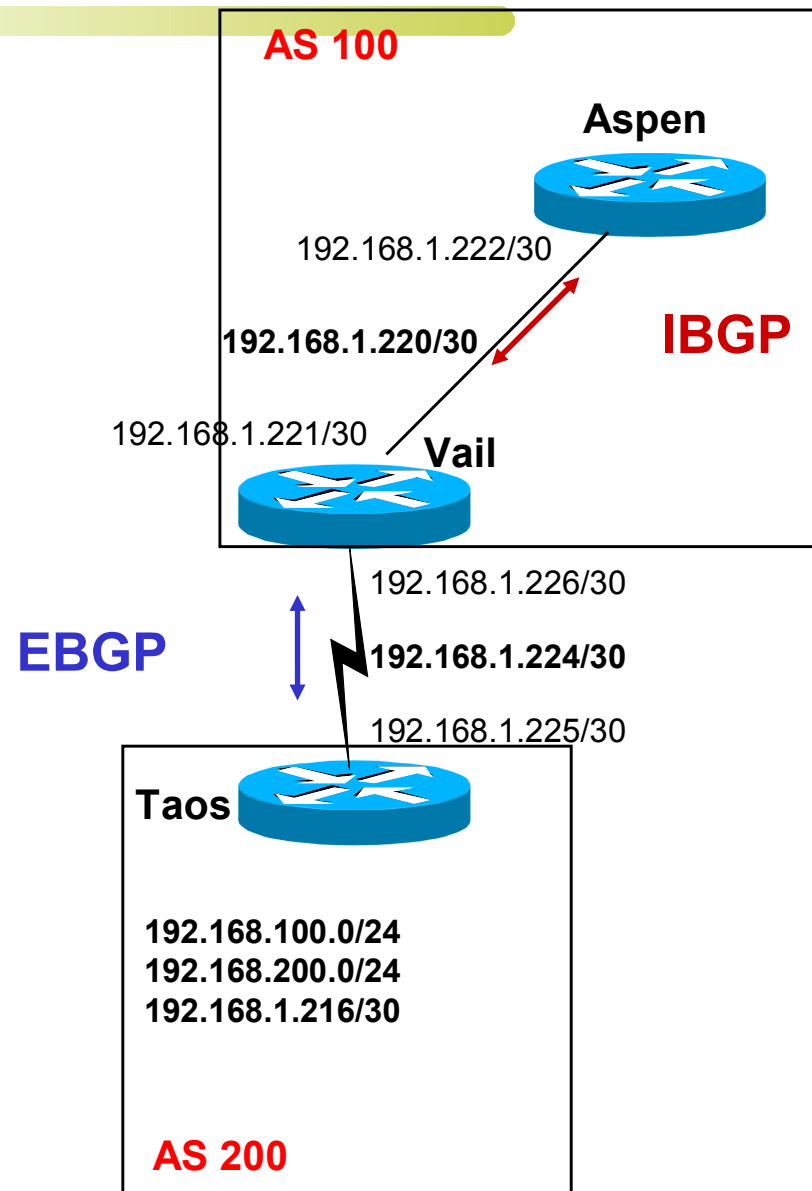
Risques avec BGP – Devenir un AS de passage



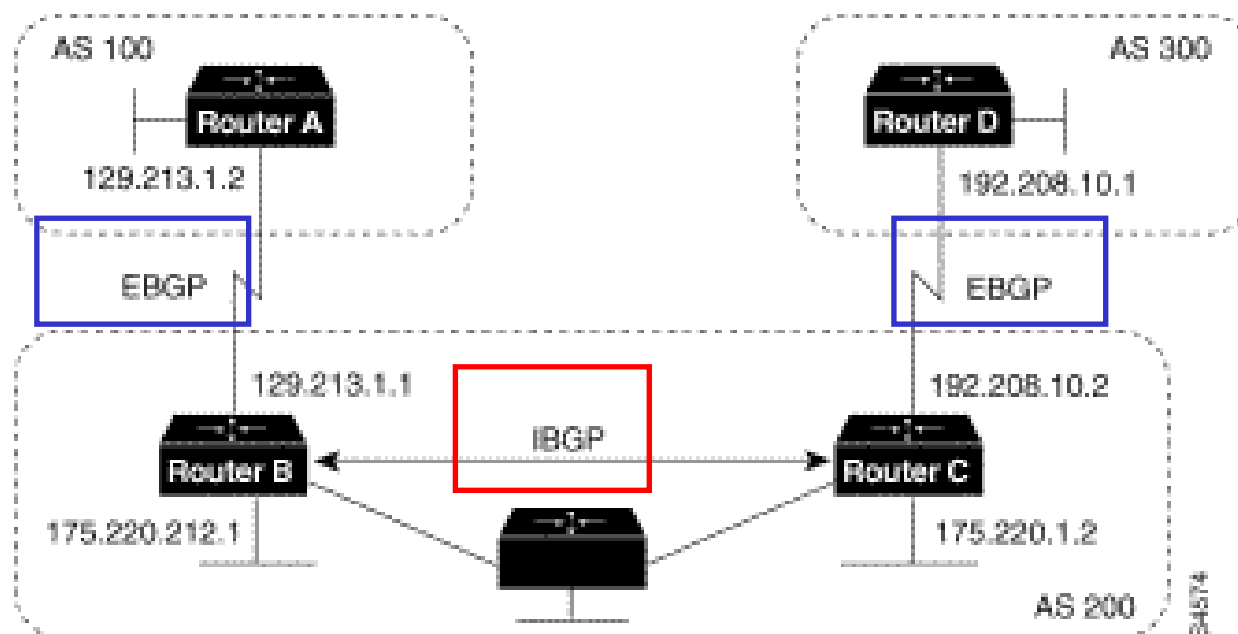
- ▶ Si AS24 annonce par mégarde des routes apprises sur ISP1 vers ISP2
- ▶ Les clients ISP2 considèrent AS24 comme une route vers ISP1
- ▶ AS24 devient un domaine de transit pour ISP1 et ISP2

IBGP versus EBGP

- ▶ Lorsque BGP s'exécute à l'intérieur d'un AS, il est appelé **Internal BGP (IBGP)**.
 - ▶ Un routeur de transit exécute IBGP
 - ▶ Utilisé presque exclusivement dans le cas des AS multi-homed
- ▶ BGP entre deux AS est appelé **External BGP (EBGP)**
 - ▶ Un routeur de frontière

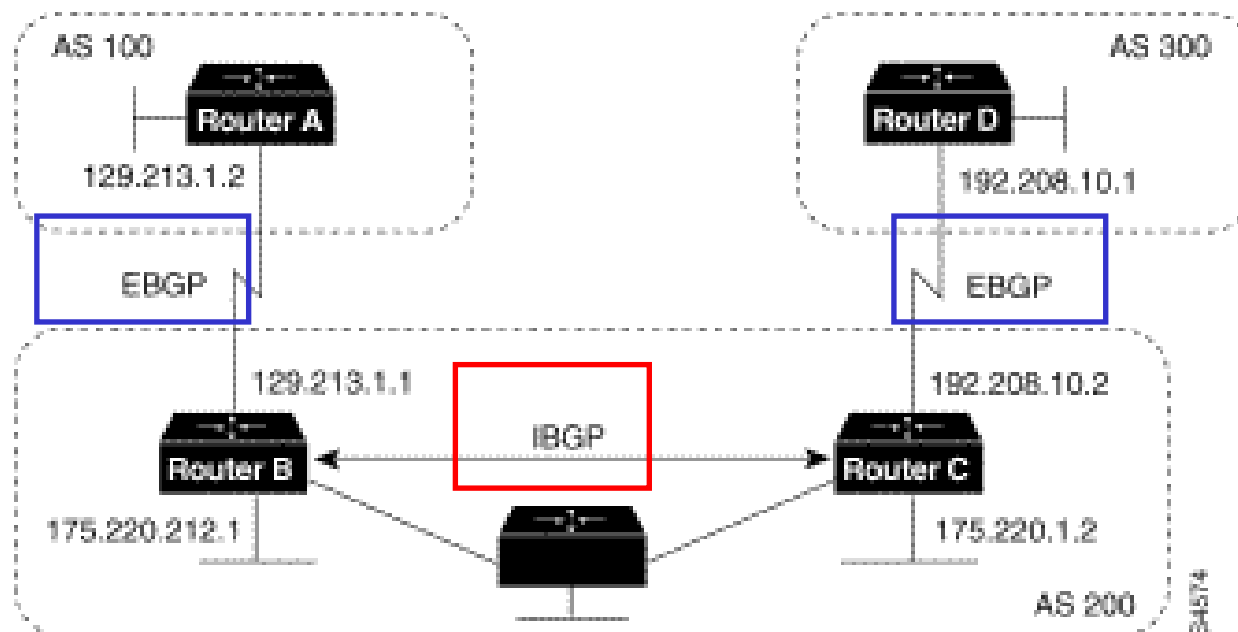


IBGP versus EBGP



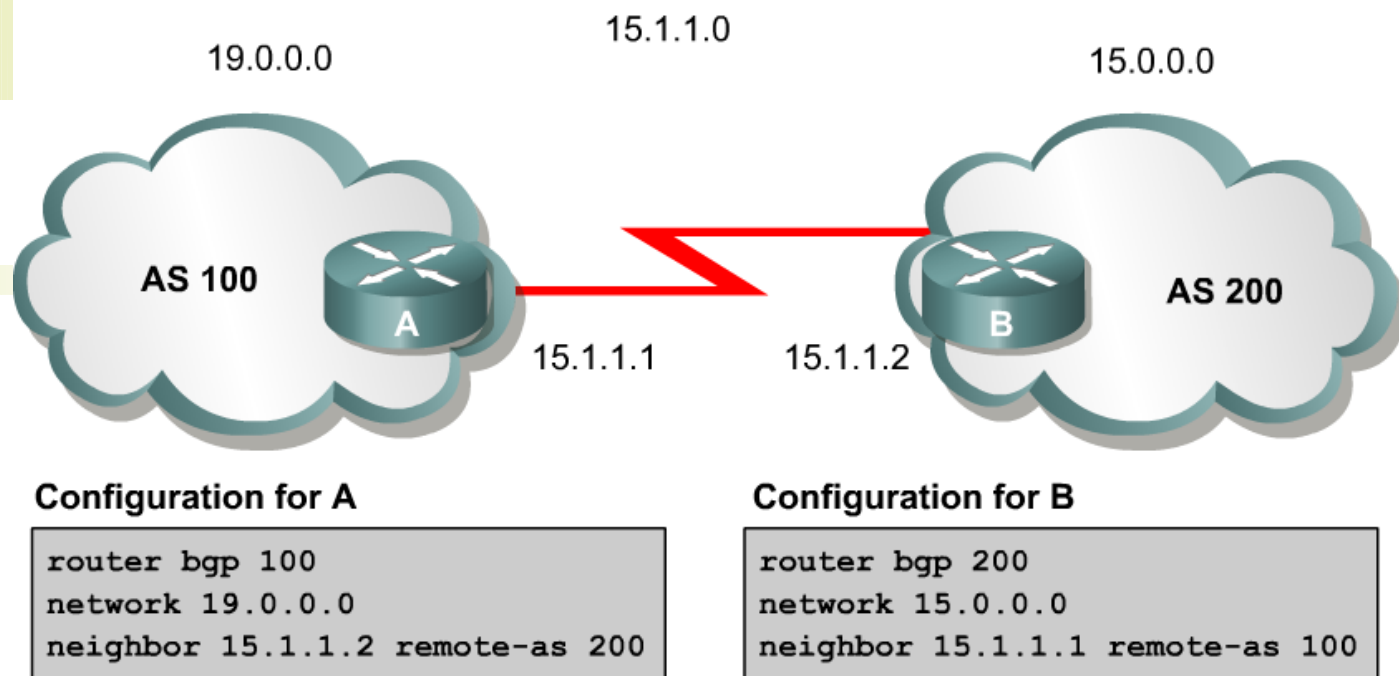
- Les routeurs A et B tournent **EBGP (BGP)** alors que les routeurs B et C tournent **IBGP**.
- On observe que :
 - ▶ les pairs **EBGP (BGP)** doivent être directement connectés
 - ▶ les pairs **IBGP** n'ont pas besoin d'être directement connectés
 - il suffit d'avoir un protocole IGP pour router les messages à l'intérieur d'un AS

