



R302

Réseaux Opérateurs

Ismail Bennis

Ismail.bennis@uha.fr

MCF, IUT de Colmar, Département Réseaux & Télécoms, bureau N°006
34 rue du Grillenbreit - 68000 Colmar Cedex



R302

- Objectifs de la ressource :
 - Comprendre les technologies des réseaux de collecte et des réseaux de transport déployés par les opérateurs
- Compétences visées :
 - RT1 Administrer les réseaux et l'Internet
 - AC21.05 Identifier les réseaux opérateurs et l'architecture d'Internet
 - RT2 Connecter les entreprises et les usagers
 - AC22.03 Mettre en place une connexion multi-site via un réseau opérateur
- Mot clés : BGP, MPLS, MPLS-VPN
- Volume horaire : 18 h : 2 CM + 3 TD + 4 TP
- Evaluation :
 - 1 DS
 - 1 TP Test



01

Plan de la ressource

- Introduction et généralité
- BGP (Border Gateway Protocol)

02

- MPLS
- MPLS-VPN L3

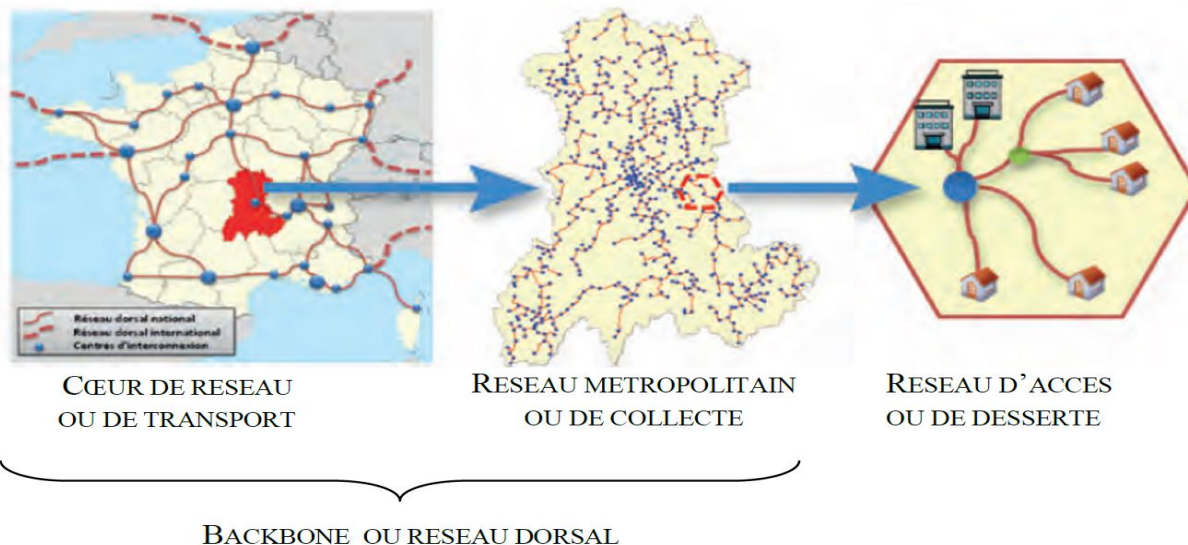
www.uha.fr



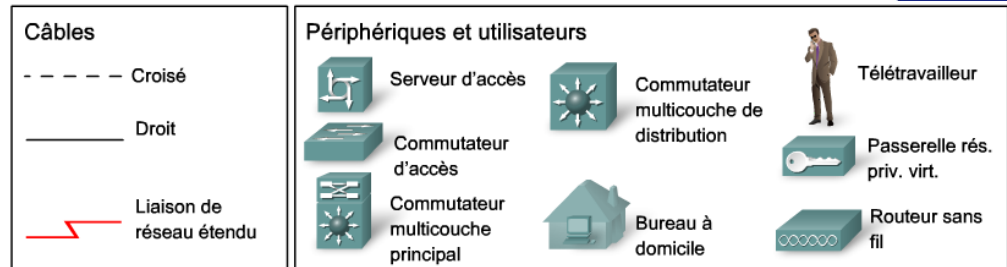
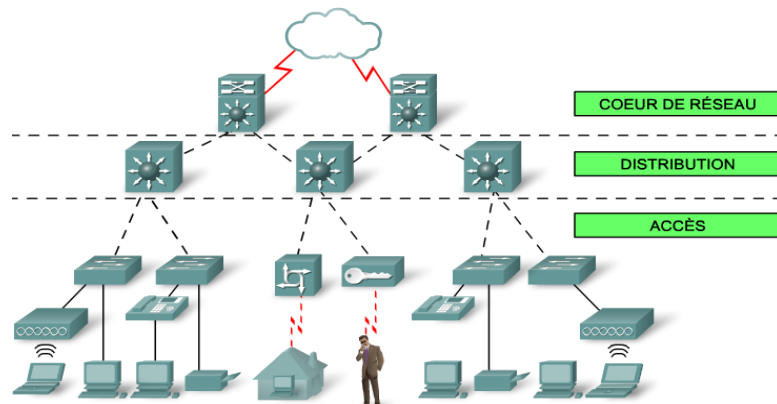
Introduction et généralité

Introduction et généralité

- C'est quoi Internet ?
- Un réseau IP constitué par des interconnexions entre les différents **opérateurs** suivant leurs tailles et leurs capacités à gérer ces interconnexions
- Un Fournisseur d'accès à internet (**FAI / ISP**) est un organisme qui offre une connexion Internet à ses client et le réseau mondial
- Les réseaux d'opérateurs sont, généralement, composés de trois niveaux :
 - Réseaux **d'accès**
 - Réseaux de **distribution**
 - **Backbone** ou réseau dorsal



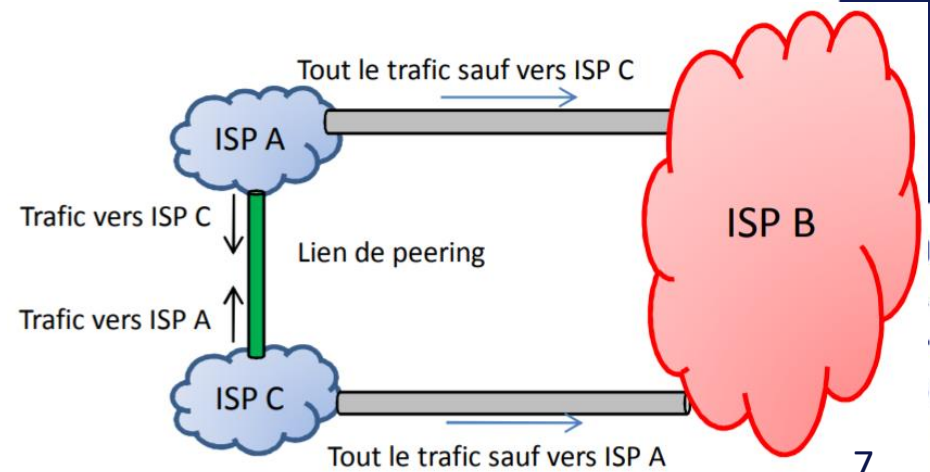
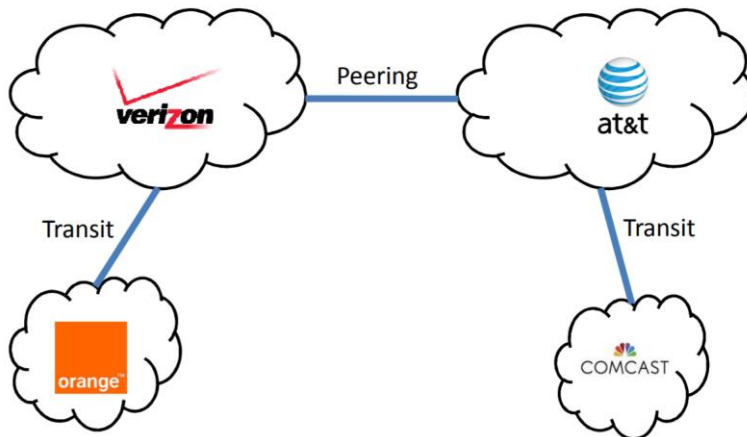
Introduction et généralité



- **Couche d'accès** : fourni un moyen de connecter des périphériques au réseau, ainsi que de contrôler les périphériques qui sont autorisés à communiquer sur le réseau tels que les modems xDSL, FTTH, Téléphones, etc
- **Couche de distribution** : gère le flux du trafic réseau à l'aide de stratégies des politiques de collectes de données et des fonctions de routage.
- **Couche cœur** : constitue le réseau fédérateur à haut débit de l'inter-réseau afin d'effectuer la transmission rapide d'importantes quantités de données. Doit assurer une disponibilité et une redondance élevée.

Introduction et généralité

- Pour mieux structurer l'Internet, les FAI utilisent des **systèmes autonomes (AS)** et du routage adapté pour échanger des routes entre les AS
- Deux types d'interconnexions entre les opérateurs :
 - **Transit** : permet d'échanger toutes les routes vers les autres opérateurs. Ainsi, chaque opérateur possède une vue globale du réseau Internet.
 - **Peering**: échanger que quelques routes, typiquement celles menant à son propre réseau et ses propres clients
- Peering est avantageux quand il y a une forte interaction entre deux opérateurs :
 - Gain en performance ; Plus court chemin ; Délais faibles

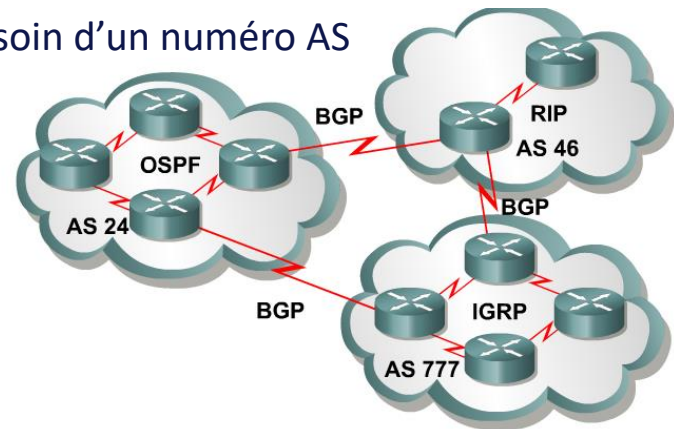


Introduction et généralité

- Donc Internet n'est rien de plus qu'un tas de systèmes autonomes (AS) connectés les uns aux autres.
- Dans un AS, nous utilisons un **IGP** et entre AS on utilise **EGP**
- **IGP** (Interior Gateway Protocol)
 - Protocoles de routage utilisés à l'intérieur d'un système autonome (AS)
 - RIP, IGRP, EIGRP, OSPF
- **EGP** (Exterior Gateway Protocol)
 - Protocole de routage utilisé pour échanger des informations de routage entre différents systèmes autonomes (AS)
 - **BGP**

Introduction et généralité : Systèmes Autonomes (AS)

- **AS** : 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 protocole spécifique (**BGP**) est utilisé pour interconnecter les AS et acheminer les information entre eux.
- Un AS peut être composé :
 - d'un ensemble de routeurs qui tourne avec le même protocole IGP
 - d'un ensemble de routeurs qui tournent avec différents protocoles de routage (une grande entreprise ou un FAI)
- Un AS est identifié par un numéro
 - Attribué par l'**IANA** (Internet Assigned Numbers Authority)
 - Entre 1 et **65535**.
 - Les adresse de **64512 à 65535** sont réservées pour un usage privée
- À cause du nombre limité d'AS, il faut justifier le besoin d'un numéro AS



Introduction et généralité : Single/Dual Homed

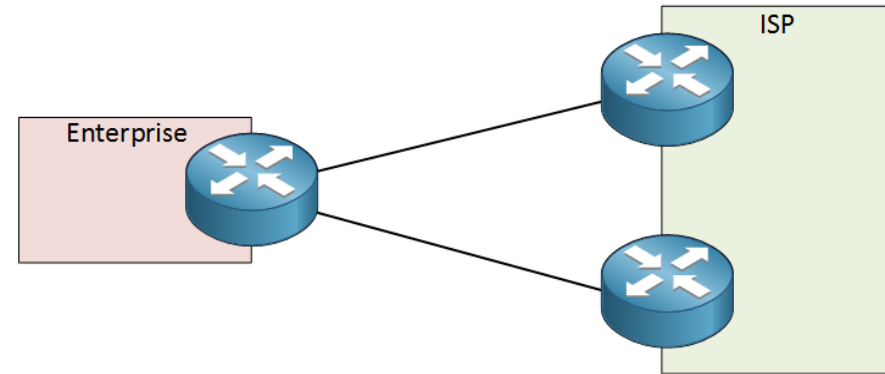
- Plusieurs cas de figures existent lorsqu'on parle des connexions entre les FAI et les clients :

Single Homed



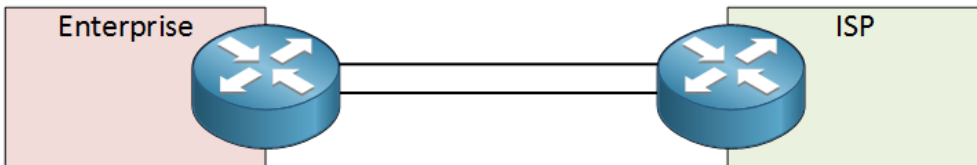
Faible coût mais pas de redondance

Dual Homed



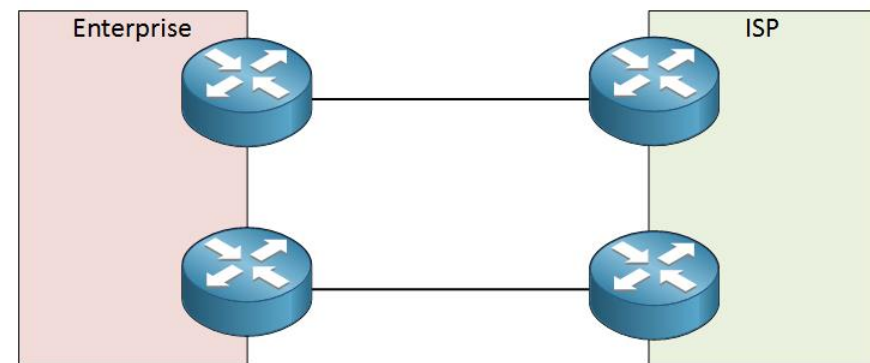
Redondance des liens et routeur côté FAI

Dual Homed



Redondance des liens

Dual homed

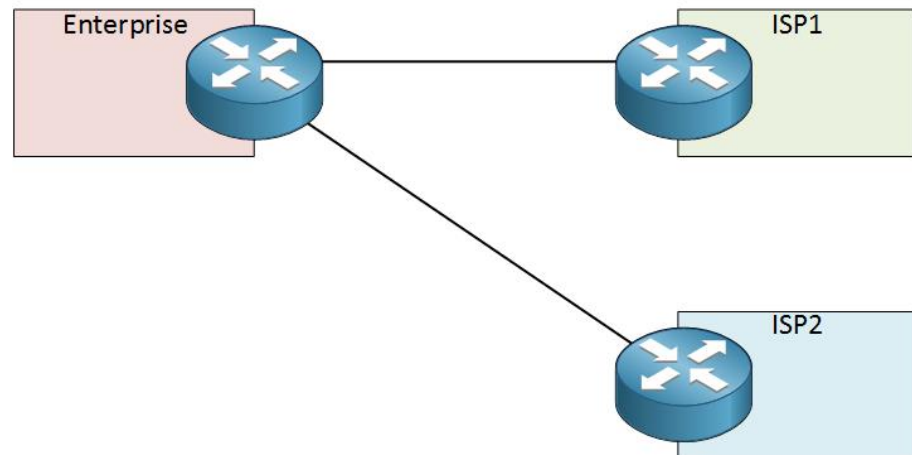


Redondance des liens et routeur côté FAI et client

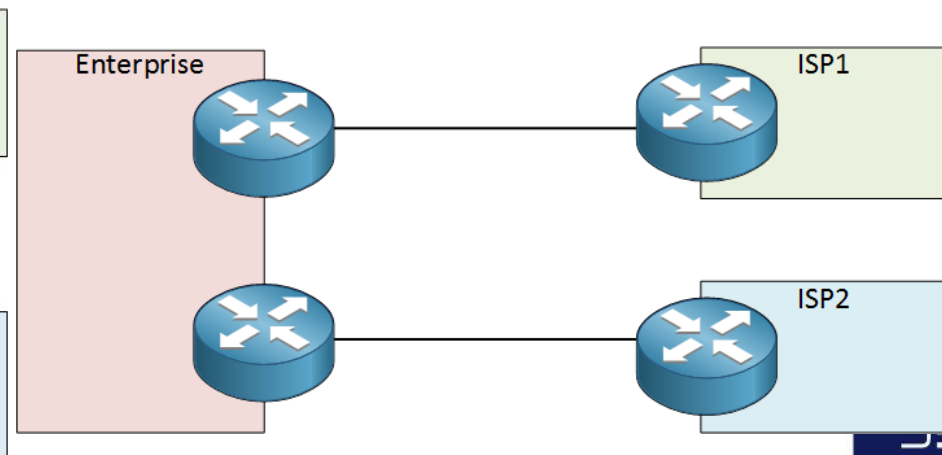
Introduction et généralité : Single/Dual Multi-homed

