

Introduction à BGP



AFNOG 2014

Border Gateway Protocol

- ❑ Un protocole de routage utilisé pour l'échange information de routage entre plusieurs réseaux
- ❑ Protocole de routage externe (EGP)
- ❑ Décrit dans le RFC4271
 - RFC4276 donne un rapport d'implémentation BGP
 - RFC4277 décrit les expériences opérationnelles BGP
- ❑ Le système autonome est fondamental dans le fonctionnement de BGP
 - Il est utilisé pour identifier de façon unique les réseaux ayant la même politique de routage.

BGP

- ❑ Path Vector Protocol
- ❑ Des mises à jour incrémentales
- ❑ Plusieurs options pour appliquer une politique de routage
- ❑ Classless Inter Domain Routing (CIDR)
- ❑ Largement utilisé sur les Backbones Internet.
- ❑ Système Autonome

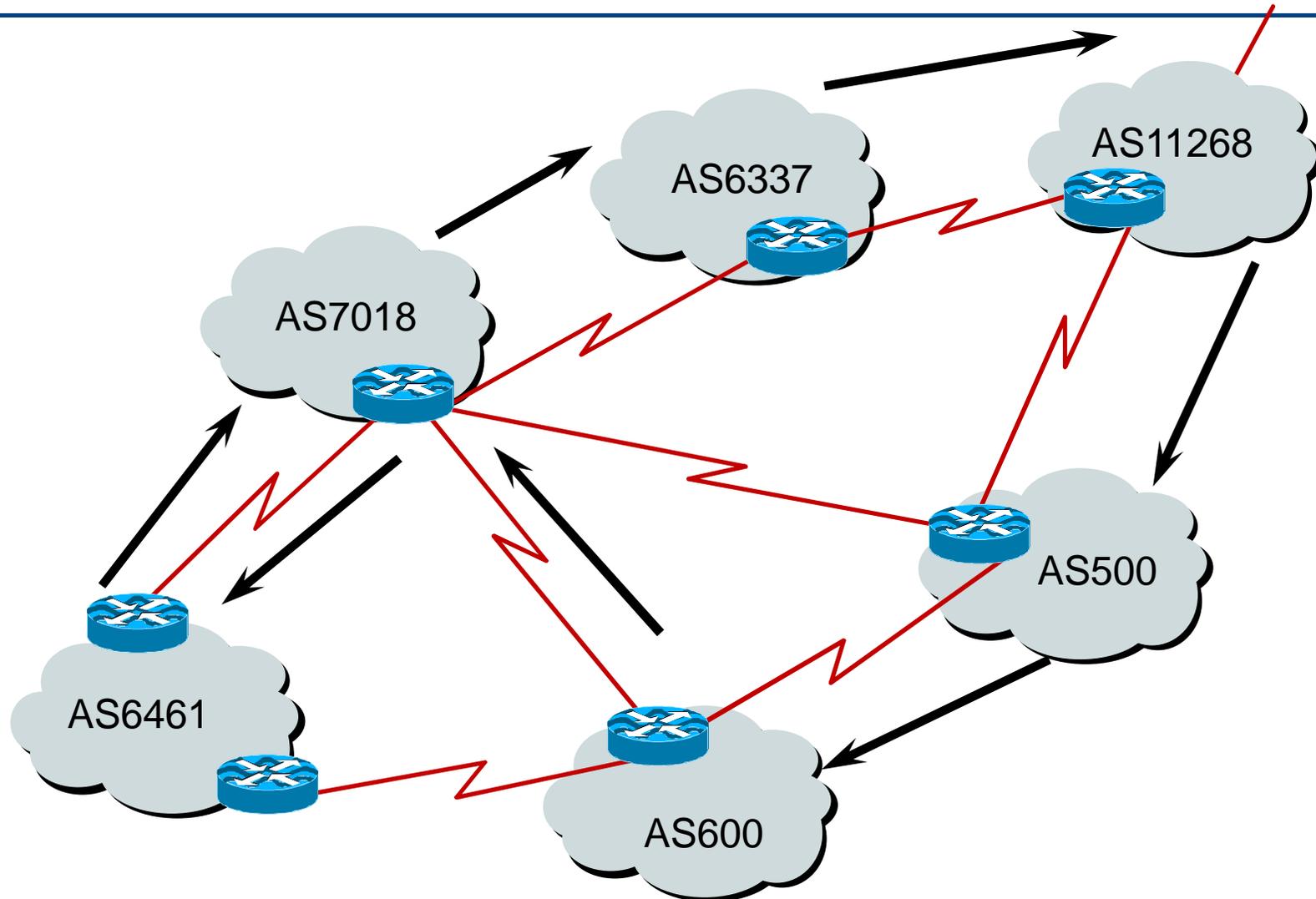
Path Vector Protocol

- BGP est un protocole de routage basé sur les *vecteurs de chemin* (path vector) (voir RFC 1322)
 - Un vecteur de chemin définit une route comme une association entre une destination et les attributs du chemin vers cette destination.

```
12.6.126.0/24 207.126.96.43 1021 0 6461 7018 6337 11268 i
```

AS Path

Path Vector Protocol



Definitions

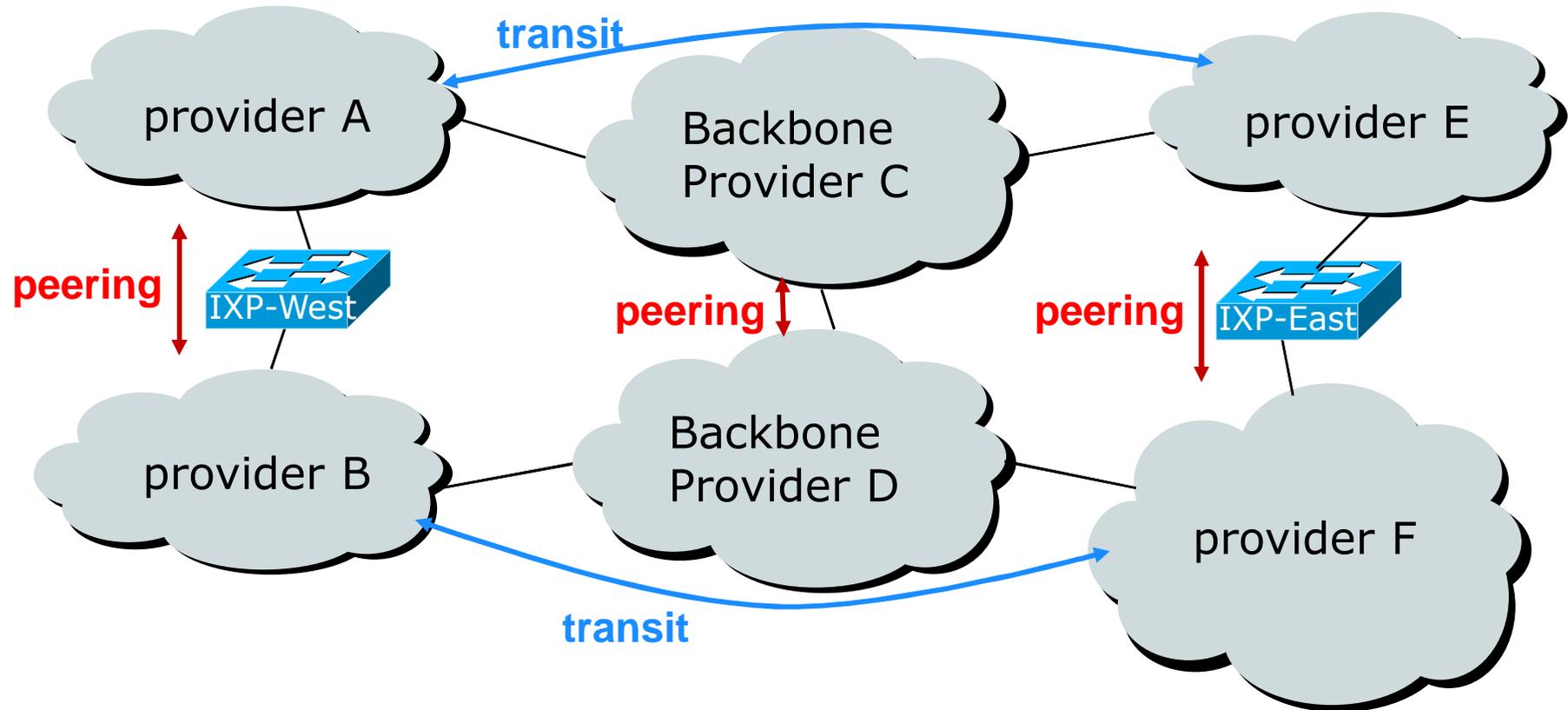
- **Transit** – Transporter du trafic à travers le réseau d'un opérateur.
- **Peering** – échange d'information de routage et de trafic.
- **Default** – Destination par défaut lorsqu'il n'y a pas de routes spécifiques dans la table de routage