IBGP versus EBGP



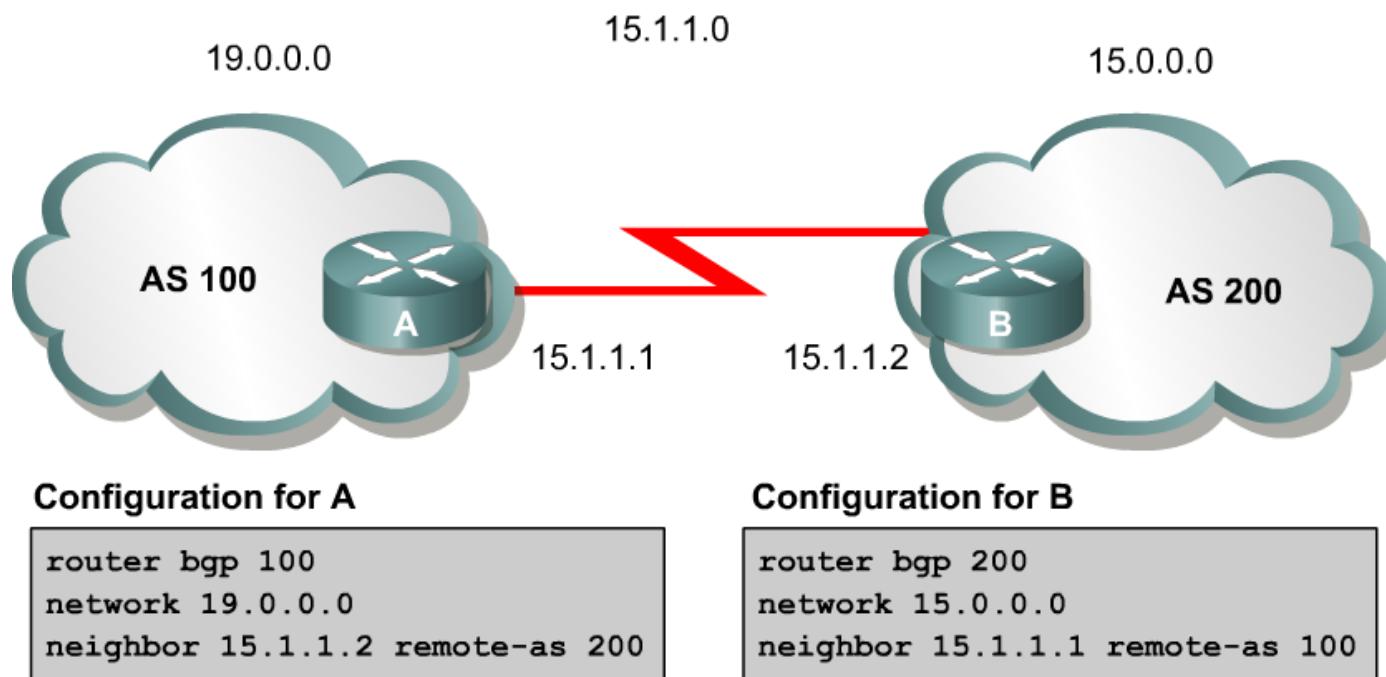
- Tout noeud **BGP** à l'intérieur d'un AS doit établir une connexion entre-eux
- Exemple : l'AS 200 est un **AS de passage** pour AS 100 et AS 300

Configurer BGP



- Pour commencer, entrer dans le mode configuration BGP
Router(config)#**router bgp AS-number**
- Attention, les commandes BGP sont presque similaires à ceux des IGP
 - Leur fonction peut être assez différente
- **Note** : L'IOS cisco ne permet qu'une seule processus BGP par routeur
 - alors un routeur n'appartient qu'à un seul AS

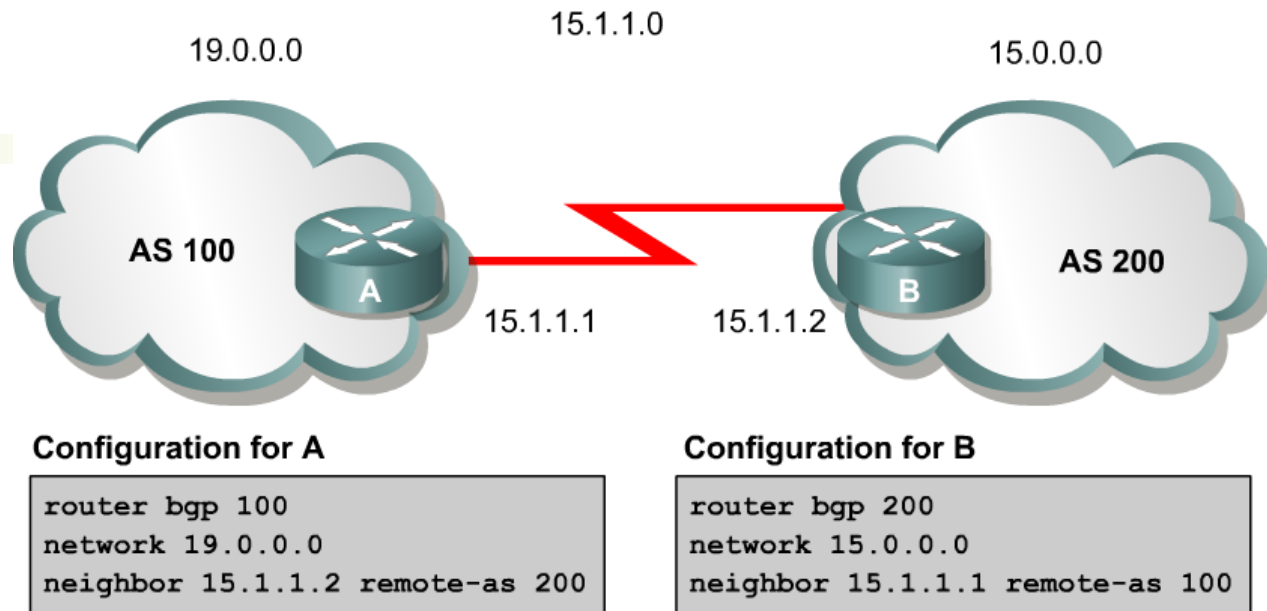
Configurer BGP



Router(config-router) #**network network-number [mask network-mask]**

- ▶ Les réseaux doivent être présents dans la table de routage (**show ip route**) pour être distribués
- ▶ Les routes apprises avec BGP sont diffusées par défaut, mais souvent elles sont filtrées par des politiques de routage

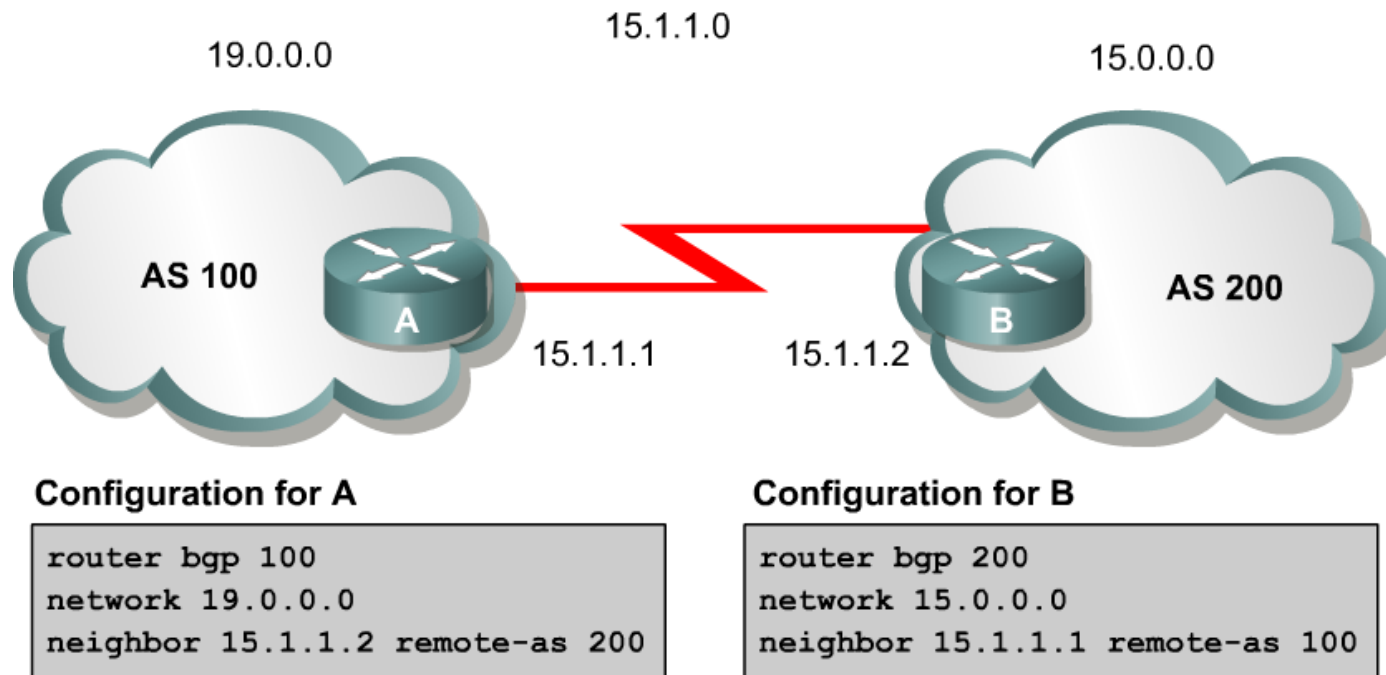
Configurer BGP



Router (config-router) #**network network-number [mask network-mask]**

- ▶ Avec les **IGPs**, la commande **network** indique les interfaces qui participent ainsi que les routes annoncées
- ▶ Dans le cas du protocole **BGP**, la commande **network** n'affecte pas les interfaces
- ▶ En BGP, **network** indique les routes apprises localement qui seront annoncées
- ▶ Les réseaux peuvent être des routes connectées, statiques ou apprises dynamiquement
 - ▶ Seulement configurer **network** ne suffit pas pour établir des connexions BGP

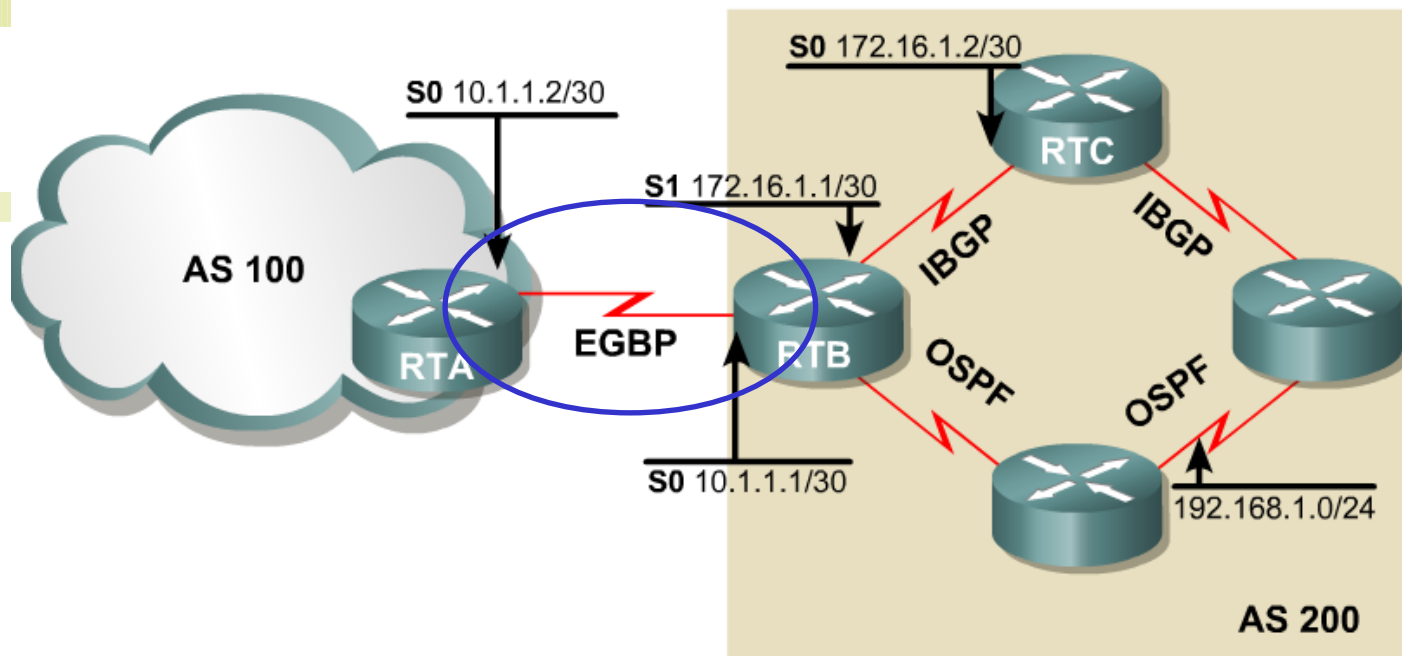
Configurer BGP



Router(config-router)#**neighbor ip-address remote-as AS-number**

- ▶ **La commande Neighbor** – Utilisée pour **établir une relation de voisinage avec un autre routeur BGP**
- ▶ Le paramètre **AS-number** indique si le routeur est un voisin EBGP ou IBGP