Single Multihomed

Single Multihomed



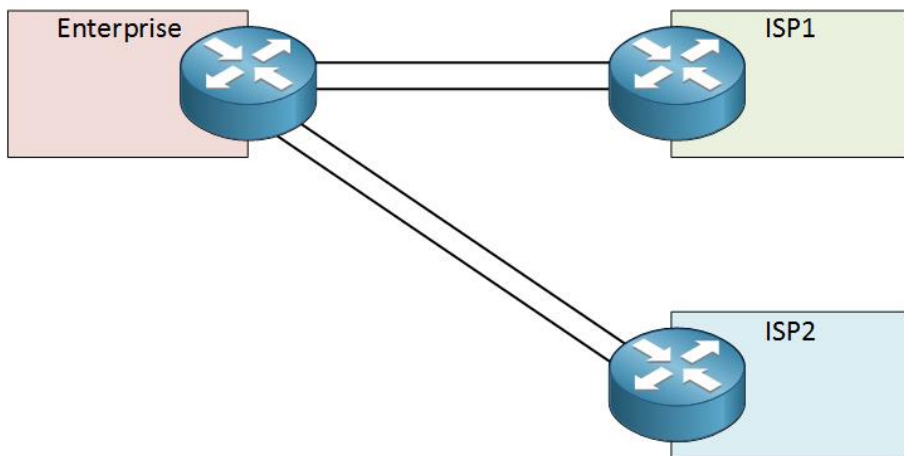
Redondance avec deux FAI mais le routeur reste un point critique !



Redondance des routeurs côté client avec deux FAI différents. Mais un seul lien avec les ISP !

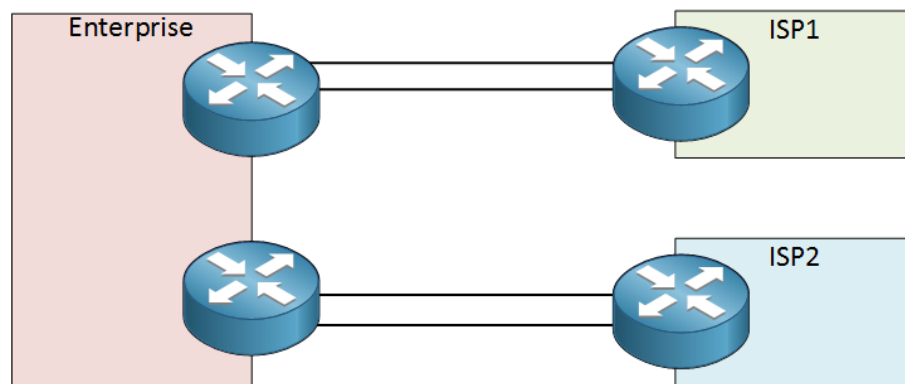
Introduction et généralité : Single/Dual Multi-homed

Dual Multihomed



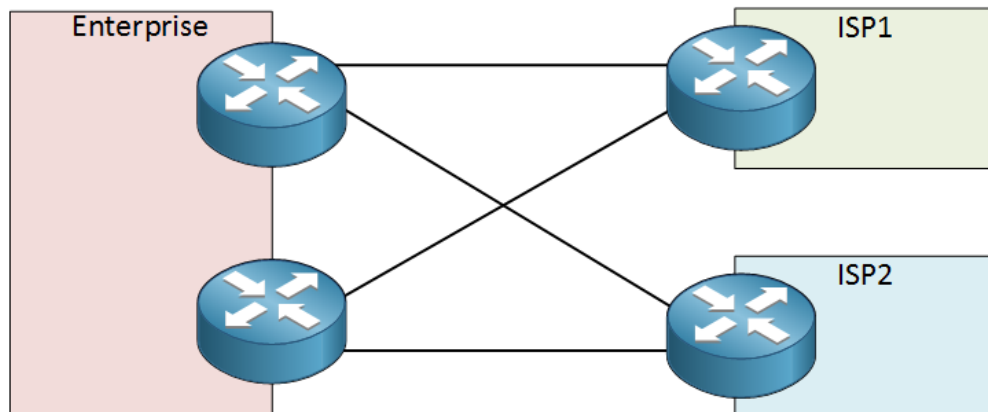
Redondance des liens avec deux FAI mais le routeur reste un point critique

Dual Multihomed



Redondance des liens et routeur avec deux FAI mais perte de la connectivé avec un FAI si le routeur correspondant et HS

Dual Multihomed



Redondance des FAI, liens et routeurs

Introduction et généralité : Politique de routage

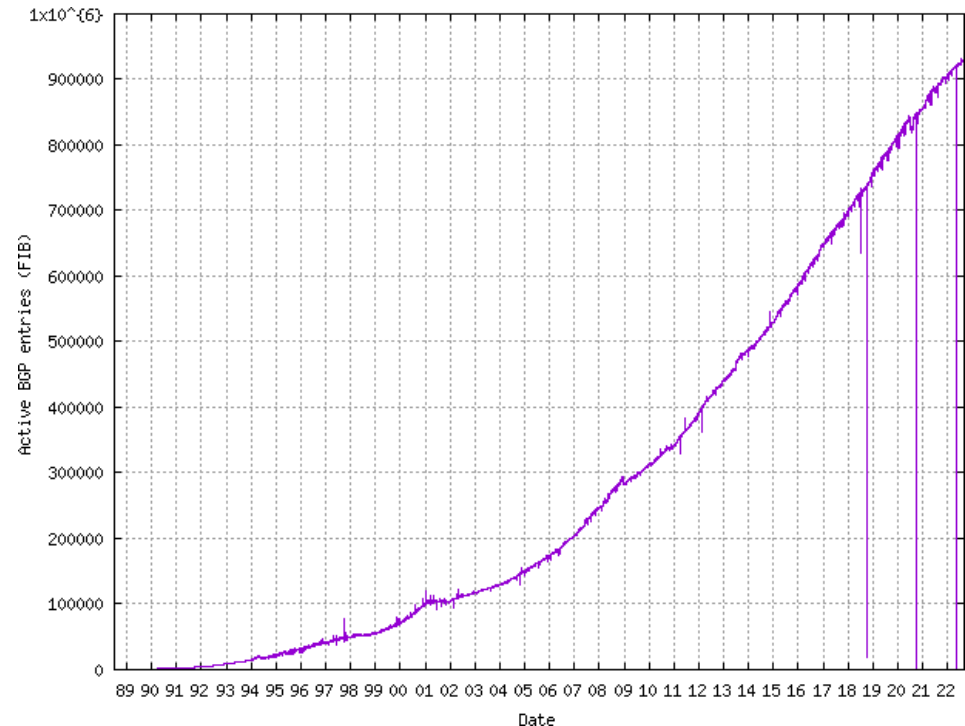
- RIP, OSPF et EIGRP sont tous différents, mais ils ont un point en commun : trouver le chemin le plus court vers la destination.
- À l'égard d'Internet, nous ne nous soucions pas que de trouver le chemin le plus court, mais plutôt d'être capable de manipuler les chemins de trafic !
- Un seul protocole de routage qui est utilisé pour cela: **BGP** (Border Gateway Protocol)
- BGP a deux variantes:
 - **External BGP (eBGP)** : utilisé entre des AS différents
 - Distance administrative : **20**
 - **Internal BGP (iBGP)** : utilisé dans le même AS
 - Distance administrative : **200**

Est-il obligatoire d'utiliser BGP ?

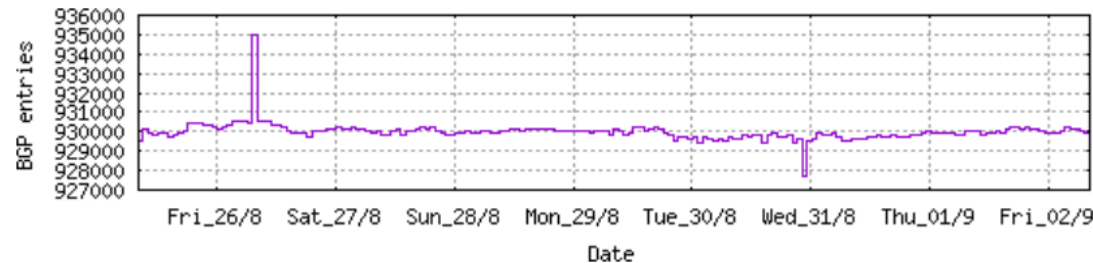
- Pourquoi un IGP n'est pas suffisant !?

Les entrées BGP actives (FIB)
30-juin-1988 au 02-sept.-2022

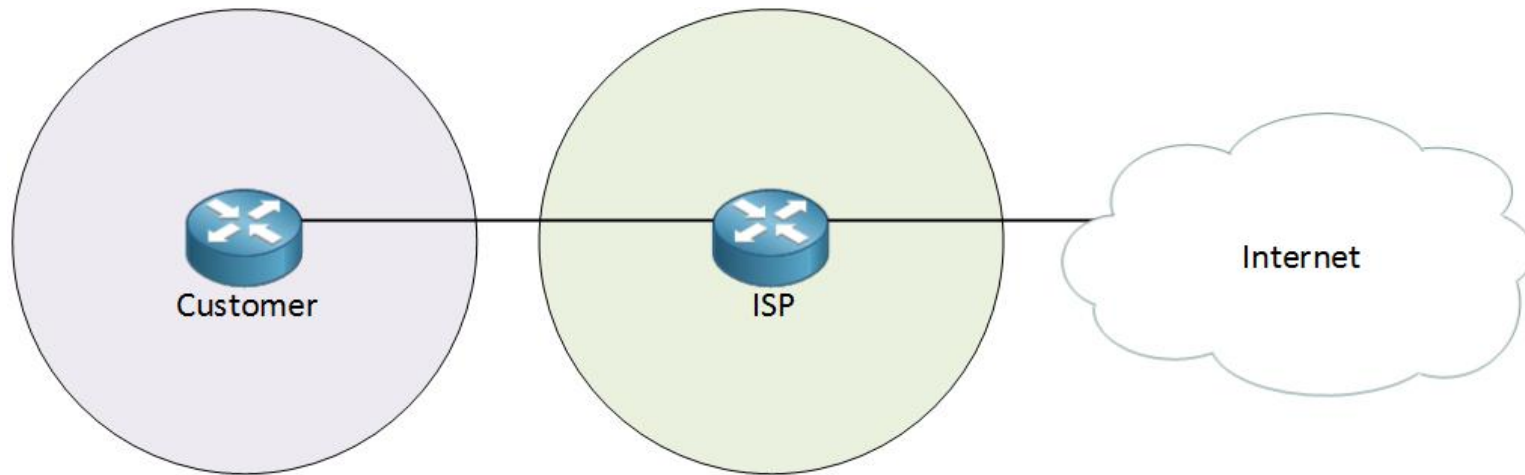
<https://www.cidr-report.org/as2.0/>



Date	Prefixes	CIDR Aggregated
26-08-22	930274	512807
27-08-22	930155	513032
28-08-22	929863	513252
29-08-22	930027	512700
30-08-22	929679	513573
31-08-22	929566	513840
01-09-22	930057	514292
02-09-22	929915	514510



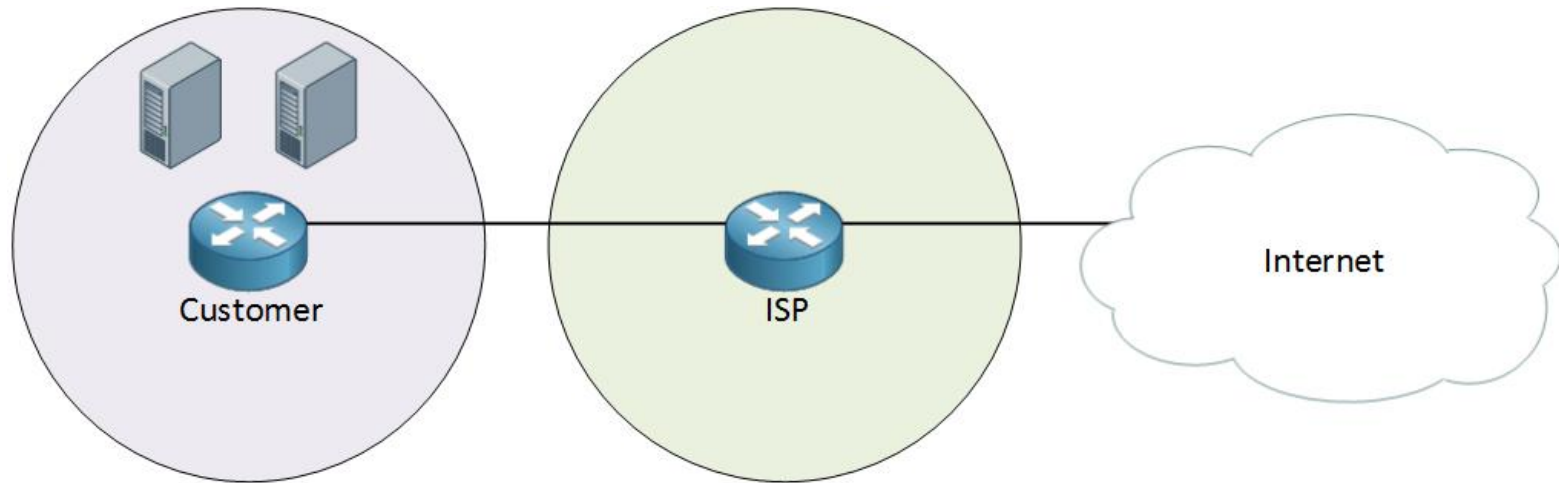
Quand avoir besoin du BGP ?



- Ici le FAI assure au client l'accès à l'Internet en lui donnant une seule adresse publique
- Les machines coté client vont utiliser du NAT
- Coté client une route par défaut est suffisante pour accéder à Internet

➔ Pas besoin de BGP !