Default Free Zone

Zone de routage constituée de routeurs ayant des routes explicites des réseaux d'Internet et donc n'ont pas besoin de route par défaut

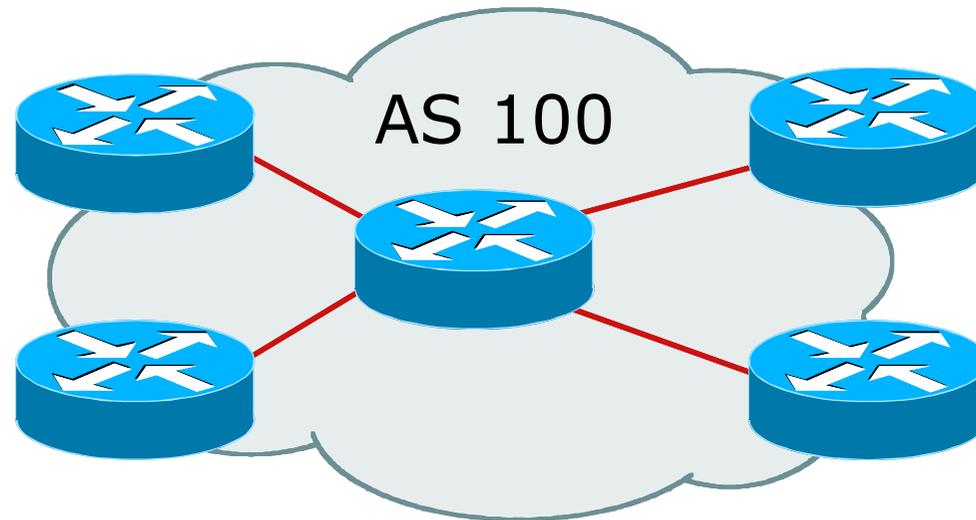
NB: Different pour les IPS liés à un upstream

Exemple de Peering et Transit



A et B ont un accord de peering gratuit, mais requierent un transit avec C et D pour communiquer avec E et F

Autonomous System (AS)



- ❑ Ensemble de réseaux ayant la même politique de routage
- ❑ Un seul protocole de routage
- ❑ En general sous la responsabilité d'une seule autorité d'administration
- ❑ Identifié par un ASN unique de 32-bit

Autonomous System Number (ASN)

- Deux plages
 - 0-65535 (original 16-bit range)
 - 65536-4294967295 (32-bit range – RFC6793)
- Usage:
 - 0 and 65535 (reserved)
 - 1-64495 (public Internet)
 - 64496-64511 (documentation – RFC5398)
 - 64512-65534 (private use only)
 - 23456 (represent 32-bit range in 16-bit world)
 - 65536-65551 (documentation – RFC5398)
 - 65552-4199999999 (public Internet)
 - 4200000000-4294967295 (private use only)
- La représentation 32-bit décrite dans le RFC5396
 - Explique “asplain” (traditional format) comme notation standard

Autonomous System Number (ASN)

- ❑ Les ASNs sont attribués par les registres régionaux (RIR)
- ❑ Ils sont aussi utilisés par les ISP affiliés à des RIR.
- ❑ Actuellement plus de 63487 16-bit ASN assignés par les RIRs
 - Environ 44500 visibles sur Internet
 - Environ 1500 restent pas attribués
- ❑ Chaque RIR a aussi reçu un bloque de 32-bit ASNs
 - Sur 4800 ASN assignés, environ 3700 sont visibles sur Internet
- ❑ See www.iana.org/assignments/as-numbers

Configuration de BGP (Cisco IOS)