EBGP

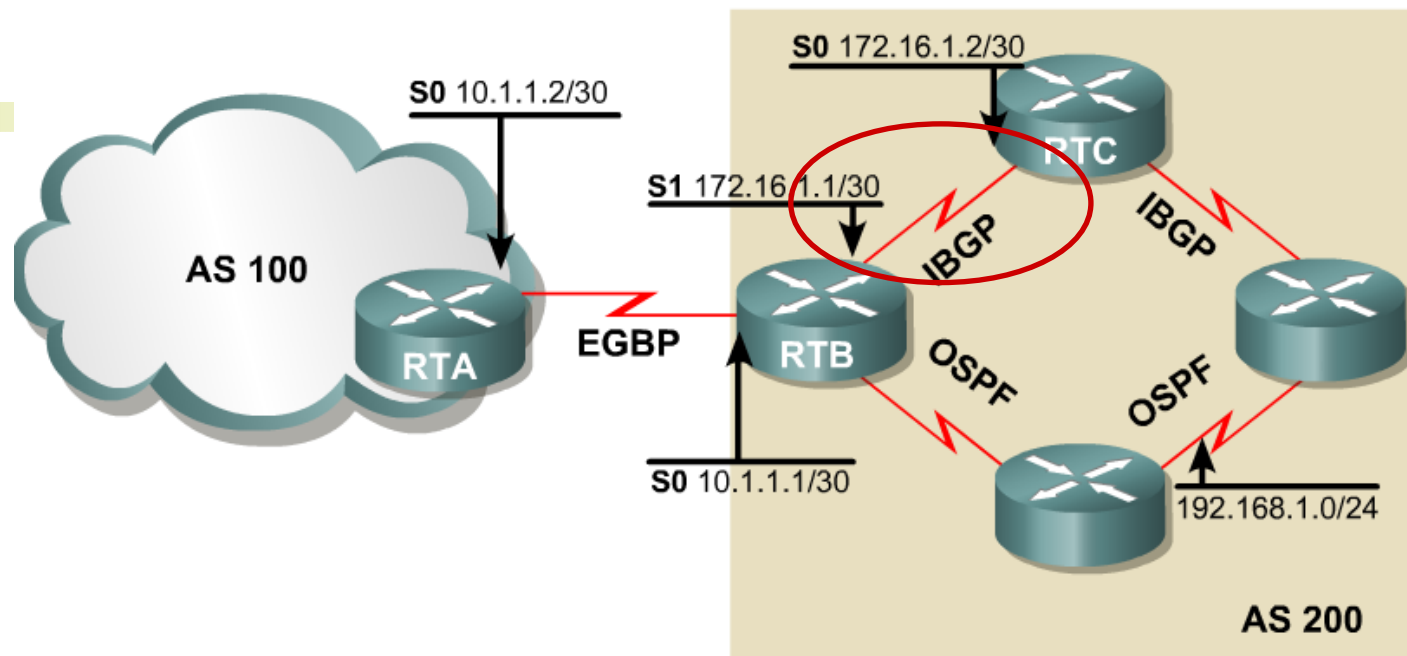


```
RTA(config)#router bgp 100
RTA(config-router)#neighbor 10.1.1.1 remote-as 200
```

```
RTB(config)#router bgp 200
RTB(config-router)#neighbor 10.1.1.2 remote-as 100
```

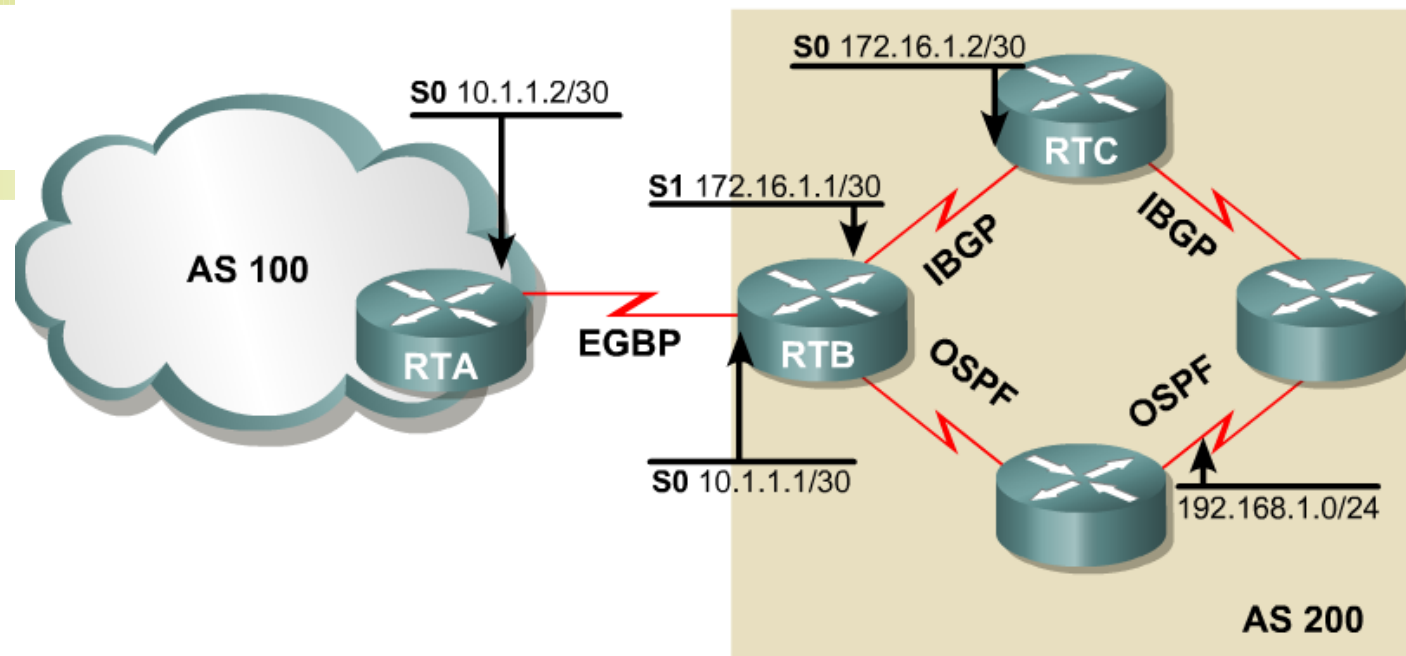
- Observez que la valeur **remote-as** 100 de la commande **neighbor** est différente de la valeur AS-number spécifiée par la commande **router bgp** (200)
- Avec deux numéros AS différents, RTB initie une connexion **EBGP** avec RTA

IBGP



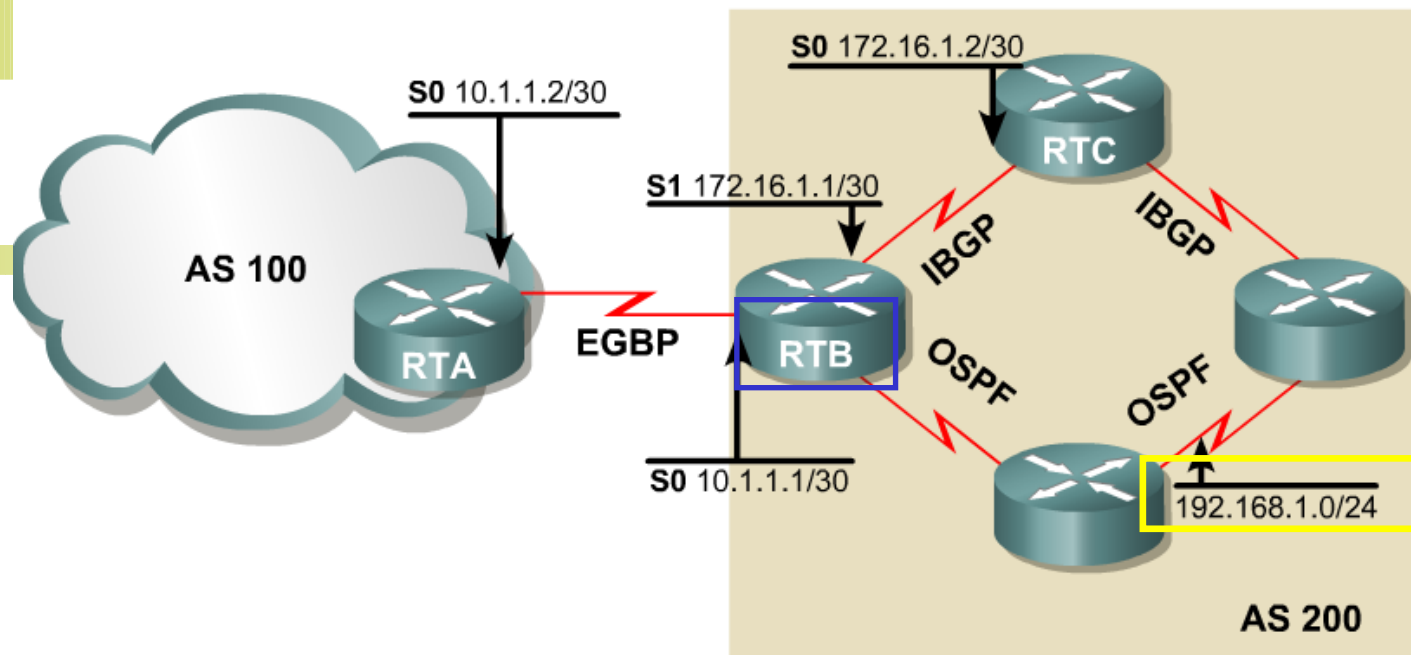
```
RTB(config)#router bgp 200
RTB(config-router)#neighbor 172.16.1.2 remote-as 200
RTB(config-router)#neighbor 172.16.1.2 update-source loopback 0
```

```
RTC(config)#router bgp 200
RTC(config-router)#neighbor 172.16.1.1 remote-as 200
RTC(config-router)#neighbor 172.16.1.1 update-source loopback 0
```



```
RTB(config-router)#neighbor 172.16.1.2 update-source loopback 0
RTC(config-router)#neighbor 172.16.1.1 update-source loopback 0
```

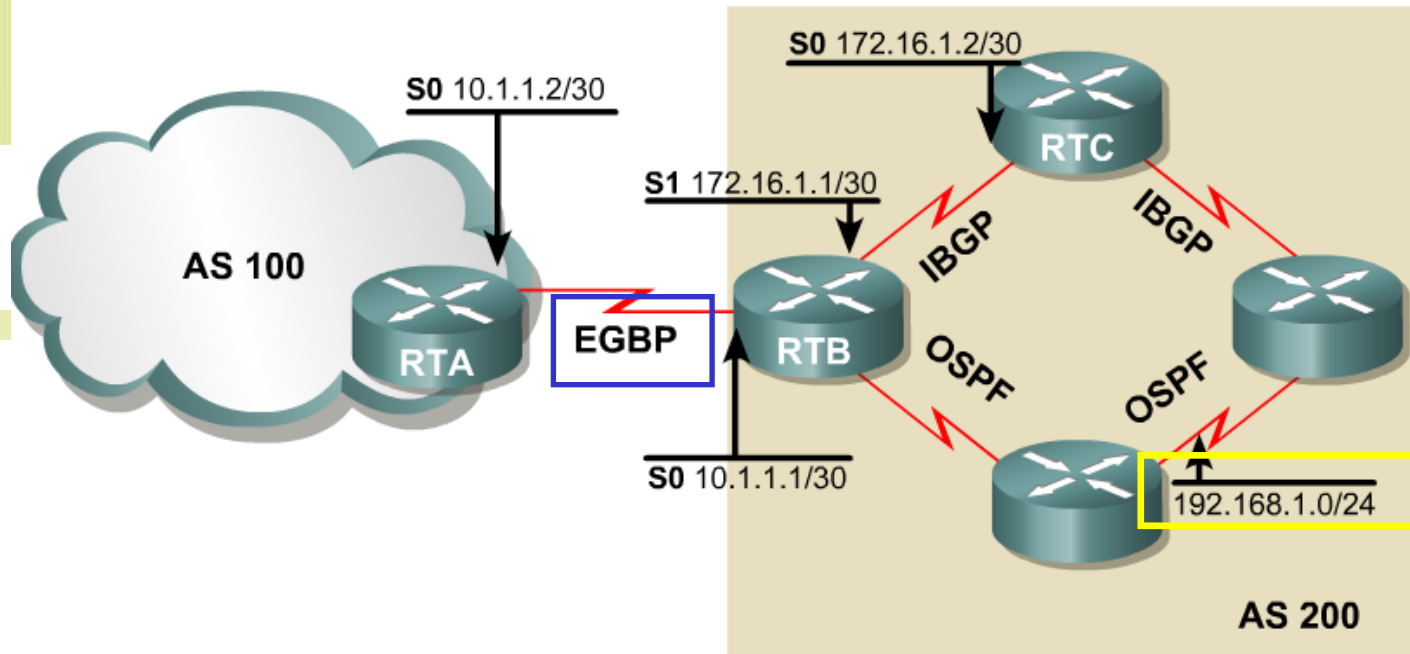
- **update-source loopback 0** - indique au routeur d'utiliser toute interface opérationnelle pour établir les connexions TCP (tant que Lo0 est active et configurée avec une adresse IP)
- Sans cette commande **update-source loopback 0**, les routeurs BGP ne peuvent utiliser que l'interface la plus proche
- Utiliser n'importe quelle interface augmente la robustesse de BGP si un lien est défaillant