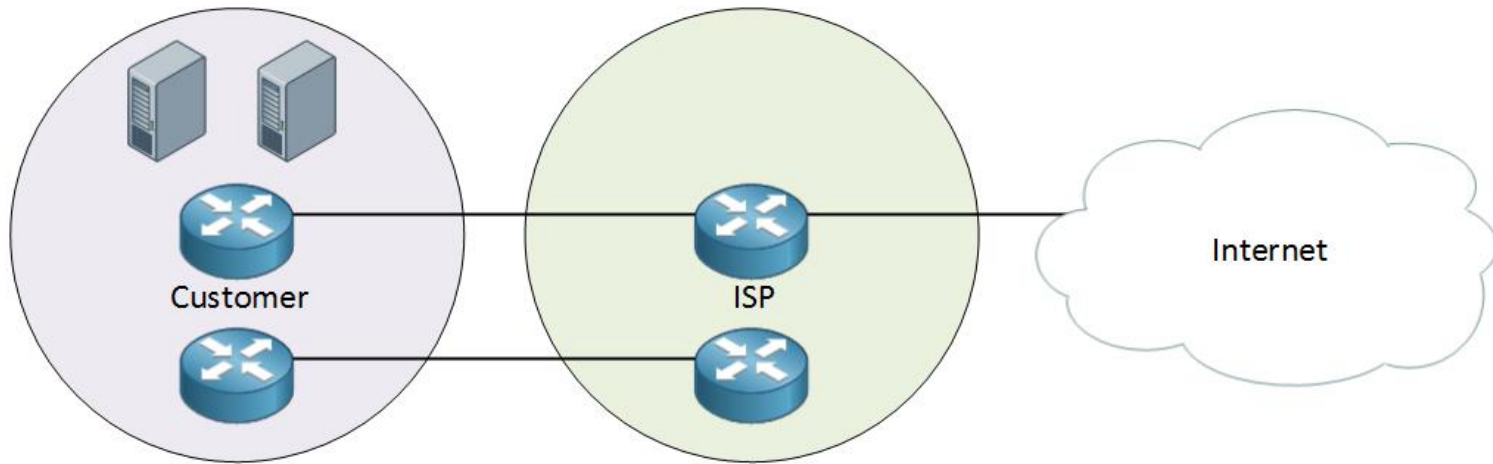
Quand avoir besoin du BGP ?



- Et si le client possède plusieurs serveurs à être joignable de l'extérieur ?
- Port forwarding ou bien avoir une autre adresse IP publique

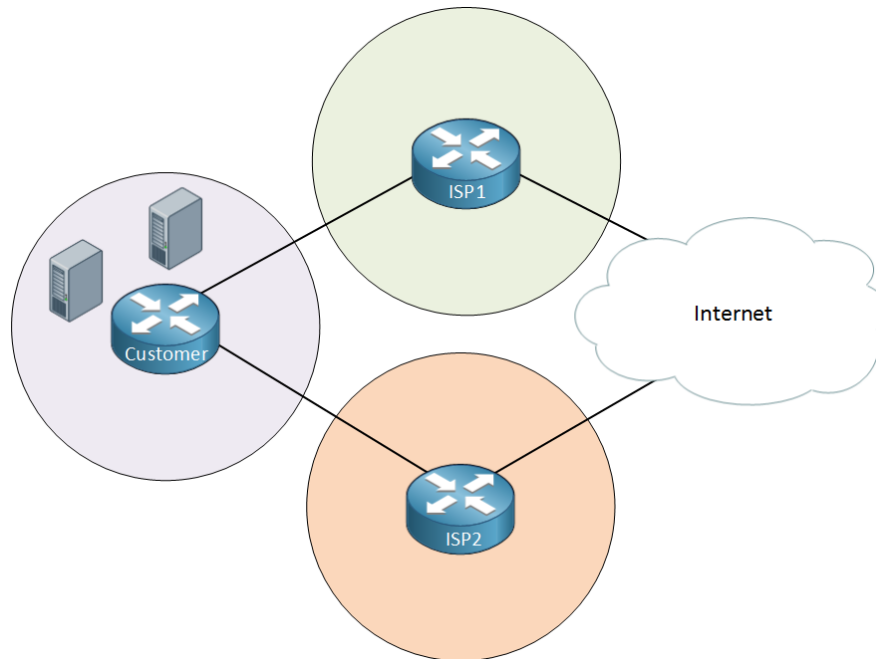
➔ Pas besoin de BGP !

Quand avoir besoin du BGP ?



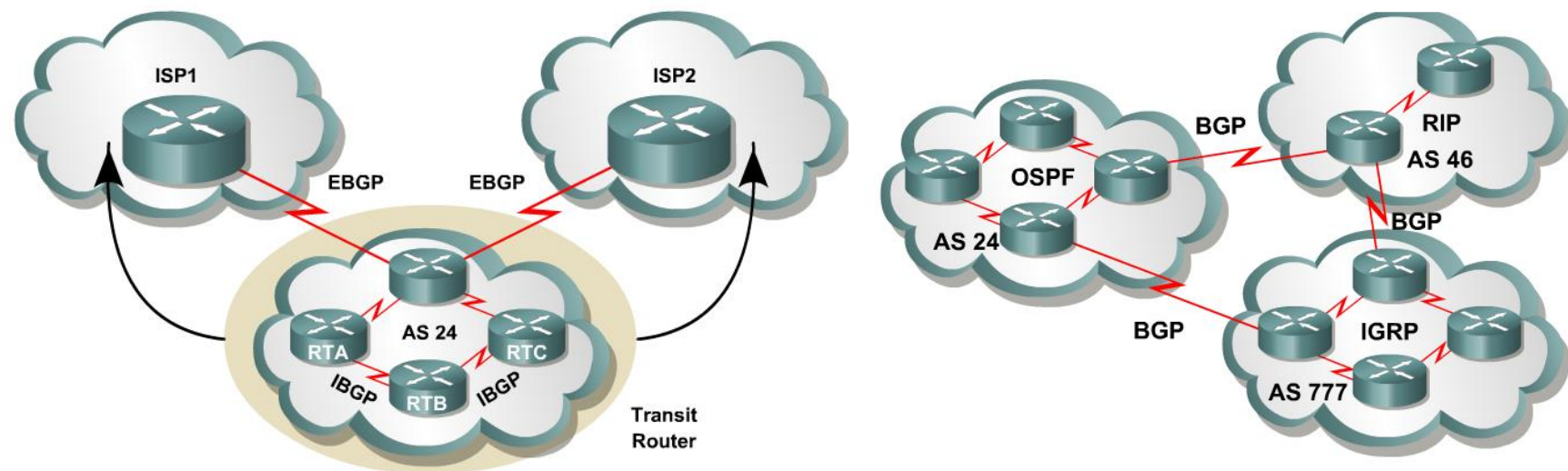
- Et si on veut assurer de la redondance ?
 - Un lien primaire et autre secondaire
 - On peut faire même de l'équilibrage de charge ! (50% / 50%)
 - Toujours pas besoin de BGP !
- ➔ mais si on veut envoyer 80% du trafic sortant sur le lien principal et 20% sur le lien de sauvegarde ? Cela est compliqué ... mais avec BGP, c'est possible

Quand avoir besoin du BGP ?



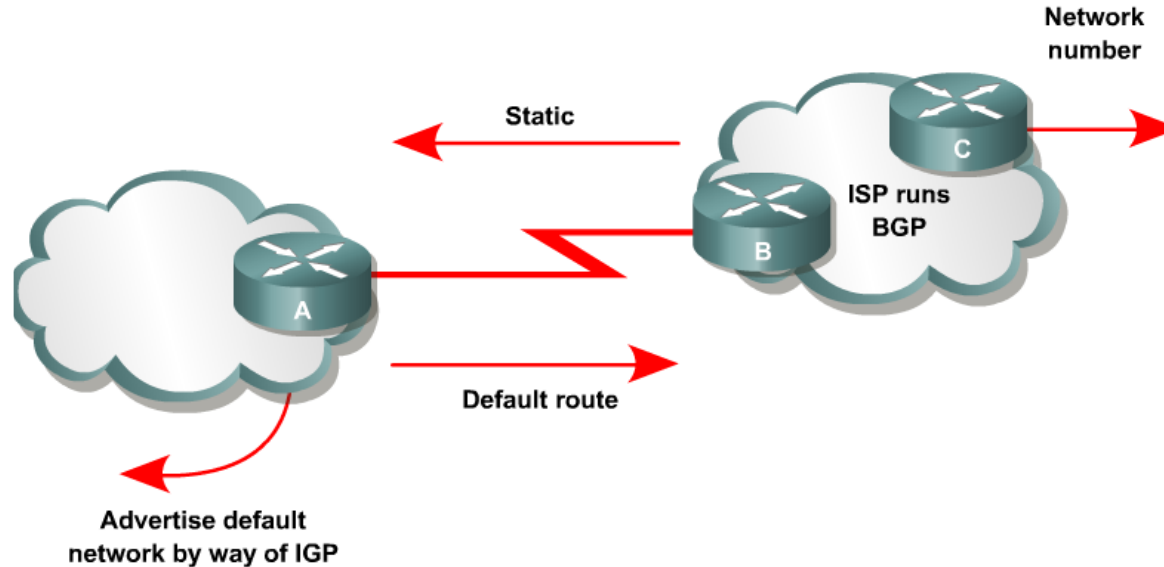
- Pour une meilleure redondance on rajoute une connexion vers un autre FAI
 - Quelle adresse publique il faut utiliser ? Celle du FAI1 ou FAI2 ?
 - Solution : on demande une adresse IP public à l'IANA, qu'on annonce à nos FAI.
- ➔ Cela se fait avec un protocole de routage qui n'est autre que **BGP**.

Quand avoir besoin du BGP ?



- **BGP est utilisé dans les situations suivantes :**
 - L'AS permet le transit de paquets d'autres AS
 - L'AS a plusieurs connexions à d'autres AS (multi-homed)
 - Les flux entrants ou sortants de l'AS doivent être manipulés

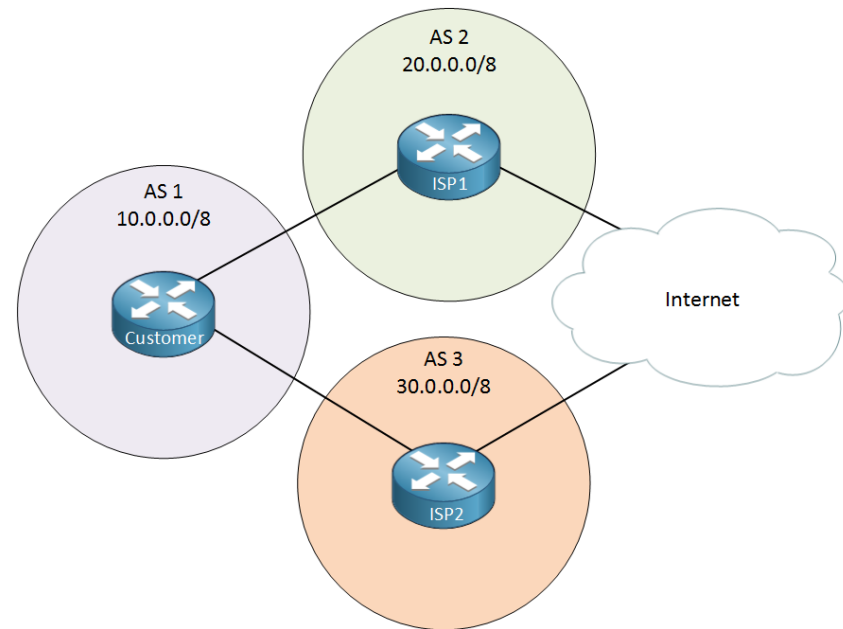
Quand avoir besoin du BGP ?



- Ne pas utiliser BGP si :
 - il a une seule connexion vers un autre AS ou FAI
 - votre AS ne requiert pas des filtrages ou manipulations des flux
 - les routeurs sont limités en mémoire ou capacité de calcul
 - la liaison entre les AS a une faible bande passante

BGP – échange de routes

- Lorsque nous exécutons BGP, cela signifie-t-il que nous devons apprendre plus de 900k préfixes ?

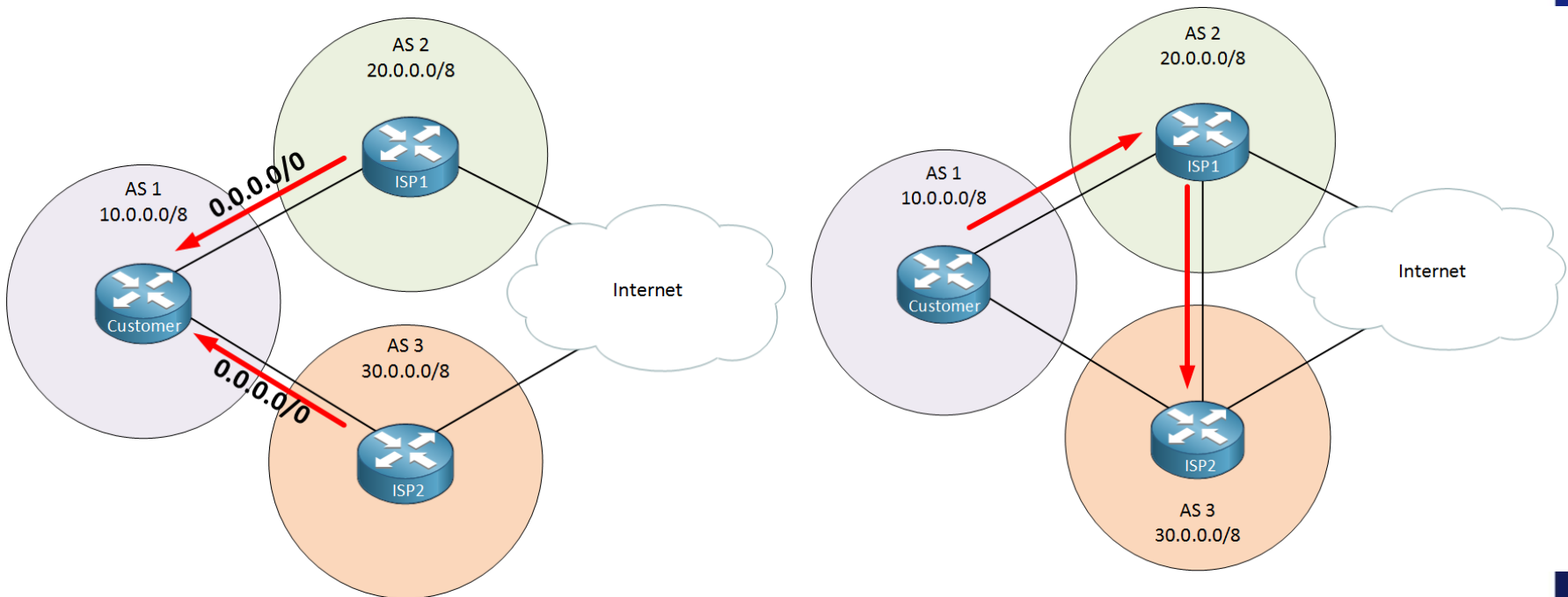


Les FAI peuvent annoncer :

- uniquement la route par défaut.
- une route par défaut et une table de routage partielle.
- la table de routage Internet complète.