- ❑ Cette commande active BGP in Cisco IOS:

```
router bgp 100
```

- ❑ Pour ASNs > 65535, l'ASN peut être configuré soit en format decimal simple ou avec des points:

```
router bgp 131076
```

ou

```
router bgp 2.4
```

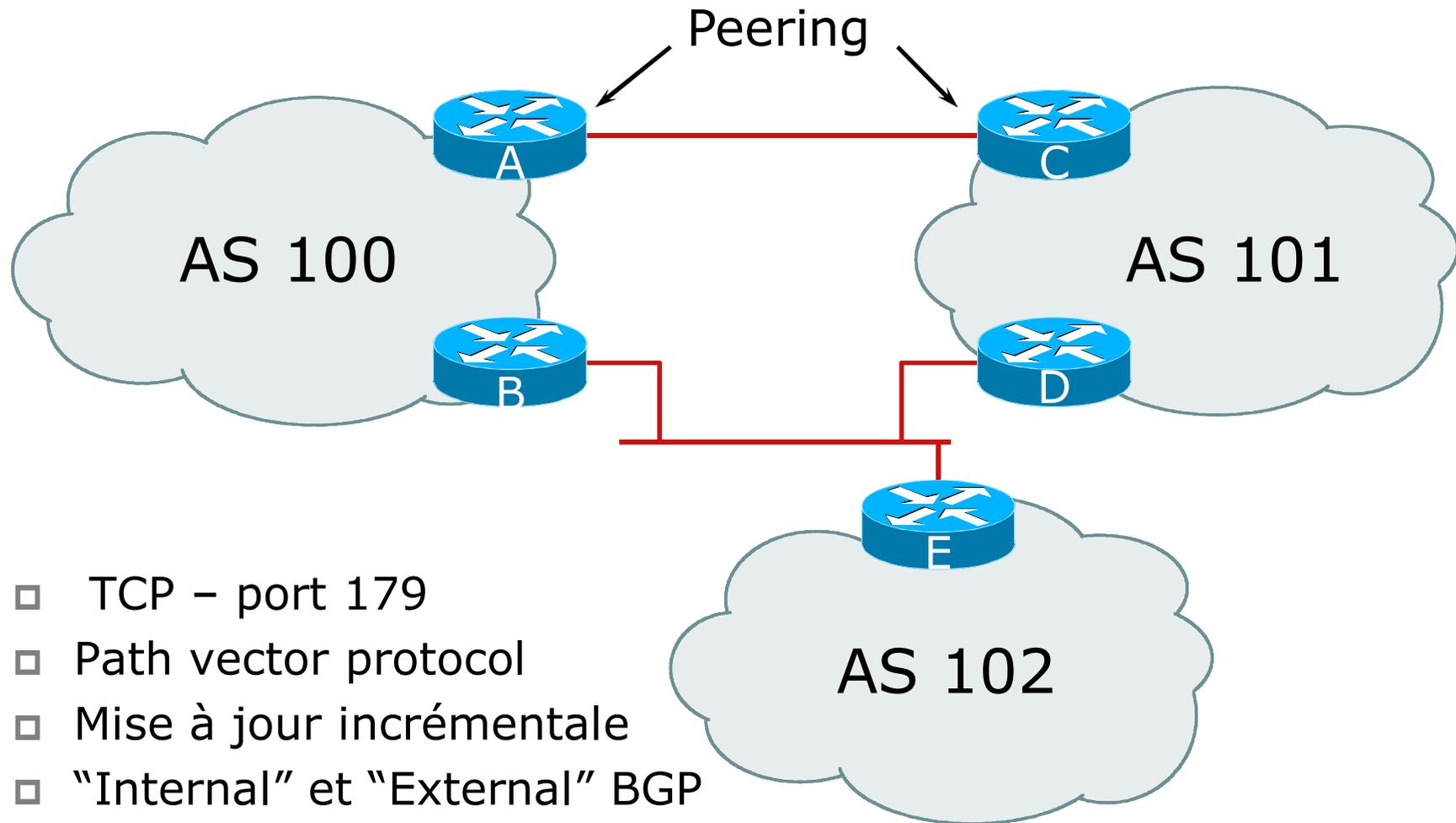
- ❑ L'IOS affiche l'ASN en format decimal par default