- Considérez que la route suivante est entrée dans la table de RTB

`0 192.168.1.0/24 [110/74] via 10.2.2.1, 00:31:34, Serial2`

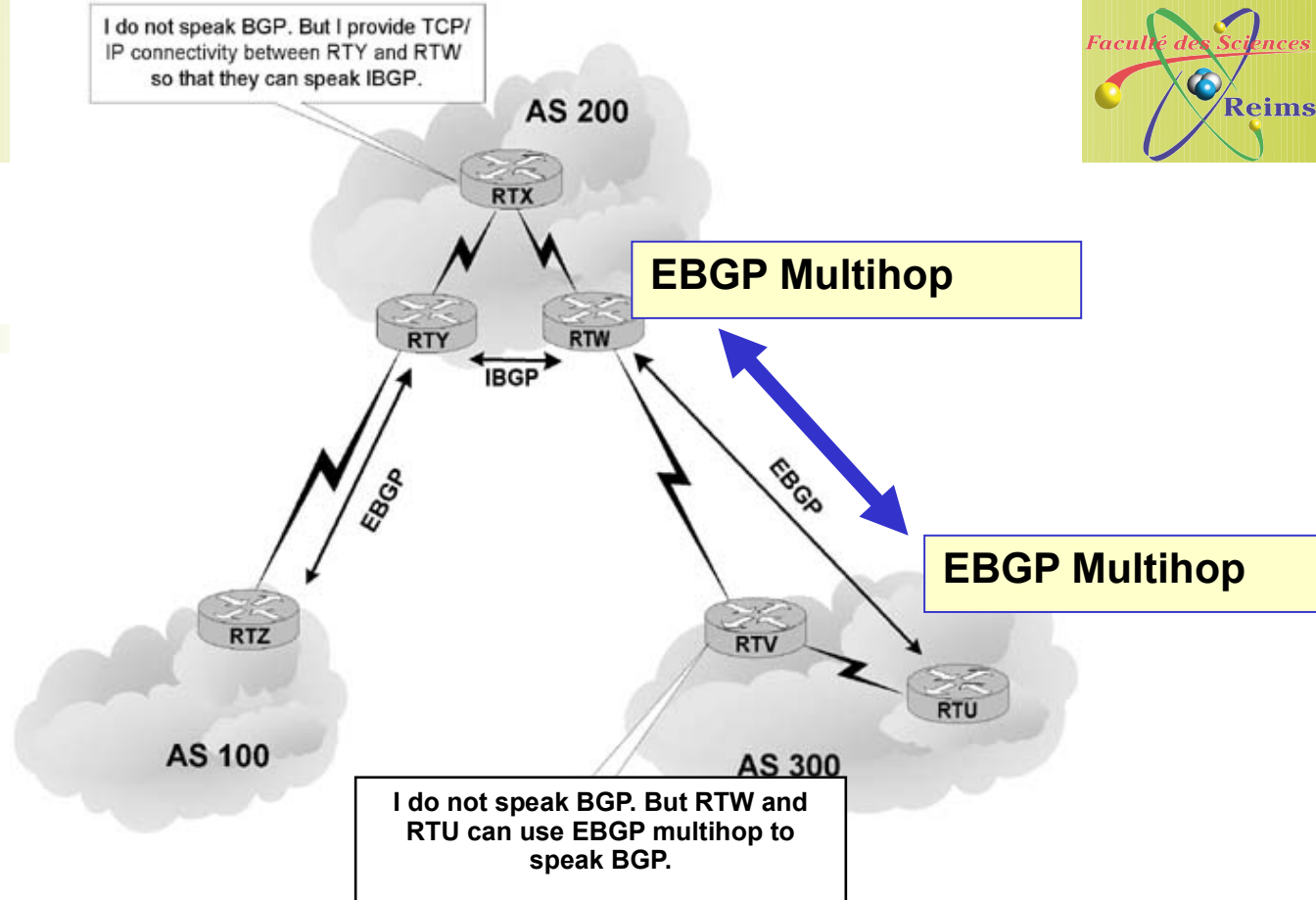
- RTB a appris cette route grâce à un IGP (OSPF)
- Cet AS utilise OSPF pour échanger les routes internes
- Comment RTB pourra annoncer cette route via BGP ?
 - Avec la redistribution des routes OSPF sur BGP
- ou
- Avec la commande BGP **network**



```
RTB(config)#router bgp 200
RTB(config-router)#network 172.16.1.0 mask 255.255.255.254
RTB(config-router)#network 10.1.1.0 mask 255.255.255.254
RTB(config-router)#network 192.168.1.0
```

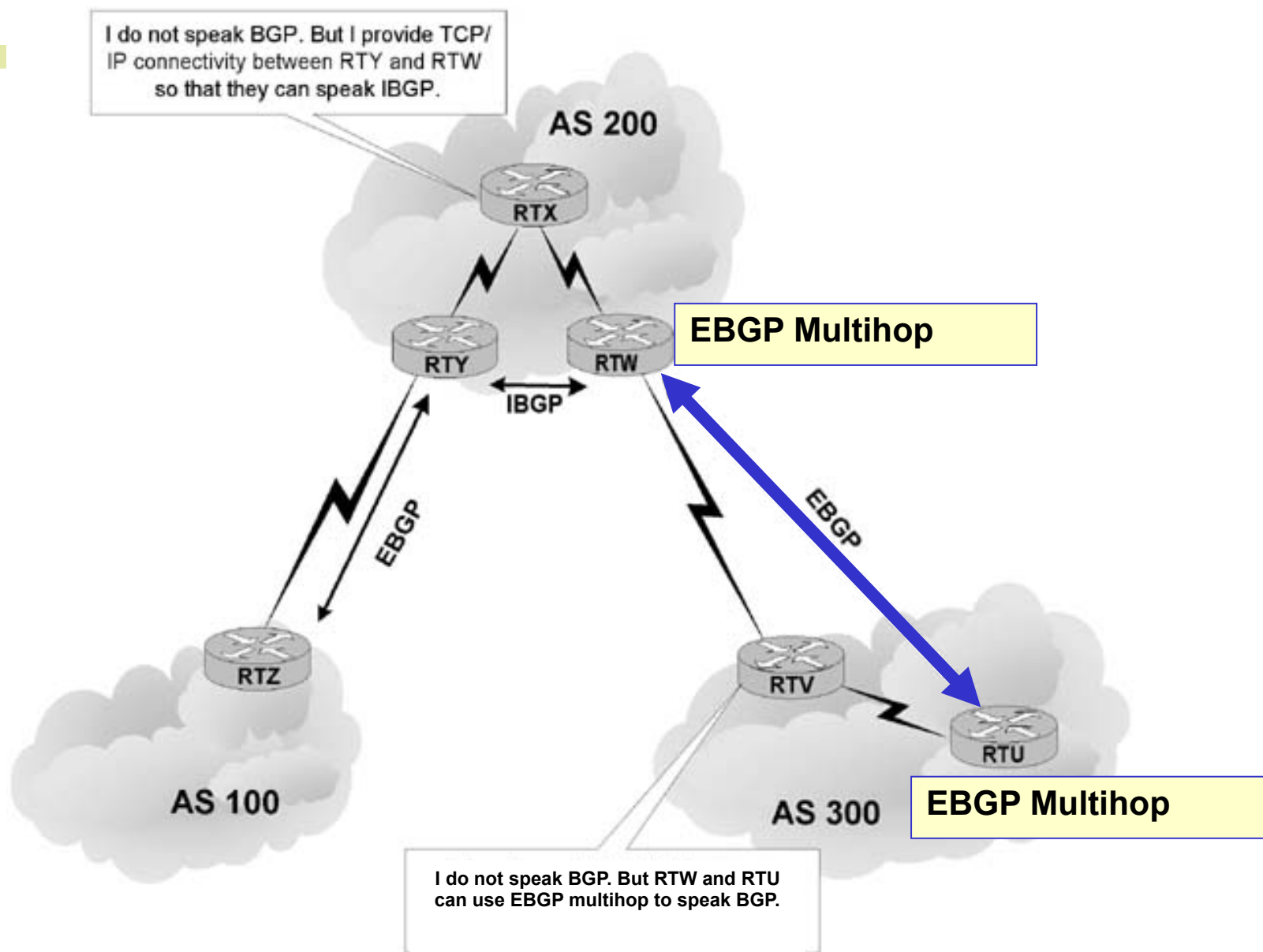
- ▶ Les deux premières commandes **network** indiquent le **masque** d'un sous-réseau spécifique
- ▶ La troisième commande **network** annonce le réseau 192.168.1.0 sans avoir besoin d'une redistribution OSPF
 - ▶ "route statique"
- ▶ Attention : network ici indique les routes qui seront annoncés sur toute connexion BGP

EBGP Multihop



- **EBGP** doit connecter directement les pairs pour établir une session
- Exception : l'IOS Cisco a une option **multihop** pour permettre la connexion logique entre RTW et RTU même si RTV n'implémente pas BGP
 - équivalent à augmenter le TTL d'une annonce BGP
- Cette option doit être configurée pour chaque pair

```
Router(config-router)#neighbor IP-address ebgp-multihop
[hops]
```

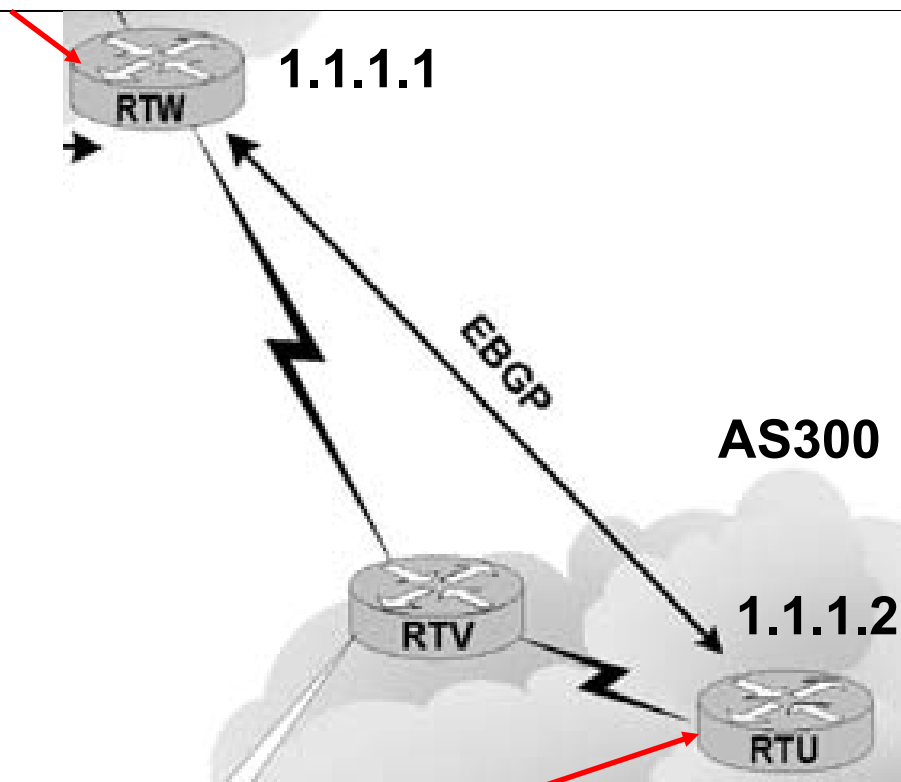


```
RTW(config)#router bgp 200
```

```
RTW(config-router)#neighbor 1.1.1.2 remote-as 300
```

```
RTW(config-router)#neighbor 1.1.1.2 ebgp-multihop 2
```