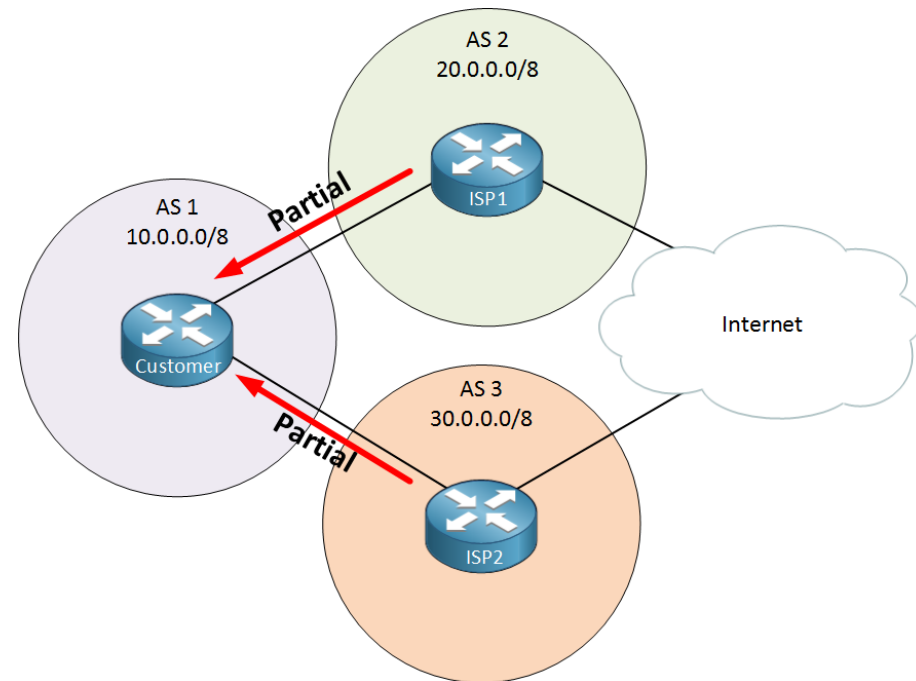
BGP – échange de routes

- Cas route par défaut :

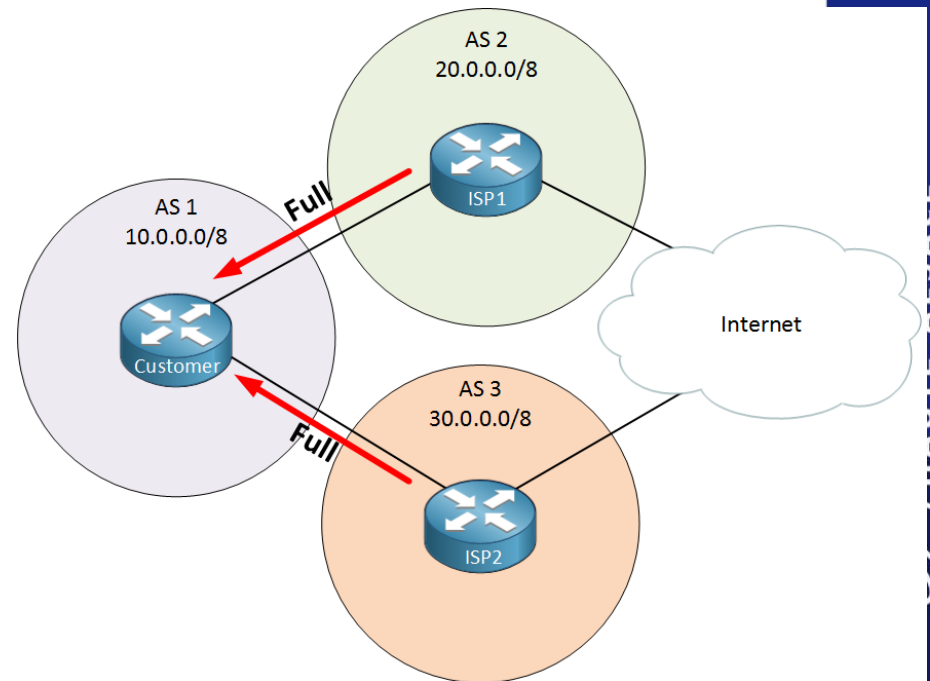


- Besoin peu de ressource mais peut conduire à un routage non optimal !
- Le client n'as pas une visibilité de ce qui se cache derrière AS2 et AS3

BGP – échange de routes



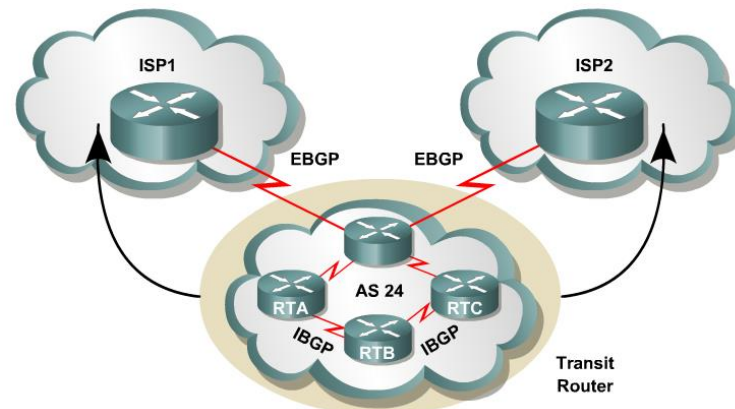
Cas partage partiel : une meilleur visibilité avec des meilleures décisions de routage



Cas partage complet : plus de ressources, mais les meilleures décisions de routage seront prises

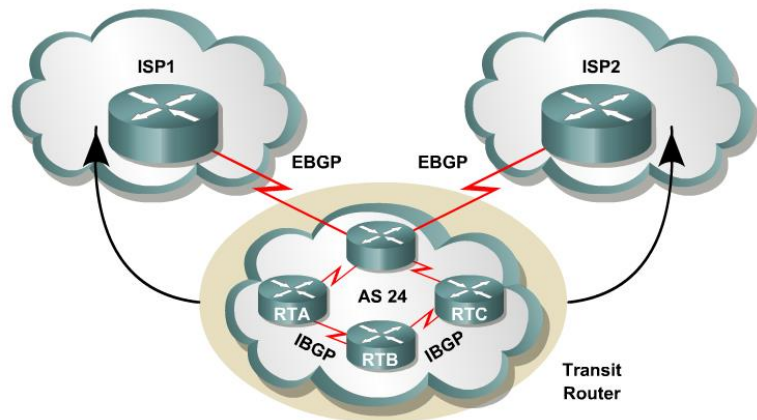
BGP – concepts de base

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



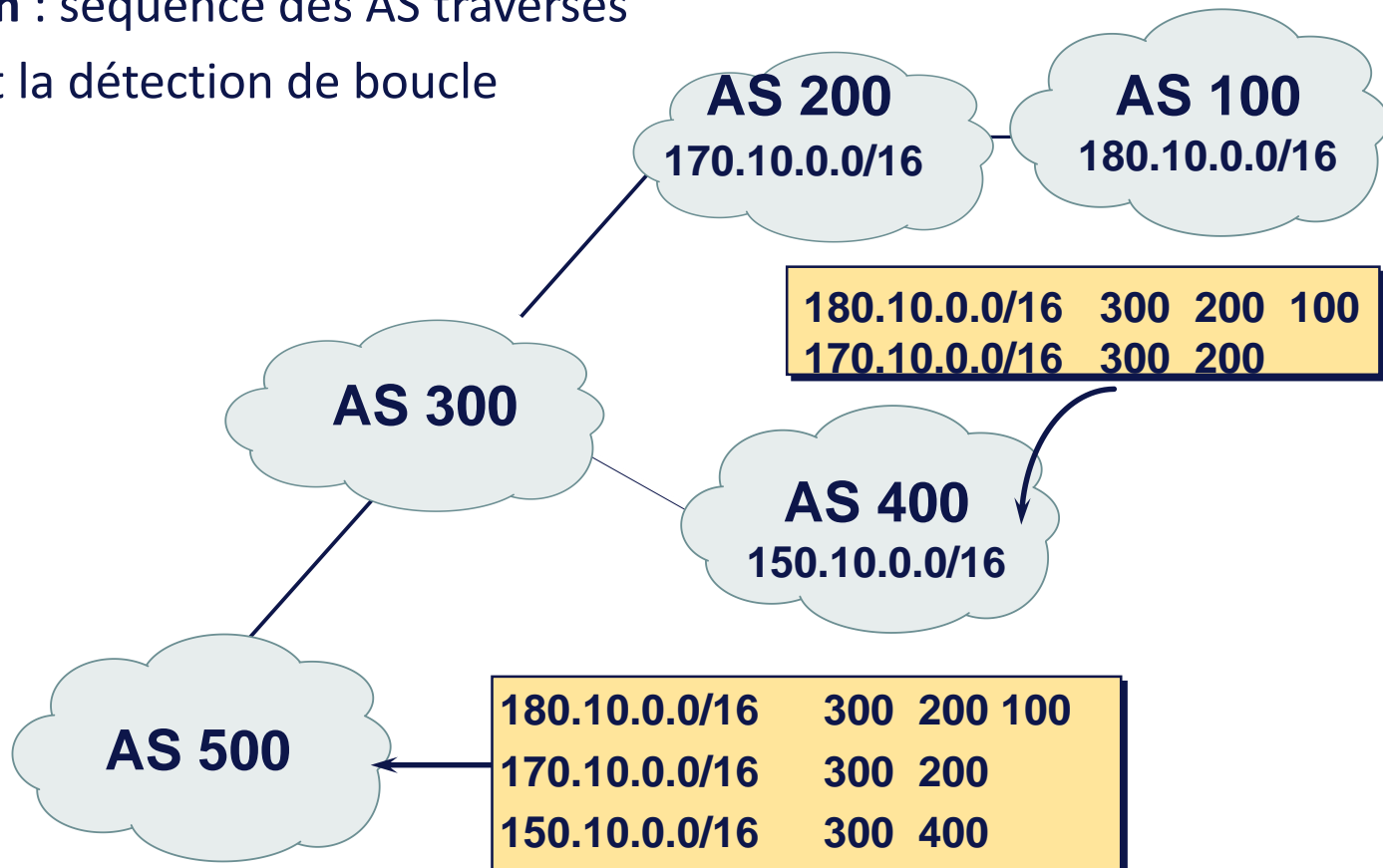
BGP – concepts de base

- BGP
 - fait les décisions selon des politiques de routage
 - n'indique pas les détails internes des ASs
 - ne représente qu'un arbre des ASs
- Les **misés à jour BGP** utilisent TCP sur le port **179**
- 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



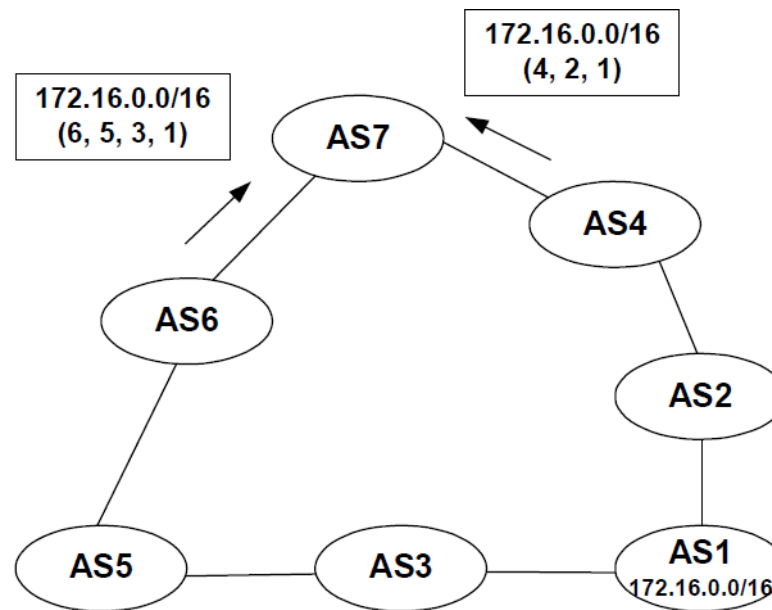
BGP - Path Vector

- BGP est un protocole de routage à **vecteur de chemin**
- Un protocole à vecteur de chemin définit une route comme un appariement entre une destination et les attributs du chemin vers cette destination
- **AS-Path** : séquence des AS traversés
- Permet la détection de boucle



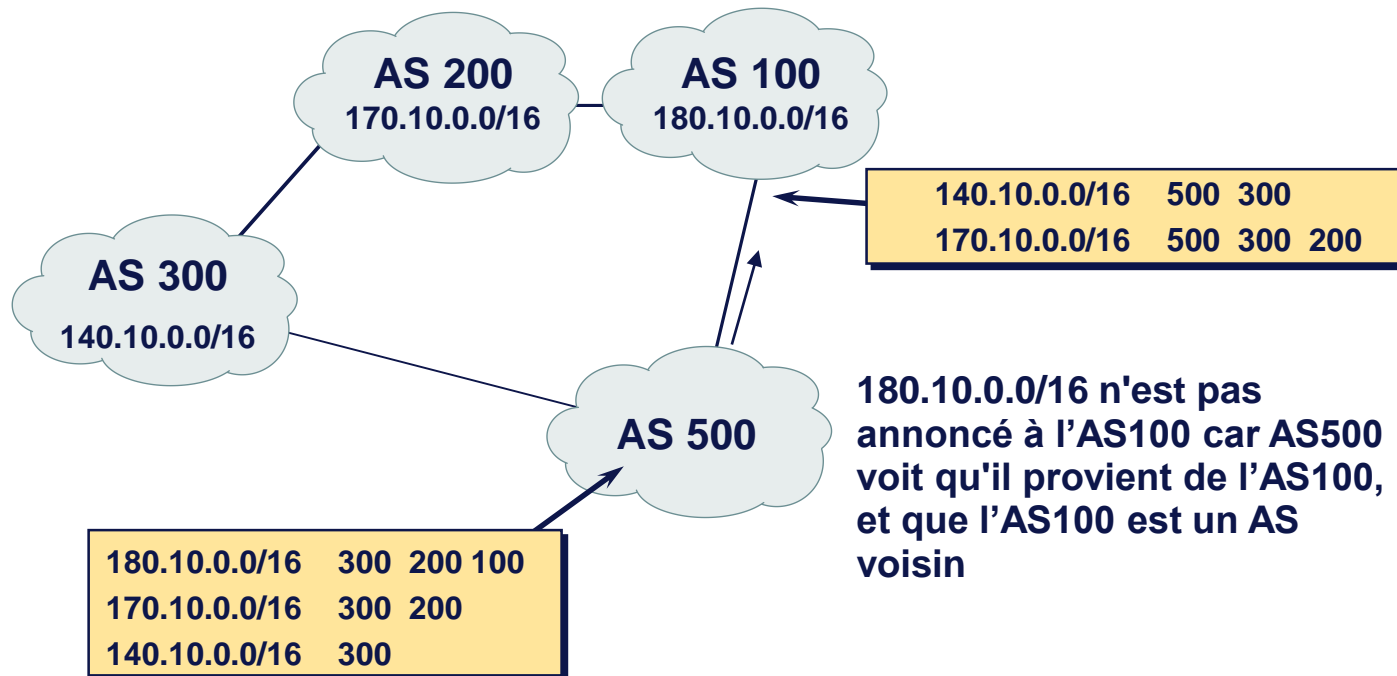
BGP - Path Vector

- **AS_PATH** est l'un des attributs associés aux routes
- La stratégie la plus simple consiste à choisir le chemin le plus court
- Pour deux annonces de réseaux identiques (mêmes préfixes), BGP préfère les routes avec le **AS_PATH** le plus petit
- Dans cet exemple, AS7 choisit le chemin le plus court (4, 2, 1)



BGP - Path Vector

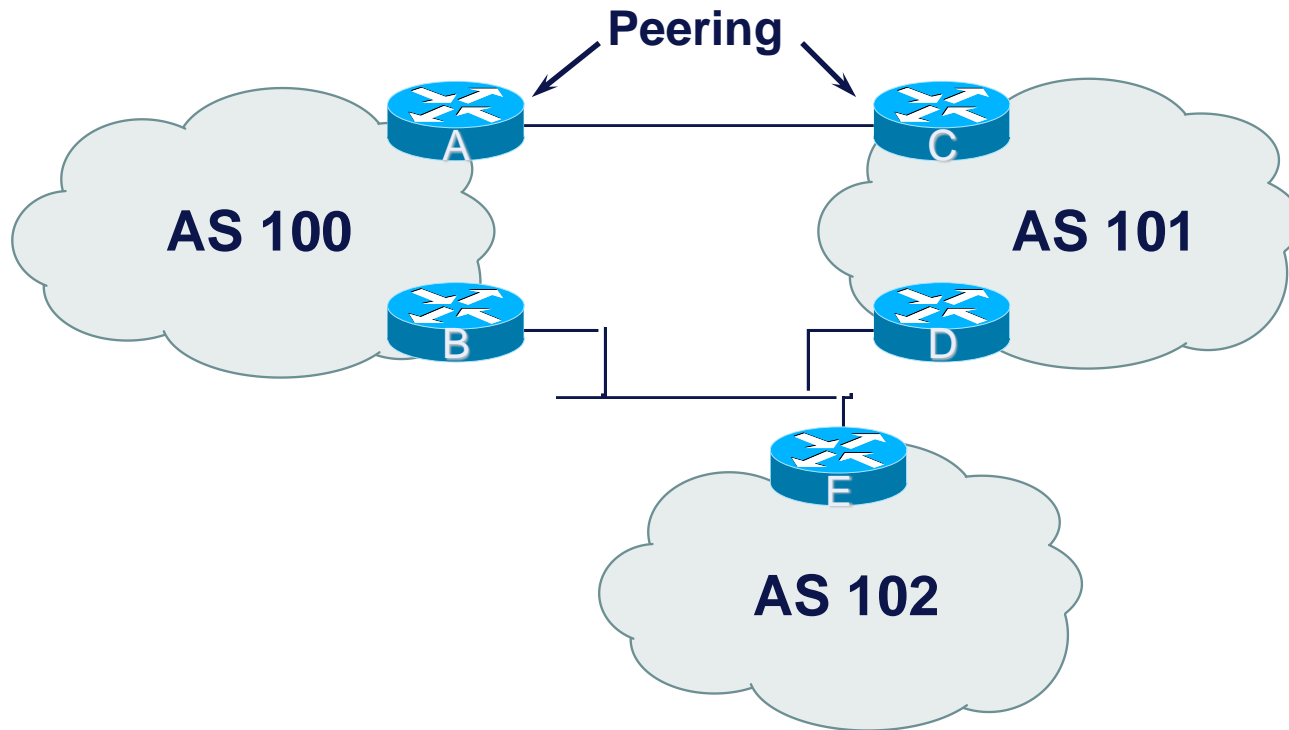
- Détection de boucle : grâce aux numéros d'AS, un système autonome peut détecter un chemin qui passe deux fois par le même AS