AS200

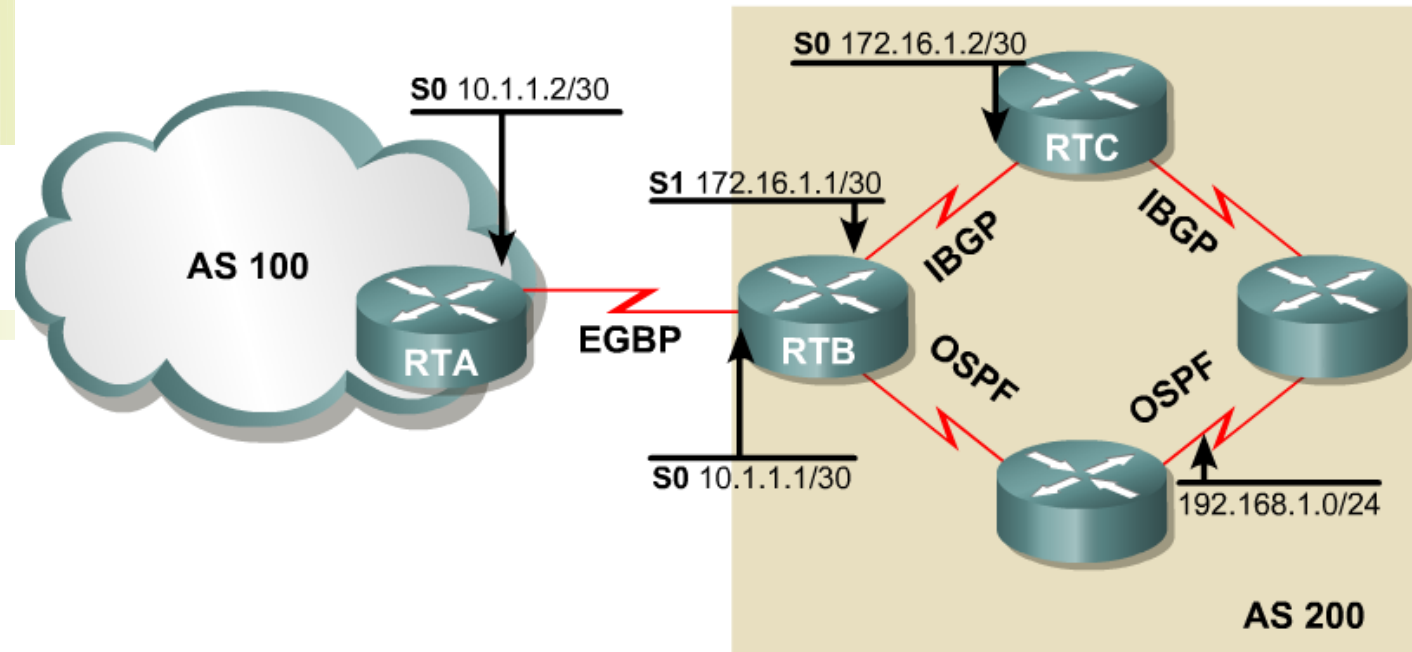


```
RTU(config)#router bgp 300
```

```
RTU(config-router)#neighbor 1.1.1.1 remote-as 200
```

```
RTU(config-router)#neighbor 1.1.1.1 ebgp-multihop 2
```

Configurer BGP



- ▶ Lors de la configuration de BGP, les modifications ne sont pas immédiatement visibles
- ▶ Pour forcer BGP de reinitialiser ses tables, utiliser la commande **clear ip bgp**

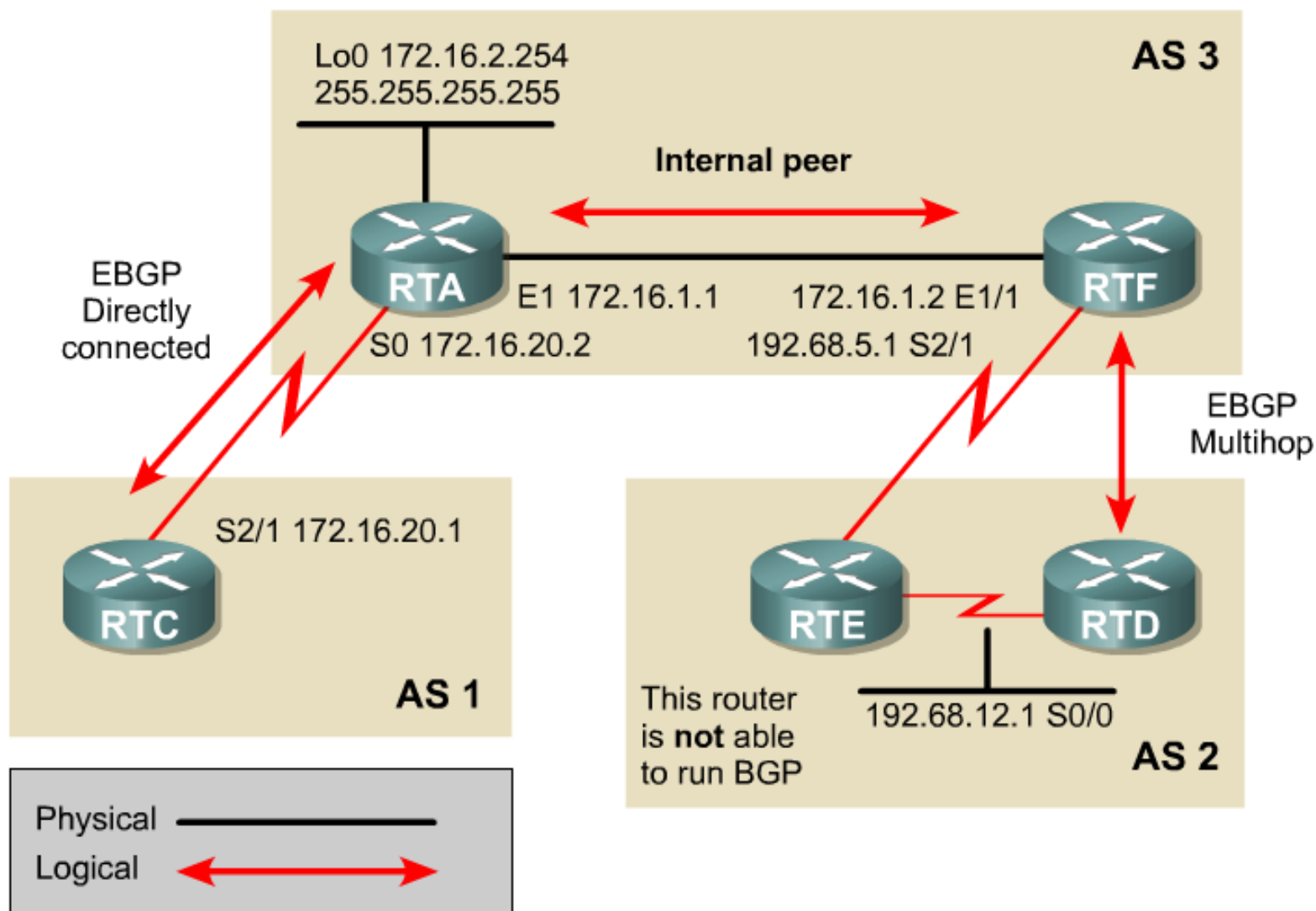
```
Router#clear ip bgp *
```

```
Router#clear ip bgp 10.0.0.0
```

Utilisez cette commande avec beaucoup d'attention, voir pas du tout sur un réseau de production

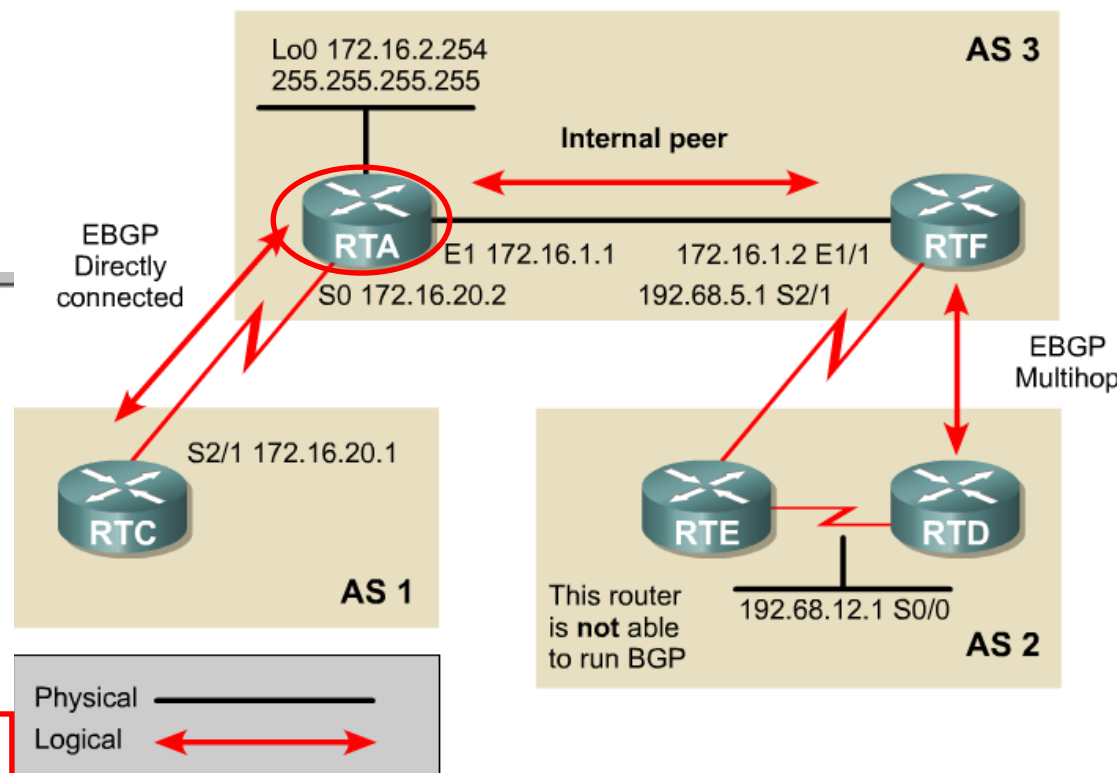
- ▶ “clear ip bgp * OOPS! Not me but a colleague who was an employee of a large ISP with a 3 letter title. Got back from a Cisco routing course and thought they would try out some commands on the core network. It took 45 minutes for the core to reconverge. P45 followed”

Exemple



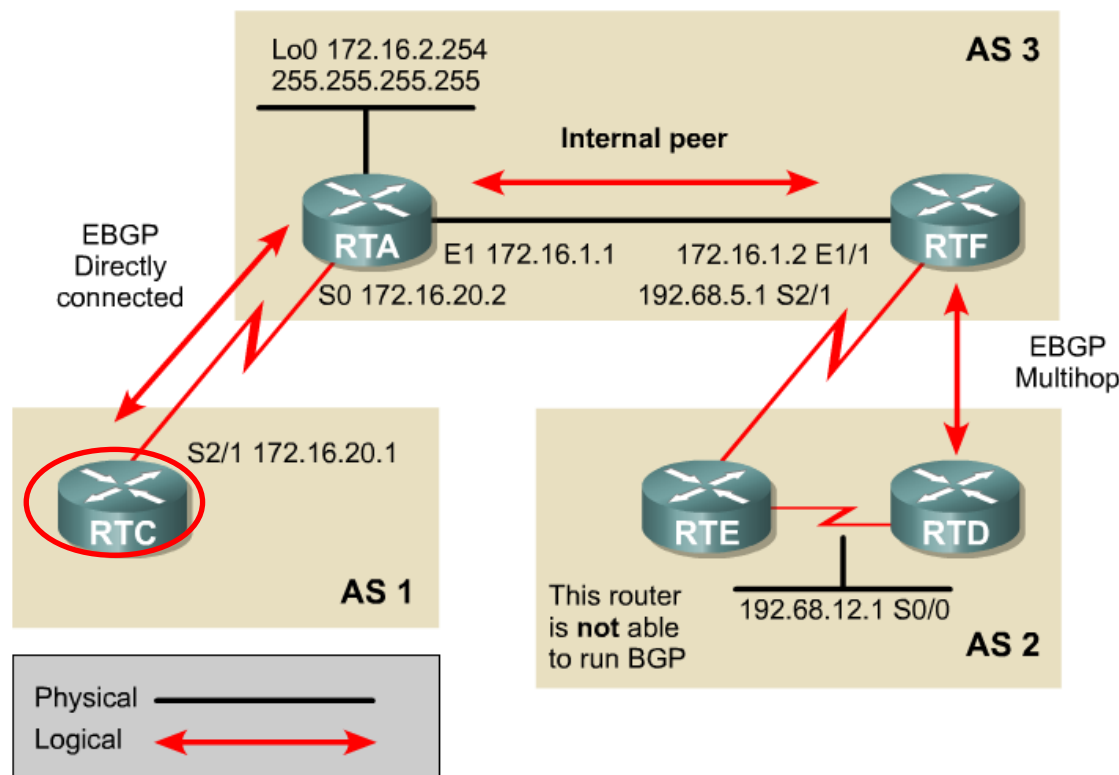
Exemple

```
RTA#show running-config
ip subnet-zero
interface Loopback0
ip address 172.16.2.254 255.255.255.255
interface Ethernet1
ip address 172.16.1.1 255.255.255.0
interface Serial0
ip address 172.16.20.2 255.255.255.0
router ospf 10
network 172.16.0.0 0.0.255.255 area 0
router bgp 3
no synchronization
neighbor 172.16.1.2 remote-as 3
neighbor 172.16.20.1 remote-as 1
no auto-summary
ip classless
RTA#
```



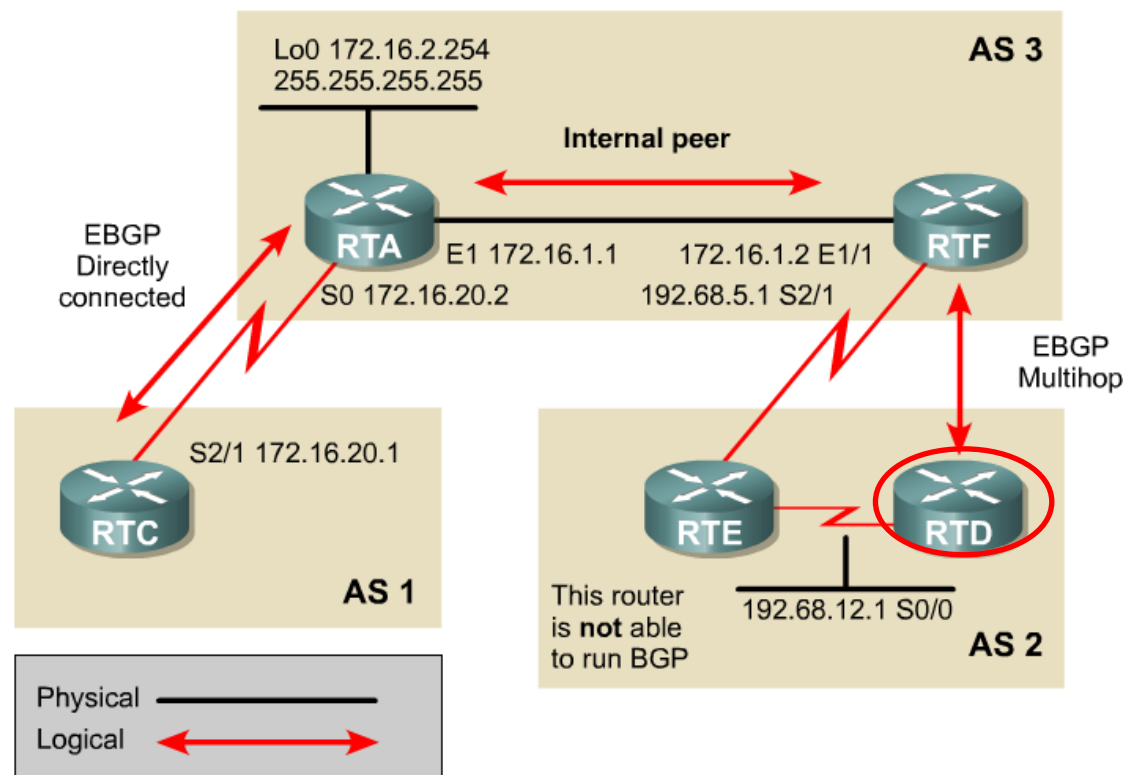
Exemple

```
RTC#show running-config
ip subnet-zero
interface Serial2/1
ip address 172.16.20.1 255.255.255.0
router bgp 1
neighbor 172.16.20.2 remote-as 3
no auto-summary
ip classless
RTC#
```

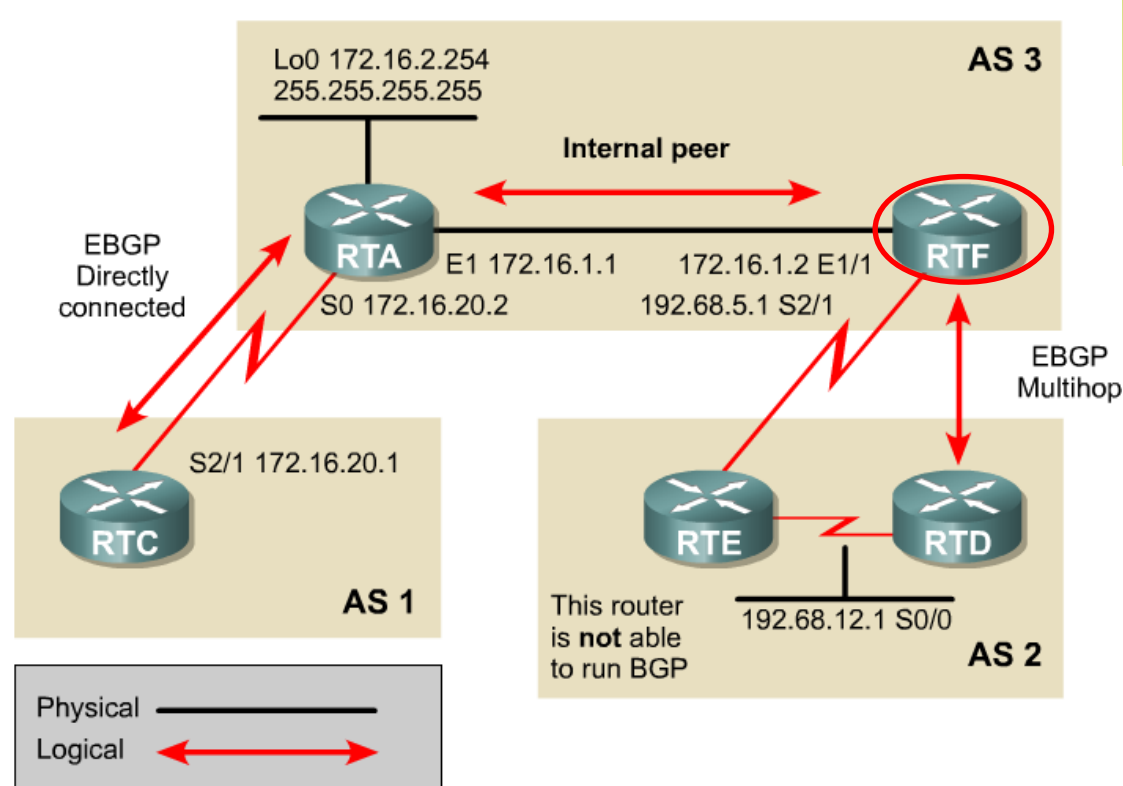


Exemple

```
RTD#show running-config
ip subnet-zero
interface Serial0/0
ip address 192.68.12.1 255.255.255.0
router ospf 10
network 192.68.0.0 0.0.255.255 area 0
router bgp 2
neighbor 192.68.5.1 remote-as 3
neighbor 192.68.5.1 ebgp-multihop 2
no auto-summary
ip classless
RTD#
```




```
RTF#show running-config
ip subnet-zero
interface Ethernet1/1
ip address 172.16.1.2 255.255.255.0
interface Serial2/1
ip address 192.68.5.1 255.255.255.0
router ospf 10
network 172.16.0.0 0.0.255.255 area 0
network 192.68.0.0 0.0.255.255 area 0
router bgp 3
no synchronization
neighbor 172.16.2.254 remote-as 3
neighbor 192.68.12.1 remote-as 2
neighbor 192.68.12.1 ebgp-multihop 2
no auto-summary
ip classless
RTF#
```



```
RTF#show ip bgp neighbor
BGP neighbor is 172.16.2.254, remote AS 3, internal link
BGP version 4, remote router ID 172.16.2.254
BGP state = Established, table version = 2, up for 22:36:09
Last read 00:00:10, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 5 seconds
Received 1362 messages, 0 notifications, 0 in queue
Sent 1362 messages, 0 notifications, 0 in queue
Connections established 2; dropped 1
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 172.16.1.2, Local port: 11008
Foreign host: 172.16.2.254, Foreign port: 179
BGP neighbor is 192.68.12.1, remote AS 2, external link
BGP version 4, remote router ID 192.68.5.2
BGP state = Established, table version = 2, up for 22:13:01
Last read 00:00:00, hold time is 180, keepalive interval is 60 seconds
Minimum time between advertisement runs is 30 seconds
Received 1336 messages, 0 notifications, 0 in queue
```

Comment vérifier la config BGP

- Si le routeur n'a pas installé les routes prévues, on peut utiliser la commande **show ip bgp** pour vérifier les routes que BGP a appris

```
RTA#show ip bgp
```

```
BGP table version is 3, local router ID is 10.2.2.2
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
* i1.0.0.0	192.168.1.6	0	100	0 200 400	e
*>i10.1.1.1/32	10.1.1.1	0	100	0	i
*>i172.16.1.0/24	10.1.1.1	0	100	0	i
* i192.168.1.32/27	192.168.1.6	0	100	0 200	i

show ip bgp

```
RouterC#show ip bgp
BGP table version is 8, local router ID is 200.200.200.66
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	11.0.0.0	0.0.0.0	0		32768	i
*>	12.0.0.0	200.200.200.65			0	300 200 i
*>	192.10.2.0	200.200.200.65	0		0	300 i

- **BGP table version** – numéro de version interne, augmente à chaque update
- **local router ID** – adresse IP du routeur
- **Status codes** – Statut des entrées dans la table
 - s —l'entrée a été supprimée
 - * —l'entrée est valide
 - > —l'entrée est le meilleur chemin pour le réseau (path vector)
 - i —l'entrée a été apprise par une session IBGP

show ip bgp

```
RouterC#show ip bgp
BGP table version is 8, local router ID is 200.200.200.66
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 11.0.0.0	0.0.0.0	0		32768	i
*> 12.0.0.0	200.200.200.65			0	300 200 i
*> 193.10.2.0	200.200.200.65	0		0	300 i

- **Origin codes** – L'origine des entrées :
 - i — entrée originée d'un IGP
 - e — entrée originée d'un EGP
 - ? — entrée avec une origine non établie. Normalement est un routeur BGP appris à partir d'un IGP
- **Network** – adresse IP du réseau
- **Next Hop** – adresse IP du prochain saut. Une entrée 0.0.0.0 indique que le routeur a une route non-BGP vers la destination

show ip bgp

```
RouterC#show ip bgp
BGP table version is 8, local router ID is 200.200.200.66
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 11.0.0.0	0.0.0.0	0		32768	i
*> 12.0.0.0	200.200.200.65			0	300 200 i
*> 193.10.2.0	200.200.200.65	0		0	300 i

- **Metric** – Si affichée, indique une métrique du système interautonome
- **LocPrf** – Préférence locale, définie avec la commande **set local-preference**. La valeur défaut est de 100
- **Weight** – Poids d'une route, définie par les filtres de l'AS
- **Path** – Le chemin vers la destination

Comment vérifier la config BGP

- Si un routeur BGP n'apparaît dans les tables BGP, on peut utiliser **show ip bgp neighbors** pour vérifier les connexions établies

```
RTA#show ip bgp neighbors
```

```
BGP neighbor is 172.24.1.18, remote AS 200, external link
```

```
BGP version 4, remote router ID 172.16.1.1
```

```
BGP state = Established, up for 00:03:25
```

```
Last read 00:00:25, hold time is 180, keepalive interval is 60 seconds
```

```
Neighbor capabilities:
```

```
Route refresh: advertised and received
```

```
Address family IPv4 Unicast: advertised and received
```

```
Received 7 messages, 0 notifications, 0 in queue
```