- Un préfixe ne sera pas accepté par un routeur BGP s'il y voit son propre numéro AS

BGP – peering

- Voisins ou pairs (**peers**) : deux routeurs qui établissent une session BGP (avec une connexion TCP)

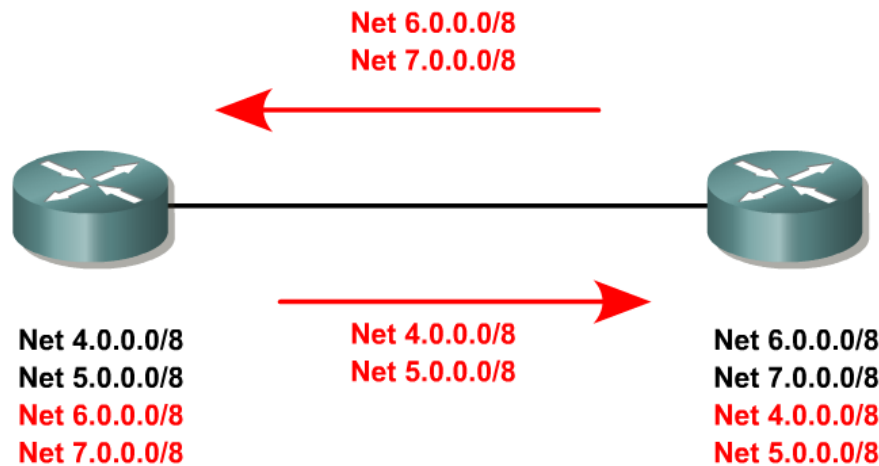


BGP – peering

- Chaque routeur doit exécuter le protocole BGP pour établir le **peering**
- Un routeur BGP :
 - Apprend plusieurs chemins via ces voisins BGP **internes** et **externes**
 - Choisit le meilleur chemin et l'installe dans la table de transfert (forwarding table)
- Politiques appliquées en influençant le choix du meilleur chemin
- Une fois la connexion TCP ouverte, BGP échange plusieurs messages avec les paramètres de la connexion et les informations de routage.
- Tous les messages BGP sont envoyés en **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

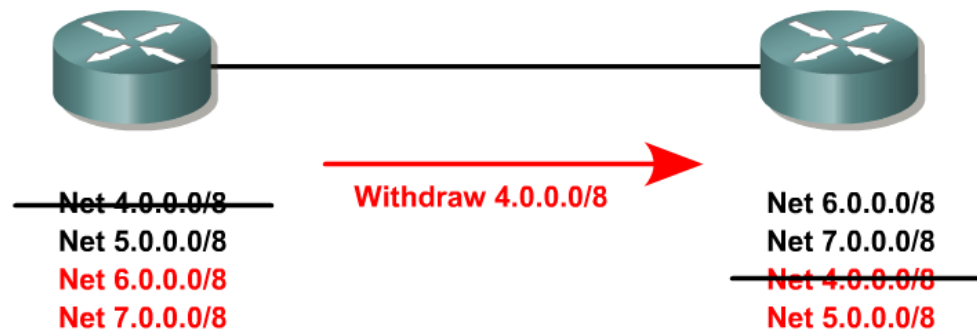
BGP – types de message

- 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



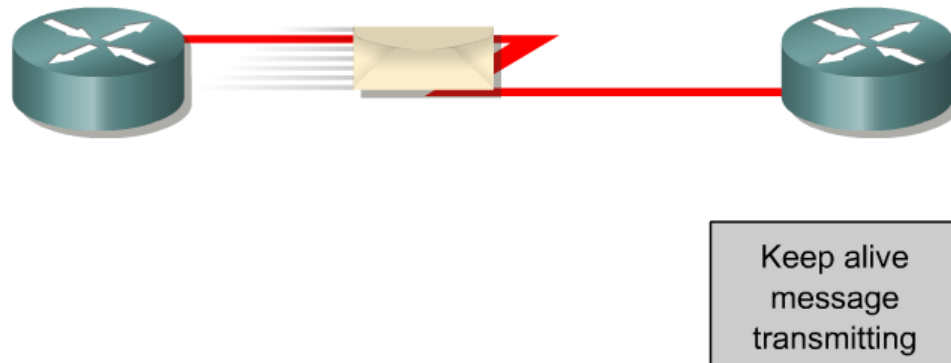
BGP – types de message

- 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 MAJ reçues de chaque routeur
 - Incrémenté à chaque modification des tables de routage BGP
 - Une incrémentation rapide indique normalement qu'un réseau instable ou mal configuré



BGP – types de message

- BGP Keepalives
 - Messages échangés entre les pairs pour garder active la connexion
 - Par défaut c'est **60 secondes** (Cisco) (RFC 1771)
 - Si l'intervalle est dépassé trois fois (**180 seconds**), un pair est considéré **down**
 - Les intervalles peuvent être configurés



BGP - Attributs d'un Chemin

- Chaque chemin 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
- Certains attributs ne sont pas reconnus par tous les fabricants



BGP - Attributs d'un Chemin

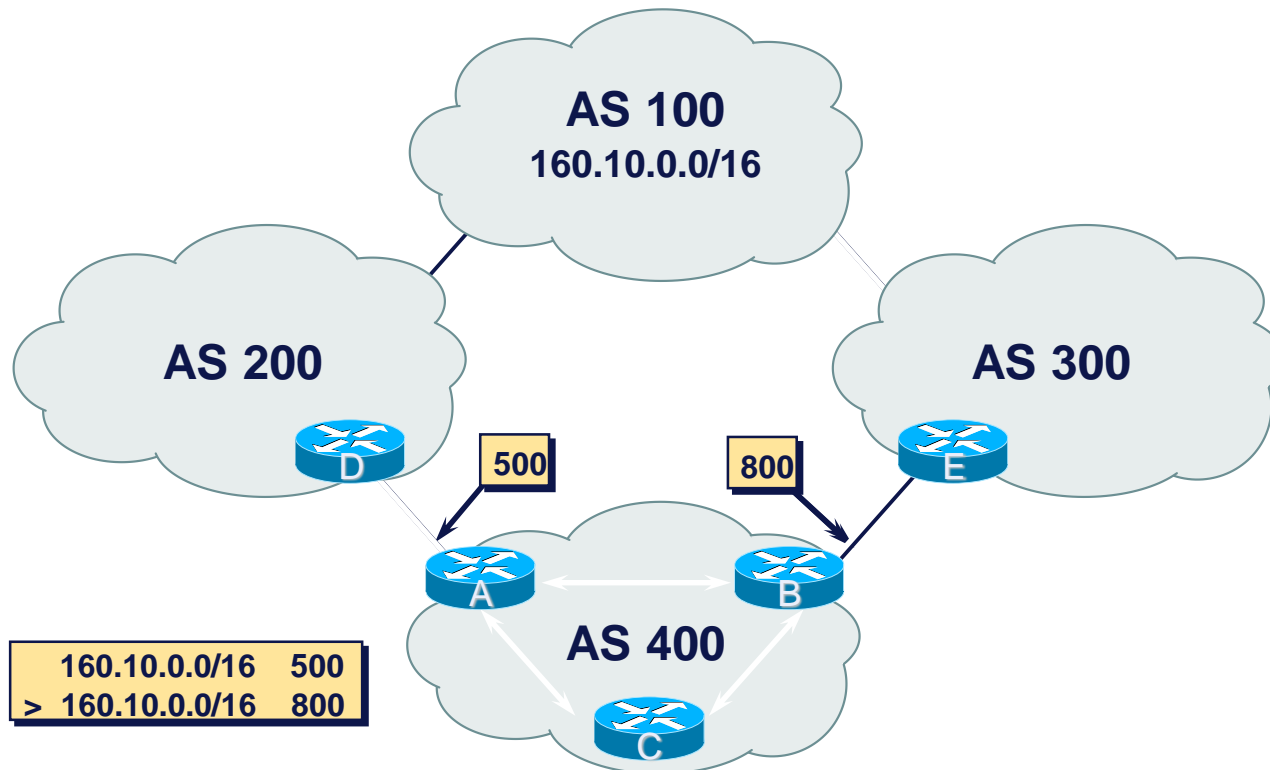
- Les attributs d'un chemin ont quatre types
 - **Well-known mandatory** : Doit être reconnu par tous les routeurs BGP, présent dans toutes les mises à jour BGP et transmis aux autres routeurs BGP.
 - **Well-known discretionary** : Doit être reconnu par tous les routeurs BGP et transmis aux autres routeurs mais n'a pas besoin d'être présent dans une MAJ
 - **Optional transitive** : Peut ou non être reconnu par un routeur BGP mais transmis à d'autres routeurs BGP.
 - **Optional non-transitive** : Peut ou non être reconnu par un routeur BGP et n'est pas transmis à d'autres routeurs

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	—

BGP - Attributs d'un Chemin

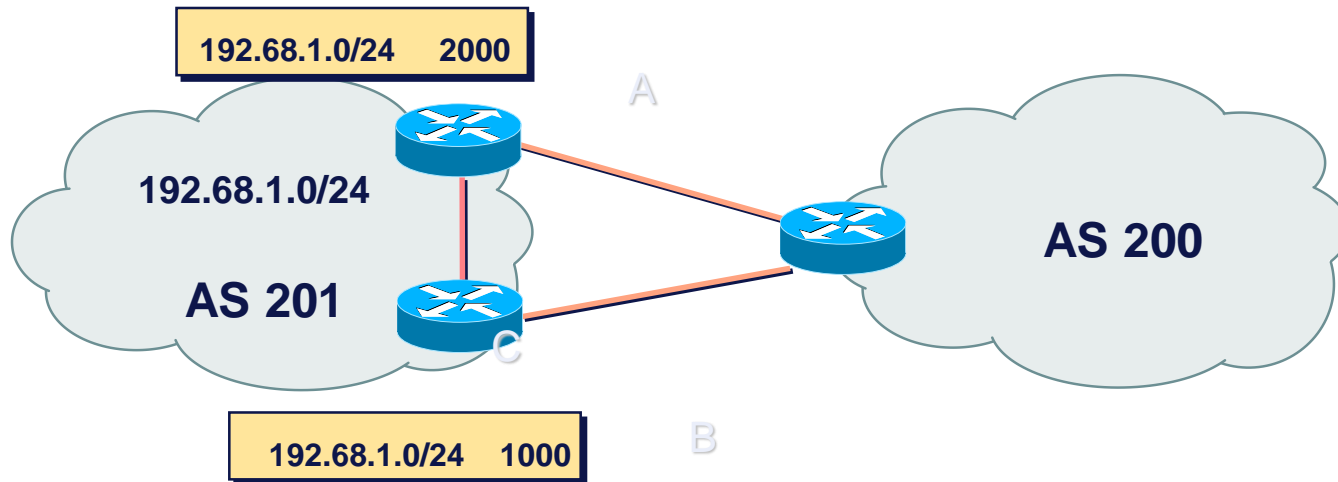
- BGP choisit un seul chemin
- Quand un chemin est choisi, **il est placé dans la table de routage et annoncé aux autres routeurs**
- L'ordre de choix :
 1. Si le prochain saut est inaccessible, ignore le chemin
 2. Préférence aux chemins avec **le plus grand poids**
 3. Préférence aux chemins avec **la 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**
 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**.

BGP – exemple d'attributs - Local Preference



- Local à l'AS – non-transitive
 - local preference est mis à 100 par défaut quand l'information est reçu d'un AS voisin
- Utilisé pour influencer le choix du chemin BGP
 - Déterminer le meilleur chemin pour le trafic sortant
- Le chemin avec la valeur la plus élevée est choisi

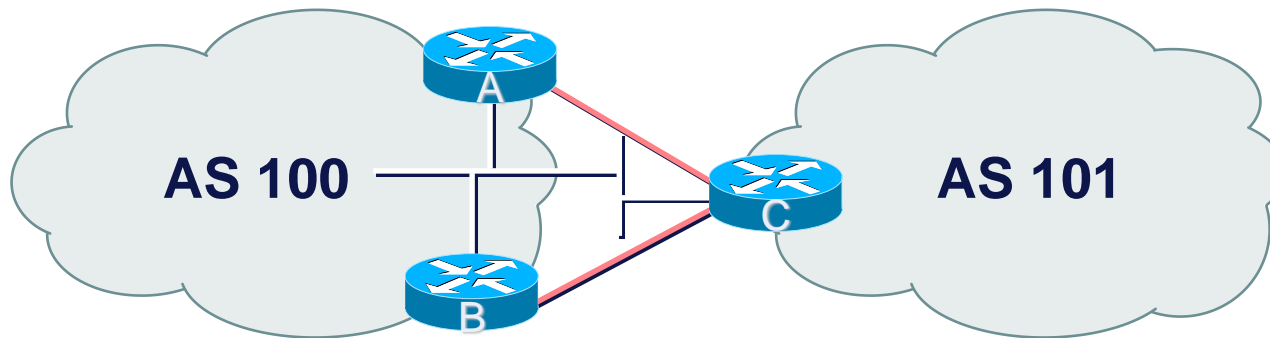
BGP – exemple d'attributs - Multi-Exit Discriminator (MED)



- Inter-AS – non-transitive
 - métrique remise à 0 lors de l'annonce au prochain AS
- Utilisé pour transmettre la préférence relative des points d'entrée à l'AS
 - détermine le meilleur chemin pour le trafic entrant, celui avec la plus petite valeur du MED
- La métrique IGP peut être transmise comme MED
 - set metric-type internal
 - permettre à BGP d'annoncer un MED qui correspond aux valeurs de métrique IGP

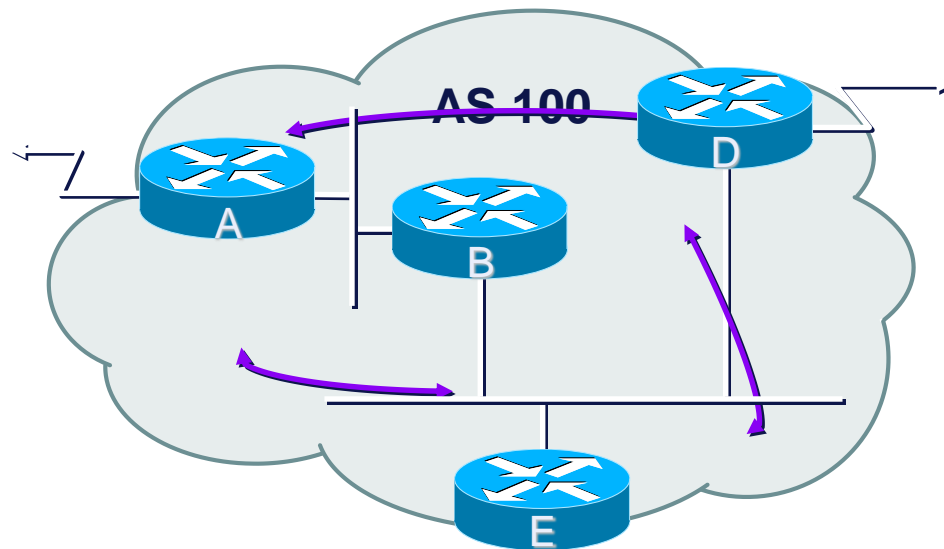
External BGP (eBGP)