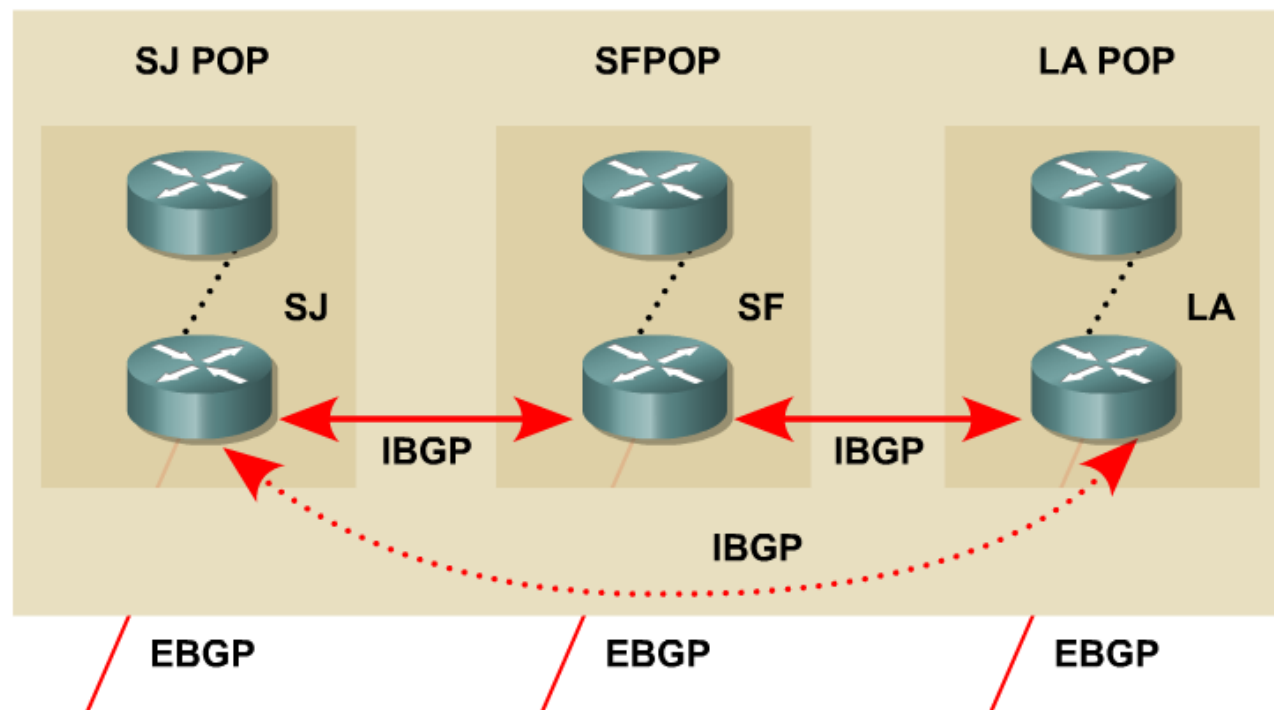
```
Sent 8 messages, 0 notifications, 0 in queue
```

```
Route refresh request: received 0, sent 0
```

```
Minimum time between advertisement runs is 30 seconds
```

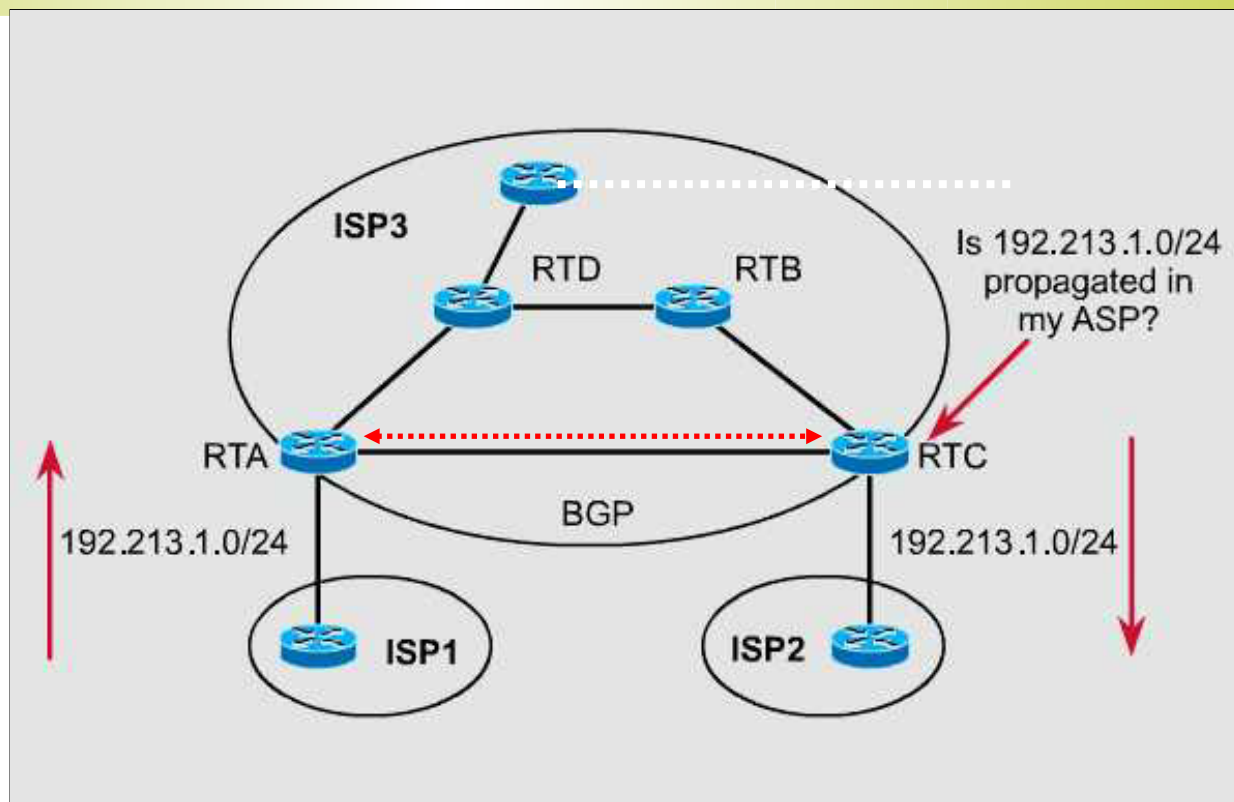
```
<output omitted>
```

Voisinage BGP



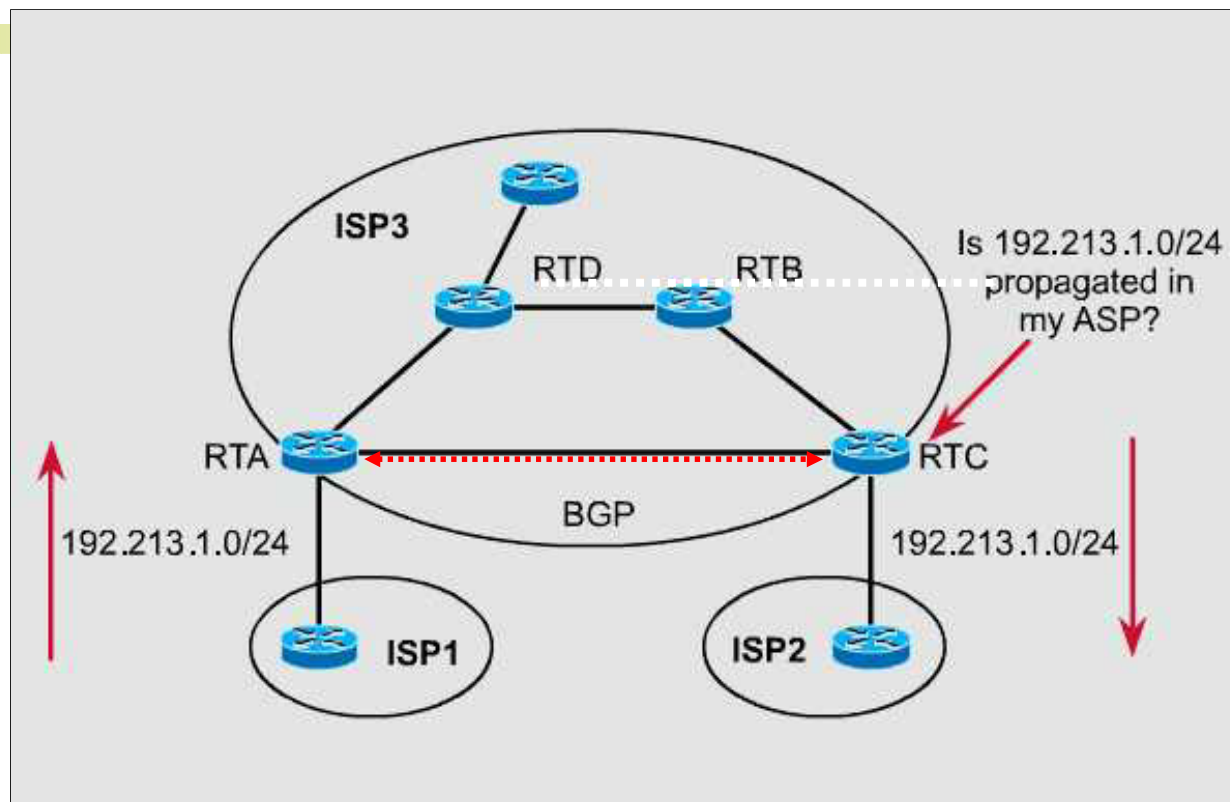
- ▶ **Split Horizon** – Des routes apprises via IBGP ne sont pas propagées vers d'autres noeuds IBGP
- ▶ Cela évite la formation de boucles de routage
- ▶ Pour qu'un routeur IBGP apprenne sur les autres routeurs IBGP à l'intérieur d'un AS, ils doivent être connectés en *full mesh*
 - ▶ Il suffit qu'une connexion TCP soit possible, pas besoin d'être directement connecté

Synchronisation des AS



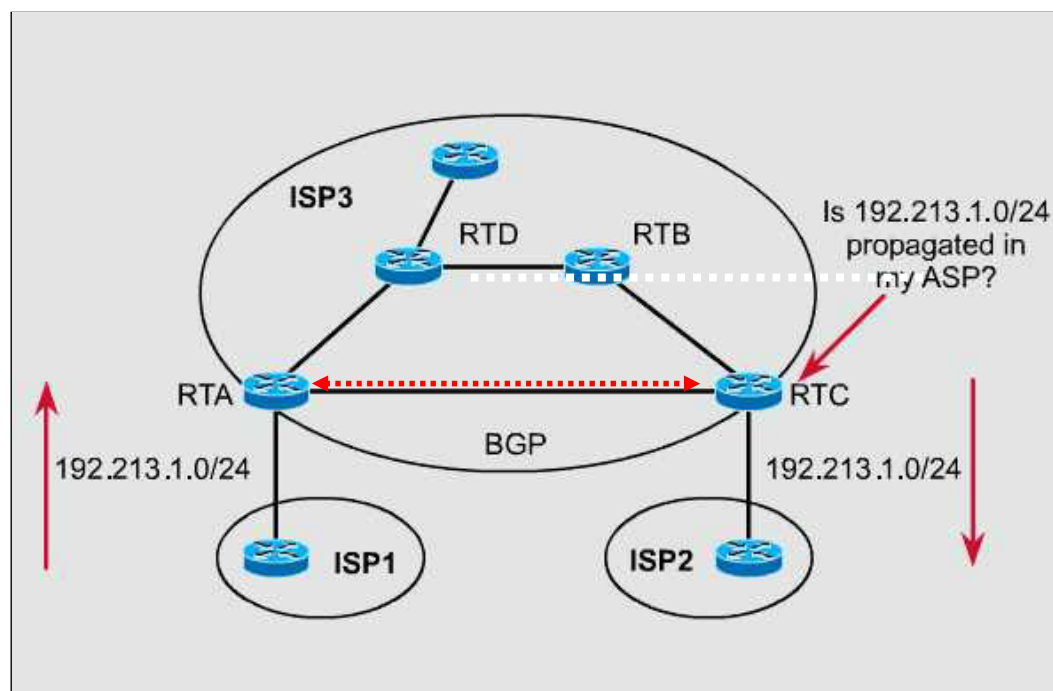
- ▶ Quand un routeur IBGP reçoit un update à partir d'un pair IBGP, il vérifie cette destination avec un **IGP**, comme EIGRP ou OSPF.
- ▶ Si le routeur IBGP ne peut pas trouver le réseau dans la table de routage **IGP**, il n'annonce pas cette destinations aux autres routeurs BGP.