- Entre des paires BGP dans des AS différents
- Doivent être directement connectés (sauf exception !)
 - Car le TTL est de 1 par défaut !
- Pas d'IGP entre les pairs eBGP !
- Appliqué plus dans les retours de frontière



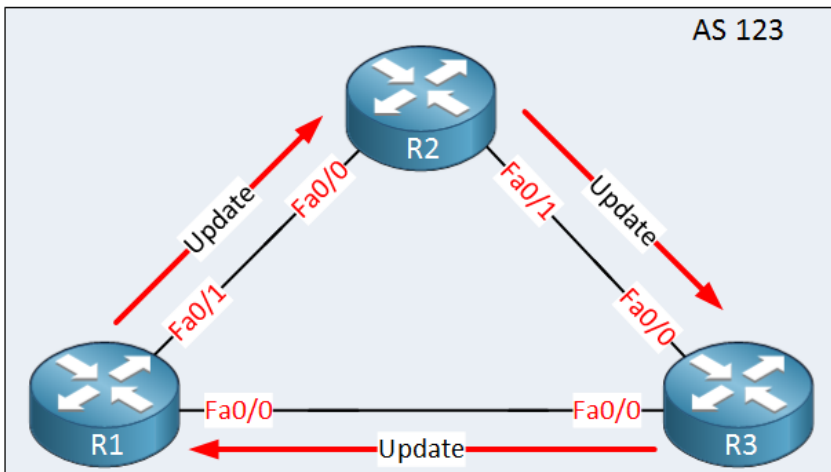
Internal BGP (iBGP)

- Entre des paires BGP dans le même AS
- Pas besoin d'avoir des connexions directes (Indépendant de la topologie)
- Chaque nœud iBGP doit former un lien de peering avec tous les autres nœuds iBGP de l'AS → full mesh
- Appliqué plus dans les retours d'un AS de transit

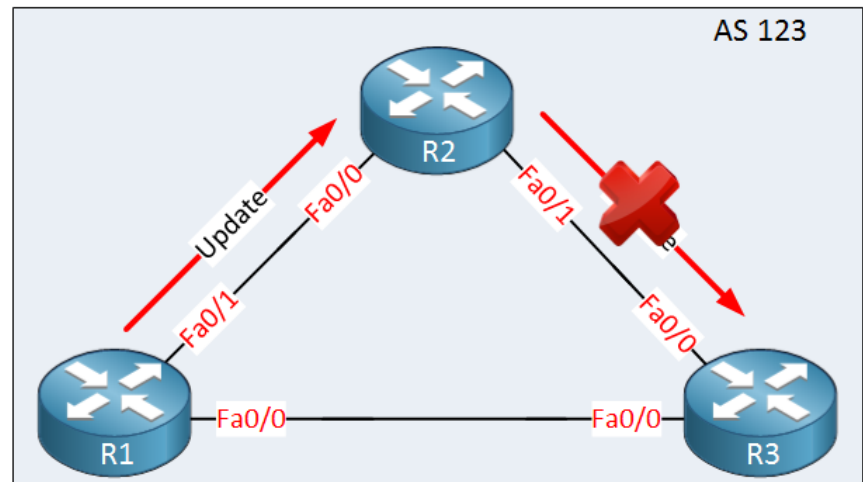


Internal BGP (iBGP)

- Comment éviter les boucles ? l'AS-PATH n'a de sens qu'entre AS !



R1 pourrait recevoir une MAJ d'un préfixe qu'il a lui-même annoncé !

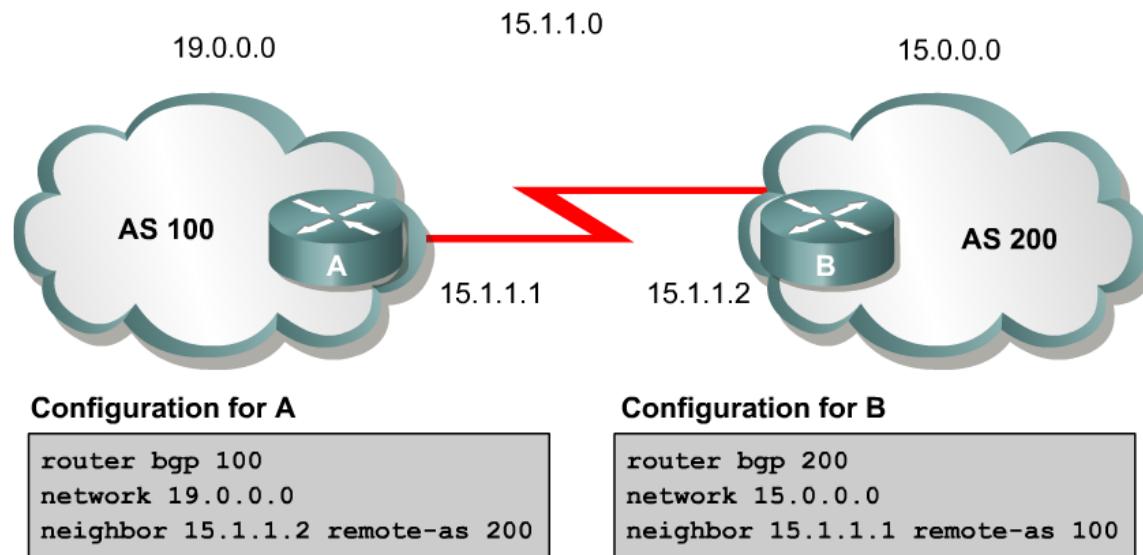


R2 ne transmettra jamais les préfixes iBGP qu'il apprend de R1 vers R3

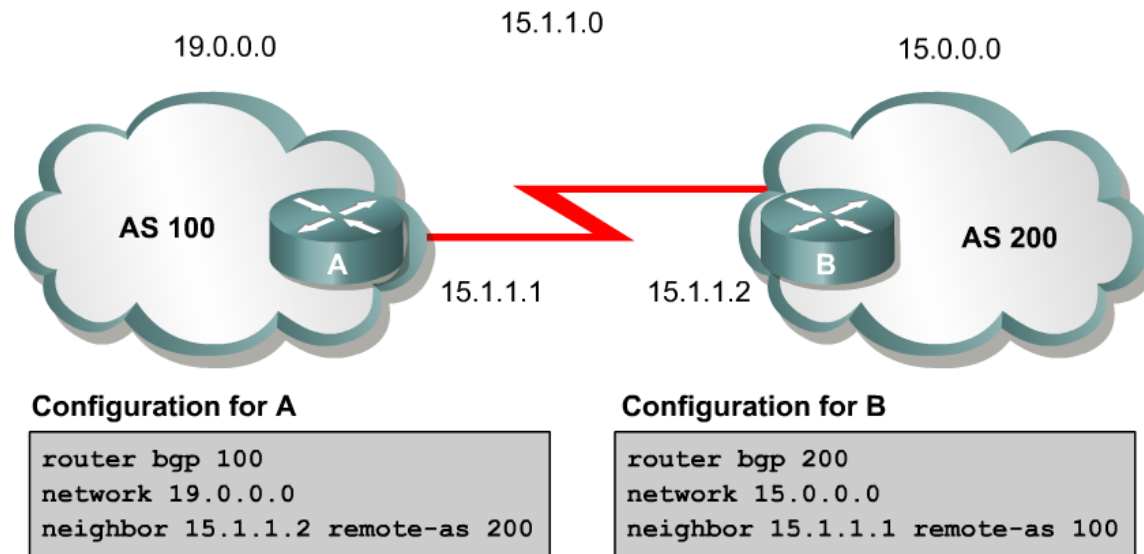
- C'est la règle iBGP Split horizon
- Du coup il faut créer des sessions iBGP avec tout le monde ! → full-mesh
- Donc si on a n nœuds, on doit créer $n(n-1)/2$ sessions BGP !?

Configurer BGP

- Pour commencer, il faut entrer dans le mode configuration BGP
 - Router(config)#**router bgp AS-number**
- Attention, les commandes BGP sont presque similaires à ceux des IGP mais le fonctionnement peut être assez différent
- **Note** : L'IOS Cisco ne permet qu'un seul processus BGP par routeur
 - donc un routeur n'appartient qu'à un seul AS



Configurer BGP

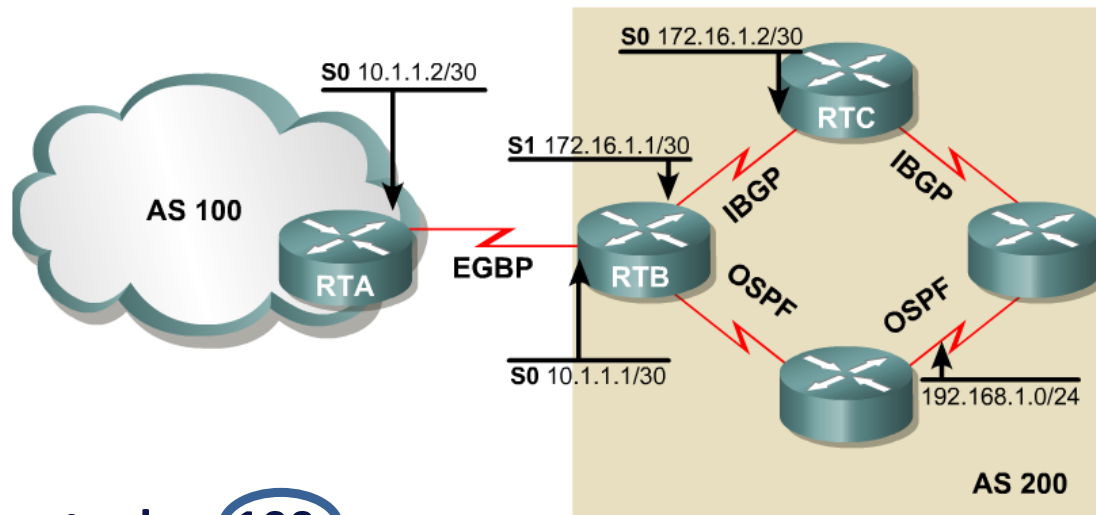


- Déclaration du voisinage
 - Router(config-router)#**neighbor ip-address remote-as AS-number**
 - La commande **neighbor** est utilisée pour établir une relation de voisinage avec un autre routeur BGP
 - Le paramètre **AS-number** va indiquer si le routeur est un voisin eBGP ou iBGP
 - Même valeur que le numéro d'AS en cours ➔ iBGP
 - Valeur différente du numéro d'AS en cours ➔ eBGP

Configurer BGP

- Announcement des reseaux :
 - Router(config-router)#**network network-number [mask networkmask]**
- 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
- Avec les **IGPs**, la commande **network** indique les interfaces concernées ainsi que les routes annoncées
- Avec **BGP**, la commande **network** n'affecte pas les interfaces. Elle indique les routes apprises localement qui seront annoncées
- Les réseaux peuvent être des routes connectées, statiques ou apprises dynamiquement

Exemple 1 : eBGP VS iBGP



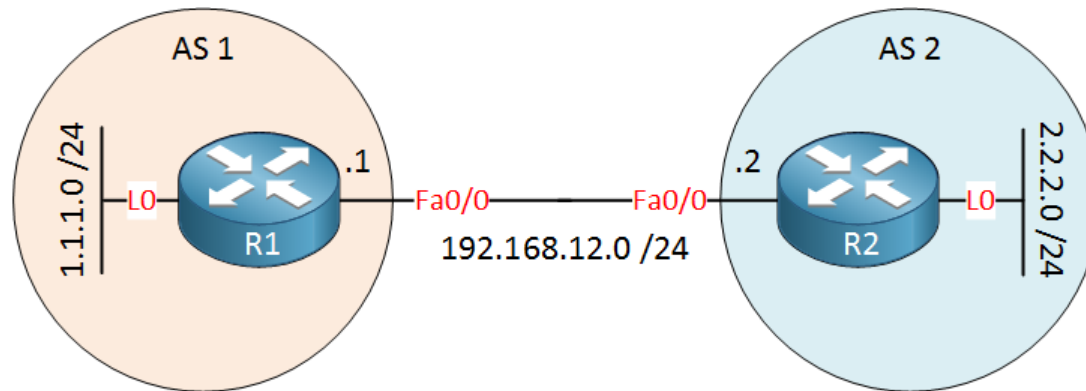
RTA(config)#**router bgp**100
 RTA(config-router)#**neighbor 10.1.1.1 remote-as**200
 RTB(config)#**router bgp**200
 RTB(config-router)#**neighbor 10.1.1.2 remote-as**100

➔ eBGP

- 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

➔ iBGP

Exemple 2 : configuration et vérification



Démarrage du processus BGP et création des voisinages :

```
R1(config)#router bgp 1
R1(config-router)#neighbor 192.168.12.2 remote-as 2
```

```
R2(config)#router bgp 2
R2(config-router)#neighbor 192.168.12.1 remote-as 1
```

```
R1# %BGP-5-ADJCHANGE: neighbor 192.168.12.2 Up
```

```
R2# %BGP-5-ADJCHANGE: neighbor 192.168.12.1 Up
```

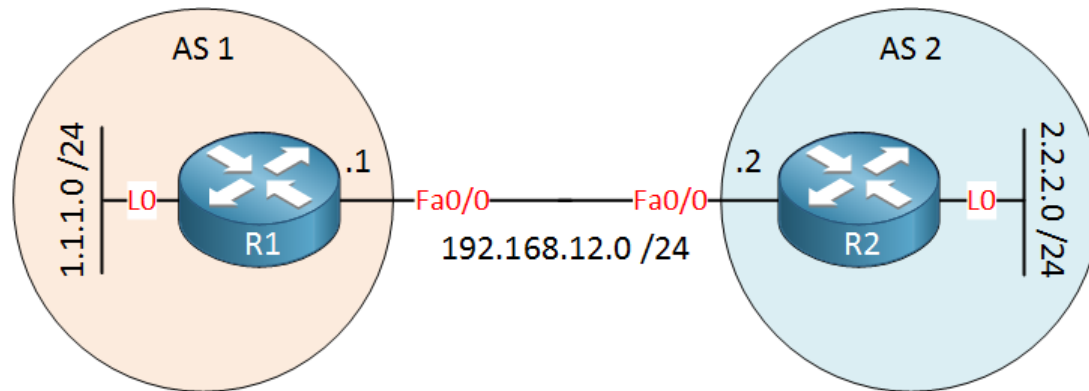
```
R1#show ip bgp summary
BGP router identifier 1.1.1.1, local AS number 1
BGP table version is 1, main routing table version 1
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
192.168.12.2	4	2	10	10	1	0	0	00:07:12	0

```
R2#show ip bgp summary
BGP router identifier 2.2.2.2, local AS number 2
BGP table version is 1, main routing table version 1
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
192.168.12.1	4	1	11	11	1	0	0	00:08:33	0

Exemple 2 : configuration et vérification



Annoncement des réseaux :

```
R1(config)#router bgp 1
```

```
R1(config-router)#network 1.1.1.0 mask 255.255.255.0
```

```
R2(config)#router bgp 2
```

```
R2(config-router)#network 2.2.2.0 mask 255.255.255.0
```

Afficher la base de données BGP :

```
R1#show ip bgp
```

BGP table version is 3, local router ID is 1.1.1.1

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
*> 1.1.1.0/24	0.0.0.0	0		32768	i
*> 2.2.2.0/24	192.168.12.2	0		0	2 i

```
R1#show ip route bgp
```

2.0.0.0/24 is subnetted, 1 subnets

```
B 2.2.2.0 [20/0] via 192.168.12.2, 00:16:13
```

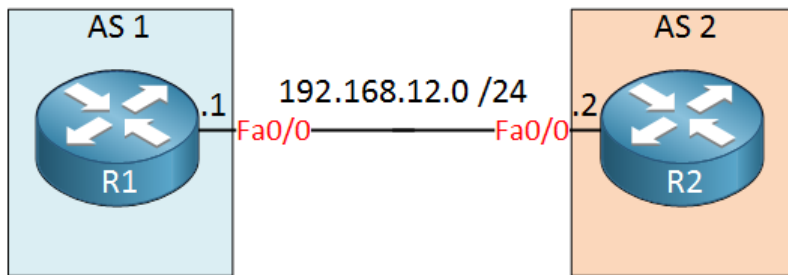
```
R2#show ip route bgp
```

1.0.0.0/24 is subnetted, 1 subnets

```
B 1.1.1.0 [20/0] via 192.168.12.1, 00:16:59
```

Comment annoncer les réseaux en BGP ?

- Il y a deux façons pour annoncer les réseaux avec BGP :
 - Commande network : besoin que le réseau soit dans la table de routage
 - Redistribution



```
R1#show running-config | section bgp
router bgp 1
bgp log-neighbor-changes
neighbor 192.168.12.2 remote-as 2
```

```
R2#show running-config | section bgp
router bgp 2
bgp log-neighbor-changes
neighbor 192.168.12.1 remote-as 1
```

R1#show ip bgp

BGP table version is 2, local router ID is 192.168.12.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
f RT-Filter

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

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 1.1.1.0/24	0.0.0.0	0		32768	i

R2#show ip bgp 1.1.1.1

BGP routing table entry for 1.1.1.0/24, version 2

Paths: (1 available, best #1, table default)

Not advertised to any peer

1

192.168.12.1 from 192.168.12.1 (192.168.12.1)

Origin IGP, metric 0, localpref 100, valid, external, best

```
R1(config)#interface loopback 1
R1(config-if)#ip address 1.1.1.1 255.255.255.0
R1(config)#router bgp 1
R1(config-router)#network 1.1.1.0 mask 255.255.255.0
```

Comment annoncer les réseaux en BGP ?

Et si le réseau qu'on veut annoncer ne figure pas dans la table ?

```
R1(config)#ip route 1.0.0.0 255.0.0.0 null 0
```

```
R1#show ip route 1.0.0.0
```

```
Routing entry for 1.0.0.0/8, 3 known subnets
```

```
Attached (3 connections)
```

```
Variably subnetted with 3 masks
```

```
S      1.0.0.0/8 is directly connected, Null0
C      1.1.1.0/24 is directly connected, Loopback1
L      1.1.1.1/32 is directly connected, Loopback1
```

```
R1(config)#router bgp 1
```

```
R1(config-router)#network 1.0.0.0 mask 255.0.0.0
```

```
R1#show ip bgp 1.0.0.0
```

```
BGP routing table entry for 1.0.0.0/8, version 6
```

```
Paths: (1 available, best #1, table default)
```

```
Advertised to update-groups:
```

```
1
```

```
Local
```

```
0.0.0.0 from 0.0.0.0 (192.168.12.1)
```

```
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local,
best
```

```
R2#show ip bgp 1.0.0.0
```

```
BGP routing table entry for 1.0.0.0/8, version 6
```

```
Paths: (1 available, best #1, table default)
```

```
Not advertised to any peer
```

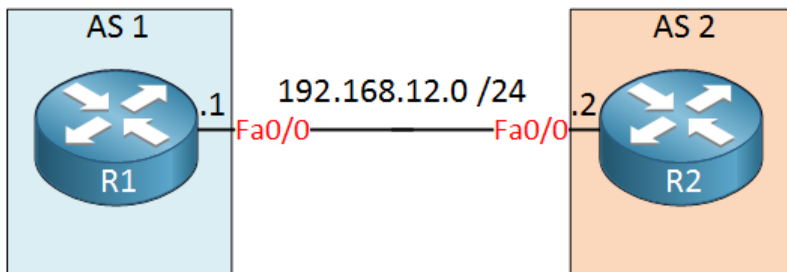
```
1
```

```
192.168.12.1 from 192.168.12.1 (192.168.12.1)
```

```
Origin IGP, metric 0, localpref 100, valid, external, best
```


Comment annoncer les réseaux en BGP ?

■ Redistribution



```
R1(config)#interface loopback 3
R1(config-if)#ip address 111.111.111.111 255.255.255.0
R1(config-if)#exit
```

```
R1(config)#router ospf 1
R1(config-router)#network 111.111.111.0 0.0.0.255 area 0
```

```
R1(config)#router bgp 1
R1(config-router)#redistribute ospf 1
```

```
R1#show ip bgp | begin Network
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 1.0.0.0	0.0.0.0	0		32768	i
*> 1.1.1.0/24	0.0.0.0	0		32768	i
*> 11.11.11.11/32	0.0.0.0	0		32768	i
*> 111.111.111.0/24	0.0.0.0	0		32768	?

```
R2#show ip bgp | begin Network
```

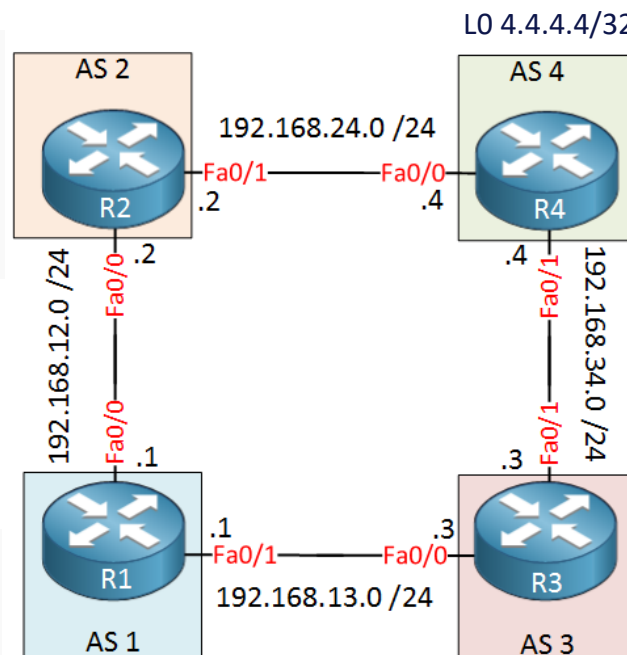
Network	Next Hop	Metric	LocPrf	Weight	Path
*> 1.0.0.0	192.168.12.1	0		0 1	i
*> 1.1.1.0/24	192.168.12.1	0		0 1	i
*> 11.11.11.11/32	192.168.12.1	0		0 1	i
*> 111.111.111.0/24	192.168.12.1	0		0 1	?

Comment lire les informations BGP ?

- Soit la configuration suivante :

```
R2#show run | section bgp
router bgp 2
no synchronization
bgp log-neighbor-changes
neighbor 192.168.12.1 remote-as 1
neighbor 192.168.24.4 remote-as 4
no auto-summary
```

```
R1#show run | section bgp
router bgp 1
no synchronization
bgp log-neighbor-changes
neighbor 192.168.12.2 remote-as 2
neighbor 192.168.13.3 remote-as 3
no auto-summary
```



```
R4#show run | section bgp
router bgp 4
no synchronization
bgp log-neighbor-changes
network 4.4.4.4 mask 255.255.255.255
neighbor 192.168.24.2 remote-as 2
neighbor 192.168.34.3 remote-as 3
no auto-summary
```

```
R3#show run | section bgp
router bgp 3
no synchronization
bgp log-neighbor-changes
neighbor 192.168.13.1 remote-as 1
neighbor 192.168.34.4 remote-as 4
no auto-summary
```

- Liaisons eBGP avec R4 qui annonce le réseau 4.4.4.4/32

Comment lire les informations BGP ?

R4#show ip bgp

BGP table version is 2, local router ID is 192.168.34.4

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
*> 4.4.4.4/32	0.0.0.0	0		32768	i

R2#show ip bgp

BGP table version is 2, local router ID is 192.168.24.2

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
*> 4.4.4.4/32	192.168.24.4	0		0 4	i

- Affichage de la BD BGP
- * : la route est valide et BGP peut l'utiliser
- > : ce chemin est choisi comme le meilleur
- **0.0.0.0** : l'entrée provient de ce retour
- **Path** : c'est l'AS path
- **Version** : incrémenté à chaque changement du meilleur chemin

Comment lire les informations BGP ?

```
R1#show ip bgp
```

```
BGP table version is 2, local router ID is 192.168.13.1
```

```
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
* 4.4.4.4/32	192.168.13.3			0 3 4	i
*>	192.168.12.2			0 2 4	i

➔ Deux chemins possibles et valides pour aller au 4.4.4.4

```
R1#show ip route bgp
```

```
4.0.0.0/32 is subnetted, 1 subnets
```

```
B 4.4.4.4 [20/0] via 192.168.12.2, 00:25:51
```

```
R1#show ip bgp 4.4.4.4
```

```
BGP routing table entry for 4.4.4.4/32, version 2
```

```
Paths: (2 available, best #2, table Default-IP-Routing-Table)
```

```
Advertised to update-groups:
```

```
1
```

```
3 4
```

```
192.168.13.3 from 192.168.13.3 (192.168.34.3)
```

```
Origin IGP, localpref 100, valid, external
```

```
2 4
```

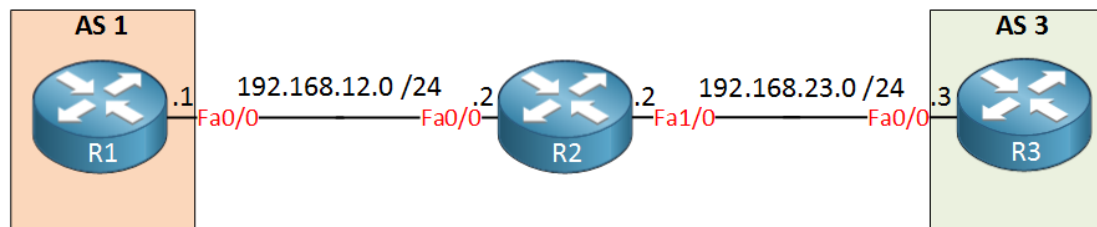
```
192.168.12.2 from 192.168.12.2 (192.168.24.2)
```

```
Origin IGP, localpref 100, valid, external, best
```



eBGP Multihop – cas 1

- **eBGP** a besoin des pairs directement connectés pour établir une session
- Exception : l'IOS Cisco a une option **multihop** pour permettre la connexion logique entre des routeurs non directement connectés et qui se trouvent dans des AS différents (même si le routeur intermédiaire n'implémente pas BGP)
 - équivalent à augmenter le TTL pour une annonce BGP
- Cette option doit être configurée pour chaque pair
 - Router(config-router)#**neighbor IP-address ebgp-multihop [hops]**



```
R1(config)#ip route 192.168.23.3 255.255.255.255 192.168.12.2
```

```
R1(config)#router bgp 1
```

```
R1(config-router)#neighbor 192.168.23.3 remote-as 3
```

```
R1#show ip bgp neighbors | include External
```

```
External BGP neighbor not directly connected.
```

```
R1(config-router)#neighbor 192.168.23.3 ebgp-multihop 2
```

```
R3(config)#ip route 192.168.12.1 255.255.255.255 192.168.23.2
```

```
R3(config)#router bgp 3
```

```
R3(config-router)#neighbor 192.168.12.1 remote-as 1
```

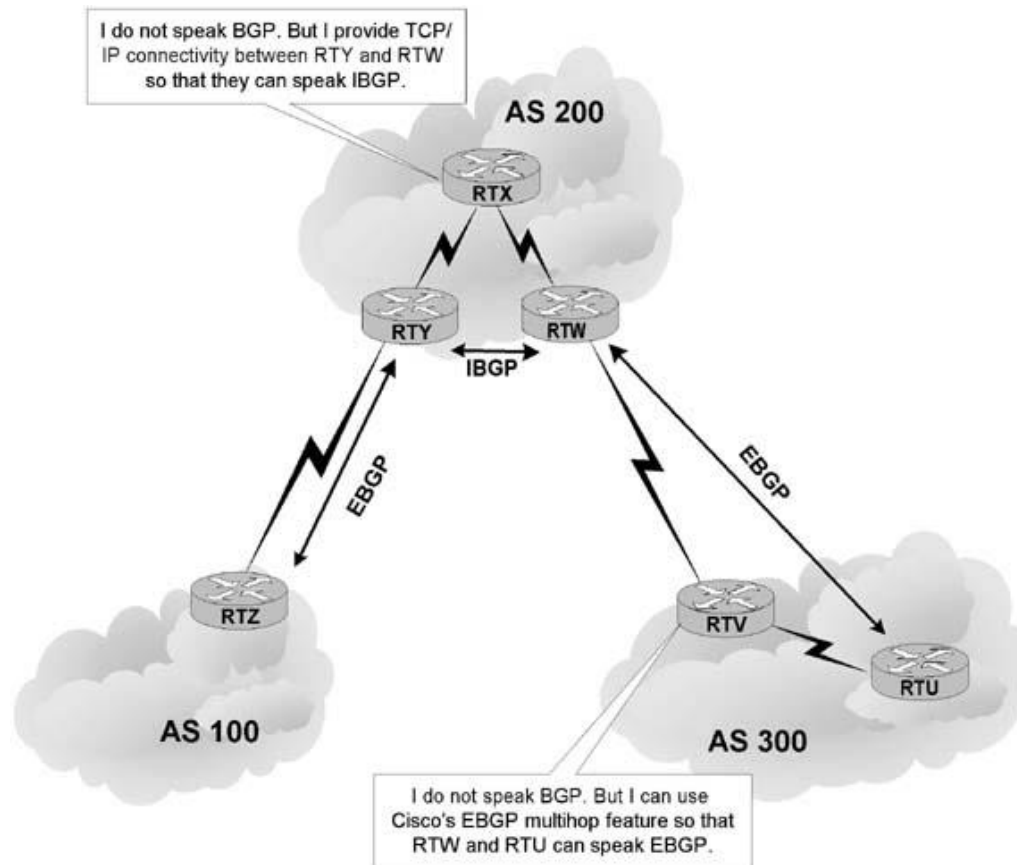
```
R3#show ip bgp neighbors | include External
```

```
External BGP neighbor not directly connected.
```

```
R3(config-router)#neighbor 192.168.12.1 ebgp-multihop 2
```

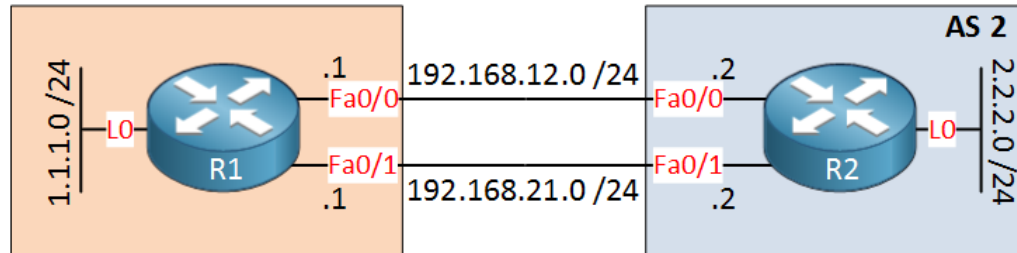
eBGP Multihop – cas 1

- L'option **multihop** va permettre la connexion logique entre RTW et RTU même si RTV n'implémente pas BGP



eBGP Multihop – cas 2

- Utilisation des interfaces loopback :



- Mise en place de l'équilibrage de charge :

```
R1(config)#ip route 2.2.2.0 255.255.255.0 192.168.12.2
R1(config)#ip route 2.2.2.0 255.255.255.0 192.168.21.2
R2(config)#ip route 1.1.1.0 255.255.255.0 192.168.12.1
R2(config)#ip route 1.1.1.0 255.255.255.0 192.168.21.1
```