Synchronisation des AS



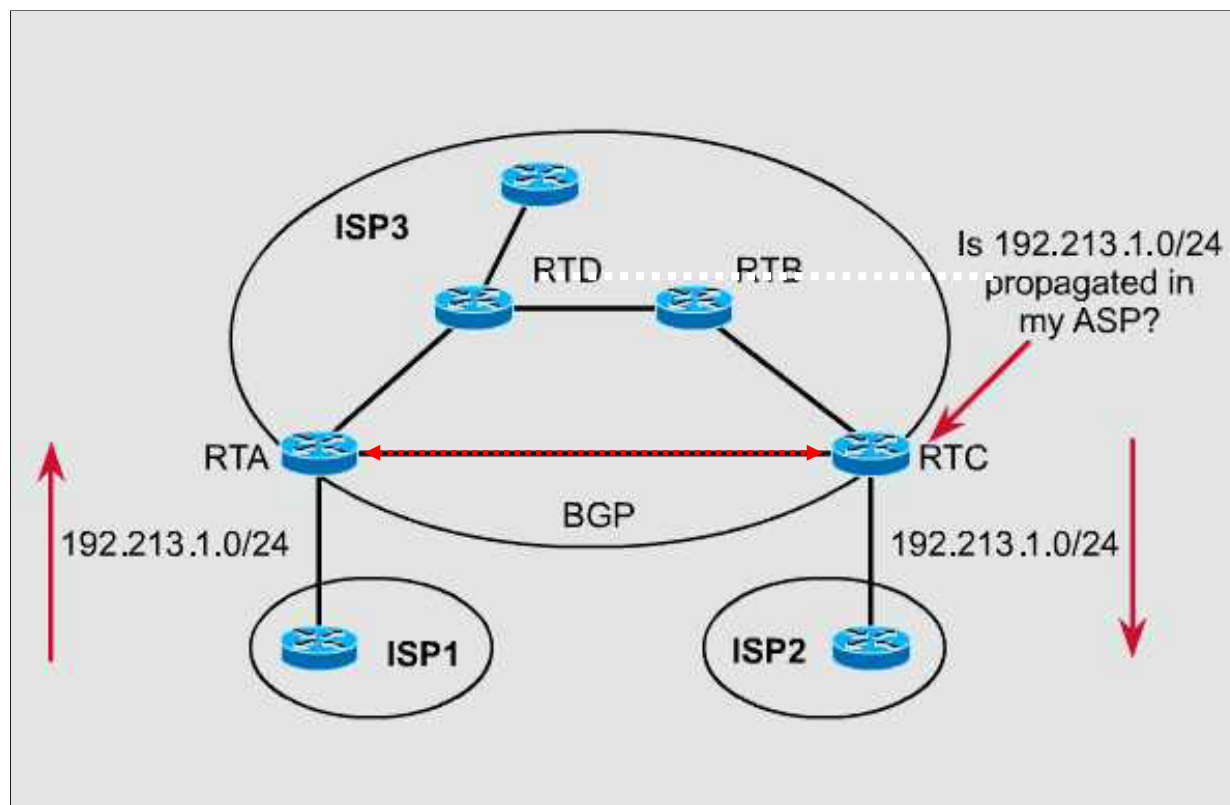
- ▶ Si la route n'est pas atteignable via **IGP**, des routeurs non-BGP ne seront pas capables de directionner le trafic qui traverse le AS vers sa destination
 - ▶ Il est inutile d'annoncer des routes aux pairs externes si les données envoyées à travers cet AS seront rejetés par un routeur non-BGP tôt ou tard

Synchronisation des AS



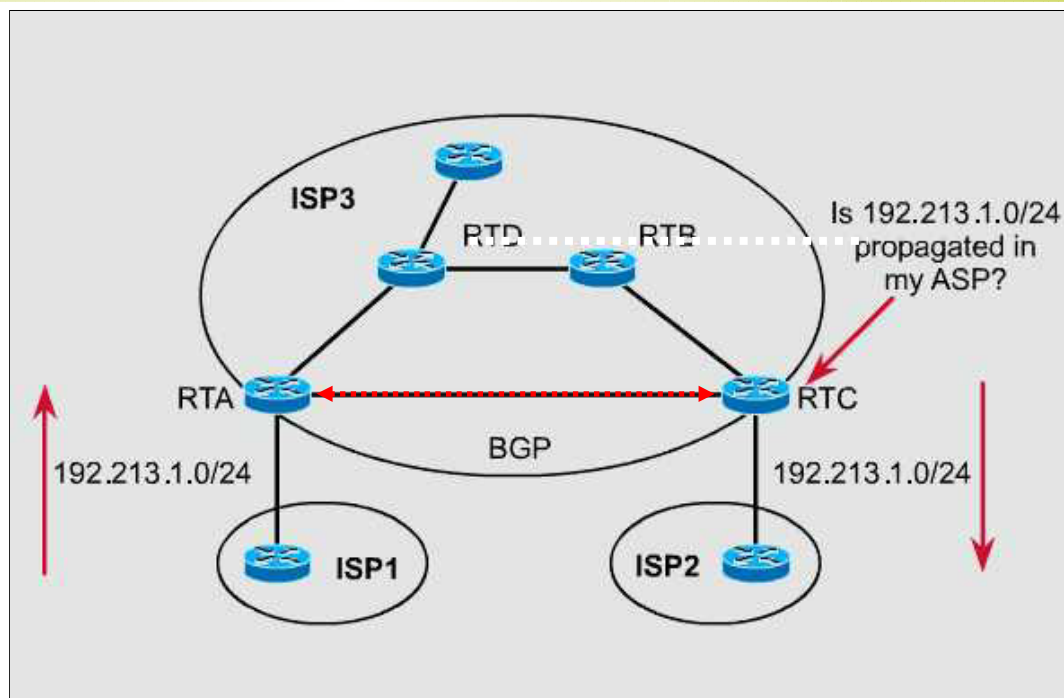
- ▶ La règle de **synchronisation BGP** empêche un routeur BGP (**RTC**) d'annoncer aux pairs extérieurs (**ISP2**) les destinations (**192.213.1.0/24**) apprises par un voisin intérieur (**RTA**) à moins que cette destination soit vérifiable via **IGP (RTD and RTB)**.
- ▶ Si un routeur vérifie la route via **IGP**, il assume que cette route a déjà été annoncé dans l'AS et que le réseau est atteignable

Synchronisation des AS



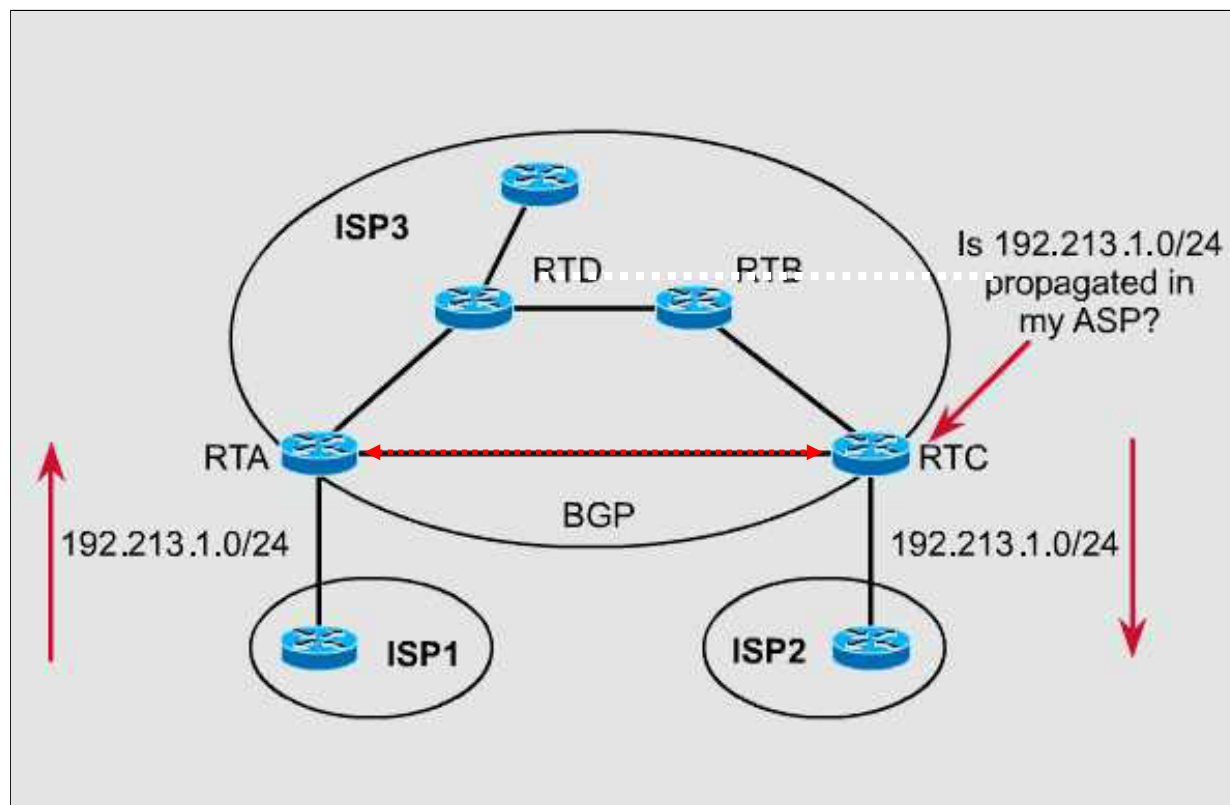
- Si un routeur IBGP (**RTC**) a une route **IGP** vers cette destination, la route est considérée **synchronisée** et la route sera annoncé aux autres pairs BGP (**ISP2**)

Synchronisation des AS



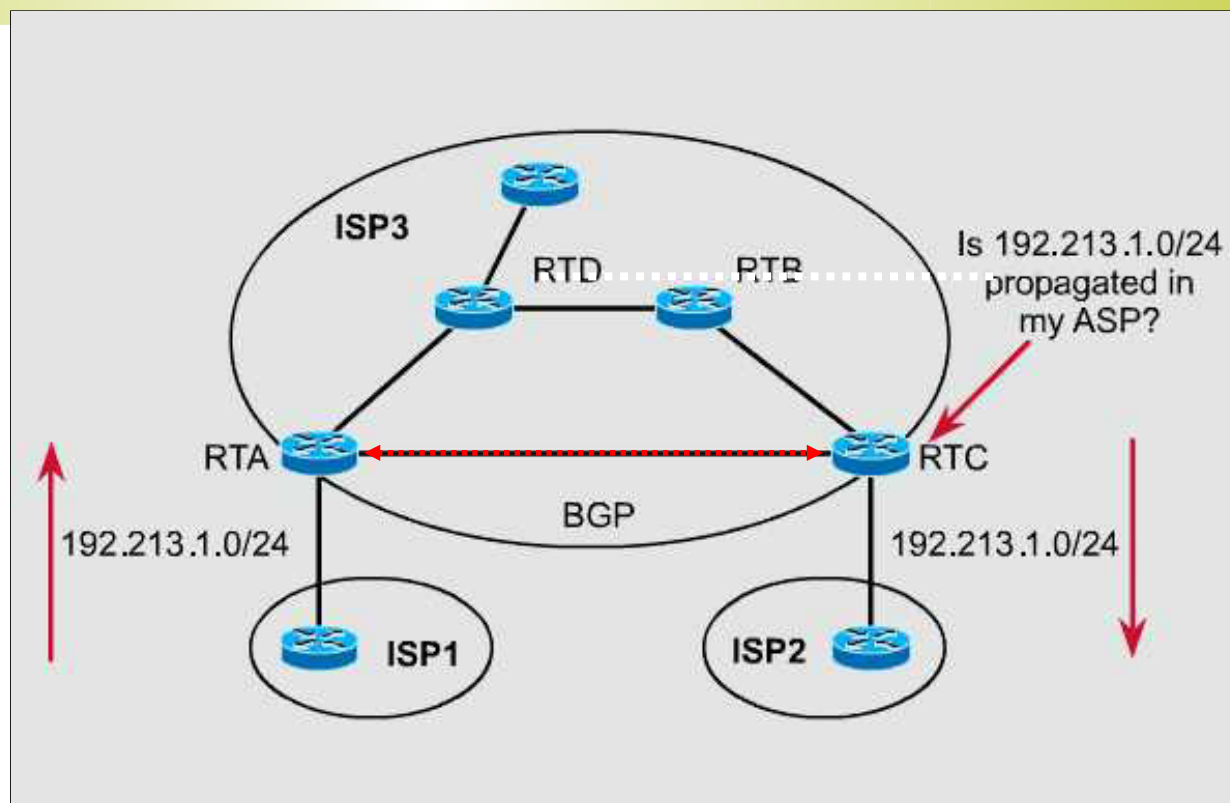
- ▶ L'injection d'une **route BGP** dans un AS est assez coûteuse
- ▶ La redistribution des routes BGP vers un **IGP** représente une surcharge que certains routeurs ne sont pas capables de tolérer
 - ▶ une table BGP représente autour de 200MB
- ▶ De plus, connaître toutes les routes BGP à l'intérieur d'un AS n'est pas nécessaire

Synchronisation des AS



- ▶ L'IOS Cisco a une option **no synchronization**
- ▶ Cette commande donne la possibilité de propager des routes sans les avoir vérifié

Synchronisation des AS



- ▶ Dans la pratique, deux situations permettent d'arrêter la synchronisation sans danger :
 - ▶ Quand tous les routeurs dans l'AS sont des routeurs IBGP (et par conséquent totalement connectés)
 - ▶ La liaison interne est garantie par l'interconnexion en *mesh* IBGP
 - ▶ Lorsque l'AS n'est pas un AS de passage