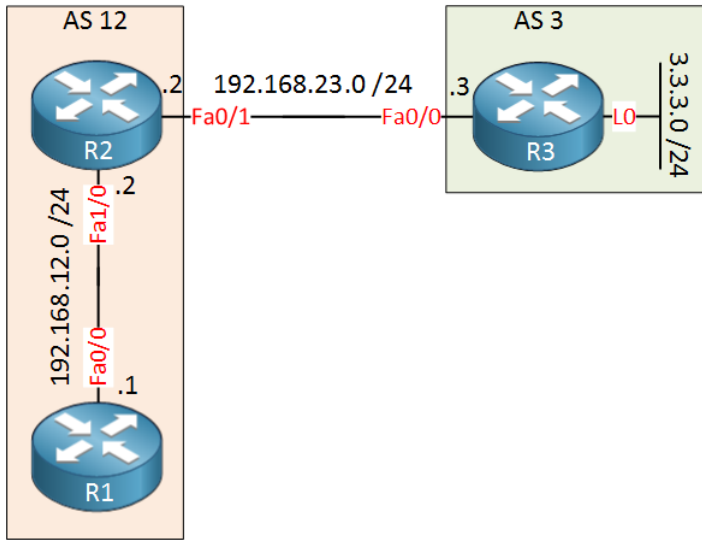
- Configuration de la session eBGP, avec quelle interface ?

```
R1(config)#router bgp 1
R1(config-router)#neighbor 2.2.2.2 remote-as 2
R1(config-router)#neighbor 2.2.2.2 update-source loopback 0
R1(config-router)#neighbor 2.2.2.2 ebgp-multihop 2
R2(config)#router bgp 2
R2(config-router)#neighbor 1.1.1.1 remote-as 1
R2(config-router)#neighbor 1.1.1.1 update-source loopback 0
R2(config-router)#neighbor 1.1.1.1 ebgp-multihop 2
```

- Avec la loopback, même si un des liens physiques est HS, la session eBGP restera opérationnelle (up)
- La commande **update-source** dit aux routeurs d'utiliser l'adresse IP sur l'**interface indiquée (lo0)** comme adresse IP source pour l'adjacence eBGP.
- Cela implique l'augmentation du nombre de saut d'où la commande **ebgp-multihop**

iBGP Next Hop Self

- Soit la configuration suivante avec une session iBGP entre R1 et R2, et une session eBGP entre R2 et R3 :



R2#show ip bgp

BGP table version is 2, local router ID is 192.168.23.2

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
*> 3.3.3.0/24	192.168.23.3	0		0 3	i

R1#show ip bgp

BGP table version is 1, local router ID is 192.168.12.1

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
* i3.3.3.0/24	192.168.23.3	0	100	0 3	i

- R1 a appris de R2 l'existence du préfixe 3.3.3.0/24 et qu'il faut passer par 192.68.23.3 ! Or R1 ne connaît pas comment y aller !

iBGP Next Hop Self

- Solution 1 : annoncer le réseau !

```
R2(config)#router bgp 12
R2(config-router)#network 192.168.23.0 mask 255.255.255.0
```

```
R1#show ip bgp
BGP table version is 3, local router ID is 192.168.12.1
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
*>i3.3.3.0/24	192.168.23.3	0	100	0	3 i
*>i192.168.23.0	192.168.12.2	0	100	0	i

```
R1#show ip route bgp
  3.0.0.0/24 is subnetted, 1 subnets
B       3.3.3.0 [200/0] via 192.168.23.3, 00:16:20
B    192.168.23.0/24 [200/0] via 192.168.12.2, 00:16:25
```

- Résolution récursive !!

iBGP Next Hop Self

- Solution 2 : modifier le next hop !

```
R2(config)#router bgp 12
```

```
R2(config-router)#neighbor 192.168.12.1 next-hop-self
```

```
R1#show ip bgp
```

```
BGP table version is 6, local router ID is 192.168.12.1
```

```
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
*>13.3.3.0/24	192.168.12.2	0	100	0 3	i

```
R1#show ip route bgp
```

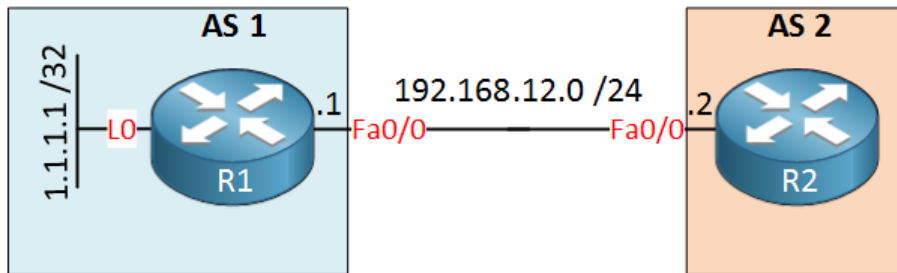
```
3.0.0.0/24 is subnetted, 1 subnets
```

```
B 3.3.3.0 [200/0] via 192.168.12.2, 00:00:33
```

- Avec **next-hop-self** R2 met son adresse comme next-hop dans les MAJ iBGP
- Une seule résolution est nécessaire

BGP Auto-Summary

- Avec BGP il faut annoncer le réseau exacte avec son masque, sinon il ne sera pas placé dans la table BGP



```
R1(config)#router bgp 1
R1(config-router)#network 1.0.0.0
```

```
R1(config-router)#auto-summary
```

- Avec **auto-summary** (activé par défaut) on peut annonce un réseau **classful** sans préciser le masque. Dans ce cas BGP ajoutera automatiquement dans sa base le réseau classful ou le sous réseau correspondant s'ils figurent dans la table de routage

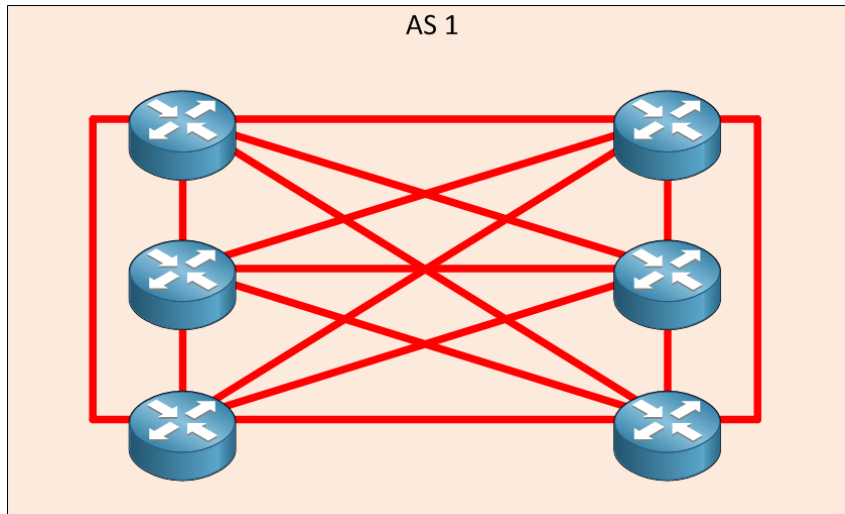
```
R1#show ip bgp 1.0.0.0
BGP routing table entry for 1.0.0.0/8, version 2
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    1
  Local
    0.0.0.0 from 0.0.0.0 (192.168.12.1)
      Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local,
best
```

```
R2#show ip bgp
BGP table version is 2, local router ID is 192.168.12.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, x best-external,
f RT-Filter
Origin codes: i - IGP, e - EGP, ? - incomplete

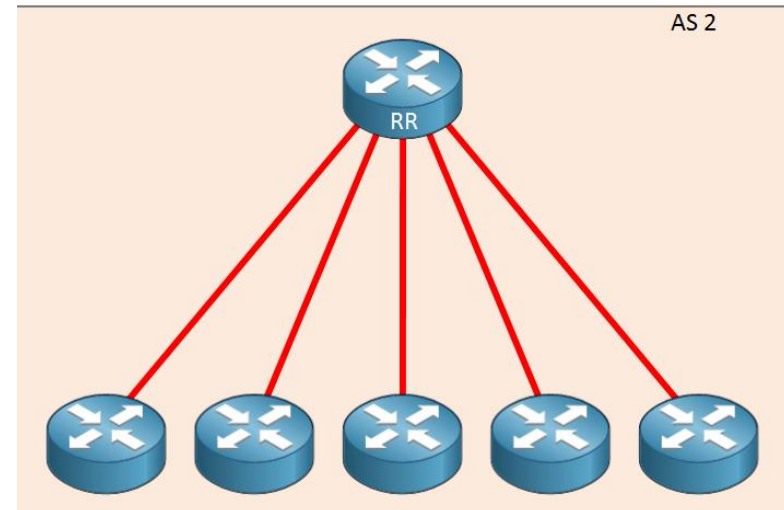
   Network        Next Hop           Metric LocPrf Weight Path
*> 1.0.0.0        192.168.12.1         0             0 1 i
```

BGP Route Reflector

- Avec **iBGP**, chaque routeur doit former un lien de peering avec tous les autres nœuds **iBGP** de l'AS → full mesh



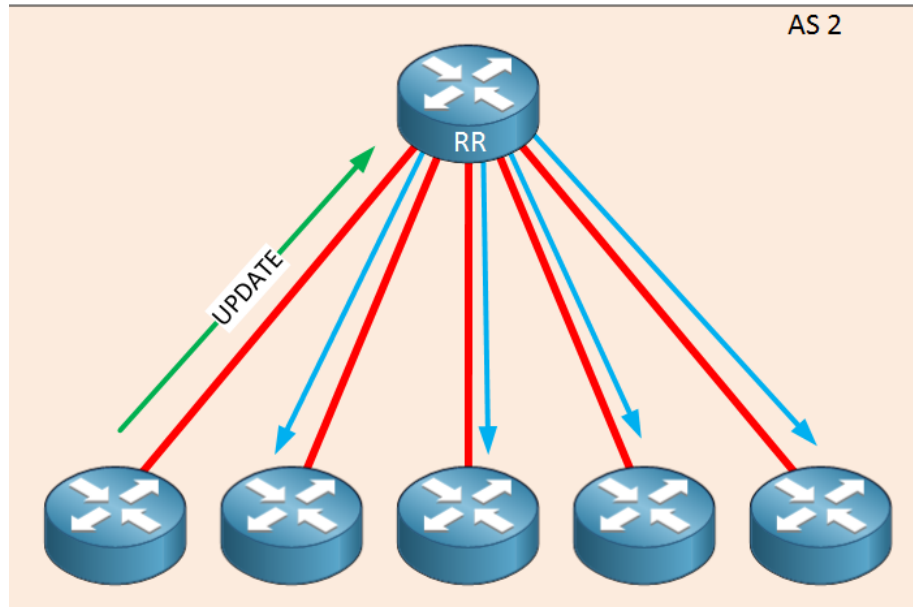
full mesh → $N*(N-1)/2$ sessions !



Résultat obtenu avec route reflector (RR)

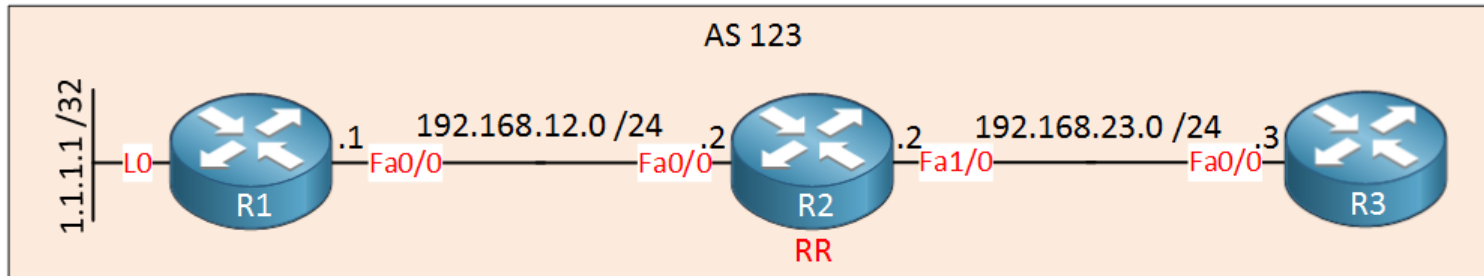
- Chaque routeur va garder une seule session **iBGP** avec le **RR**

BGP Route Reflector



- Quand un routeur iBGP annonce un chemin vers le RR, il sera «reflété» vers tous les autres routeurs iBGP
- Cela simplifie beaucoup la configuration iBGP
- Et si le RR crache !? ➔ utiliser plusieurs RR

BGP Route Reflector : Configuration



```
R1(config)#router bgp 123
R1(config-router)#neighbor 192.168.12.2 remote-as 123
```

```
R3(config)#router bgp 123
R3(config-router)#neighbor 192.168.23.2 remote-as 123
```

- Déclarer les clients du RR :

```
R2(config)#router bgp 123
R2(config-router)#neighbor 192.168.12.1 remote-as 123
R2(config-router)#neighbor 192.168.12.1 route-reflector-client
R2(config-router)#neighbor 192.168.23.3 remote-as 123
R2(config-router)#neighbor 192.168.23.3 route-reflector-client
```

- Annoncer le réseau 1.1.1.1 :

```
R1(config)#router bgp 123
R1(config-router)#network 1.1.1.1 mask 255.255.255.255
```

```
R2#show ip bgp 1.1.1.1
BGP routing table entry for 1.1.1.1/32, version 2
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Flag: 0x820
  Advertised to update-groups:
    1
  Local, (Received from a RR-client)
    192.168.12.1 from 192.168.12.1 (192.168.12.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

```
R2#show ip bgp neighbors 192.168.23.3 advertised-routes
BGP table version is 2, local router ID is 192.168.23.2
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
  *>i1.1.1.1/32      192.168.12.1          0    100     0  i
```

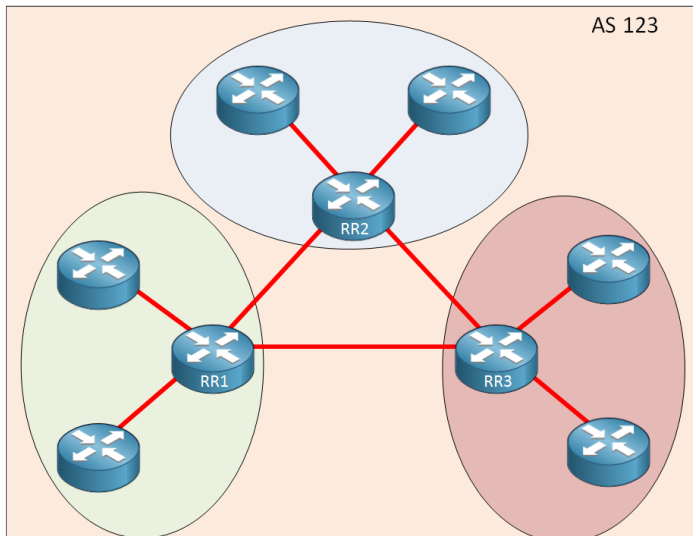
- Sans le RR, R2 ne vas pas partager l'information sur 1.1.1.1 vers R3

BGP Route Reflector : Configuration

```
R3#show ip bgp 1.1.1.1
BGP routing table entry for 1.1.1.1/32, version 0
Paths: (1 available, no best path)
Not advertised to any peer
Local
  192.168.12.1 (inaccessible) from 192.168.23.2 (192.168.23.2)
    Origin IGP, metric 0, localpref 100, valid, internal
    Originator: 192.168.12.1, Cluster list: 192.168.23.2
```

R3 a appris la route de R2 (RR) avec en plus deux nouvelles informations importantes :

- **Originator ID** : pour éviter les boucles comme la règle de split horizon a été cassée
- **Cluster List** : pour identifier le RR et ses clients



- Exemple de topologie avec 3 RR desservant deux voisins iBGP chacun
- Il est possible qu'une boucle se crée entre ces RR, donc lorsqu'un RR voit son propre ID de cluster, il abandonne le paquet

BGP – commandes de vérification

Command	Description
<code>show ip bgp</code>	Affiche les entrées dans la table BGP
<code>show ip bgp neighbors</code>	Affiche des informations sur la connexion TCP et BGP avec les voisins
<code>show ip bgp summary</code>	Affiche l'état des connexions
<code>show ip bgp neighbors {address} advertised-routes</code>	Affiche toutes les routes qui ont été annoncées par les voisins
<code>show ip bgp rib-failure</code>	Affiche les routes BGP qui n'ont pas été installées dans la RIB (routing information base), et la raison pour cela
<code>debug ip bgp</code> <code>[dampening events </code> <code>keepalives updates]</code>	