- Le format decimal avec des points est optionel :

```
router bgp 2.4
```

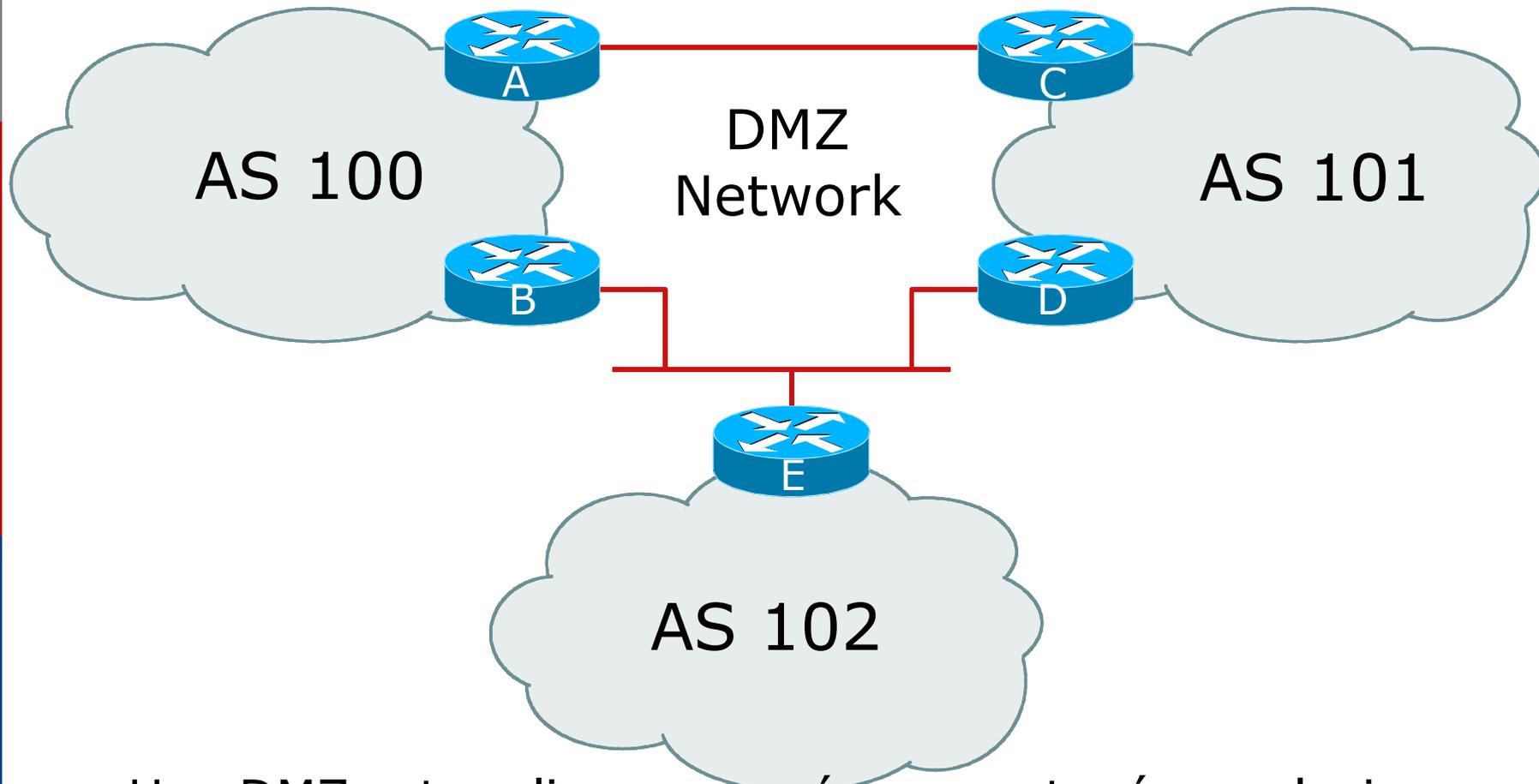
```
bgp asnotation dot
```

Les bases de BGP



- ❑ TCP – port 179
- ❑ Path vector protocol
- ❑ Mise à jour incrémentale
- ❑ "Internal" et "External" BGP

Zone de Demarcation (DMZ)



- Une DMZ est un lien ou un réseau partagé par plusieurs AS

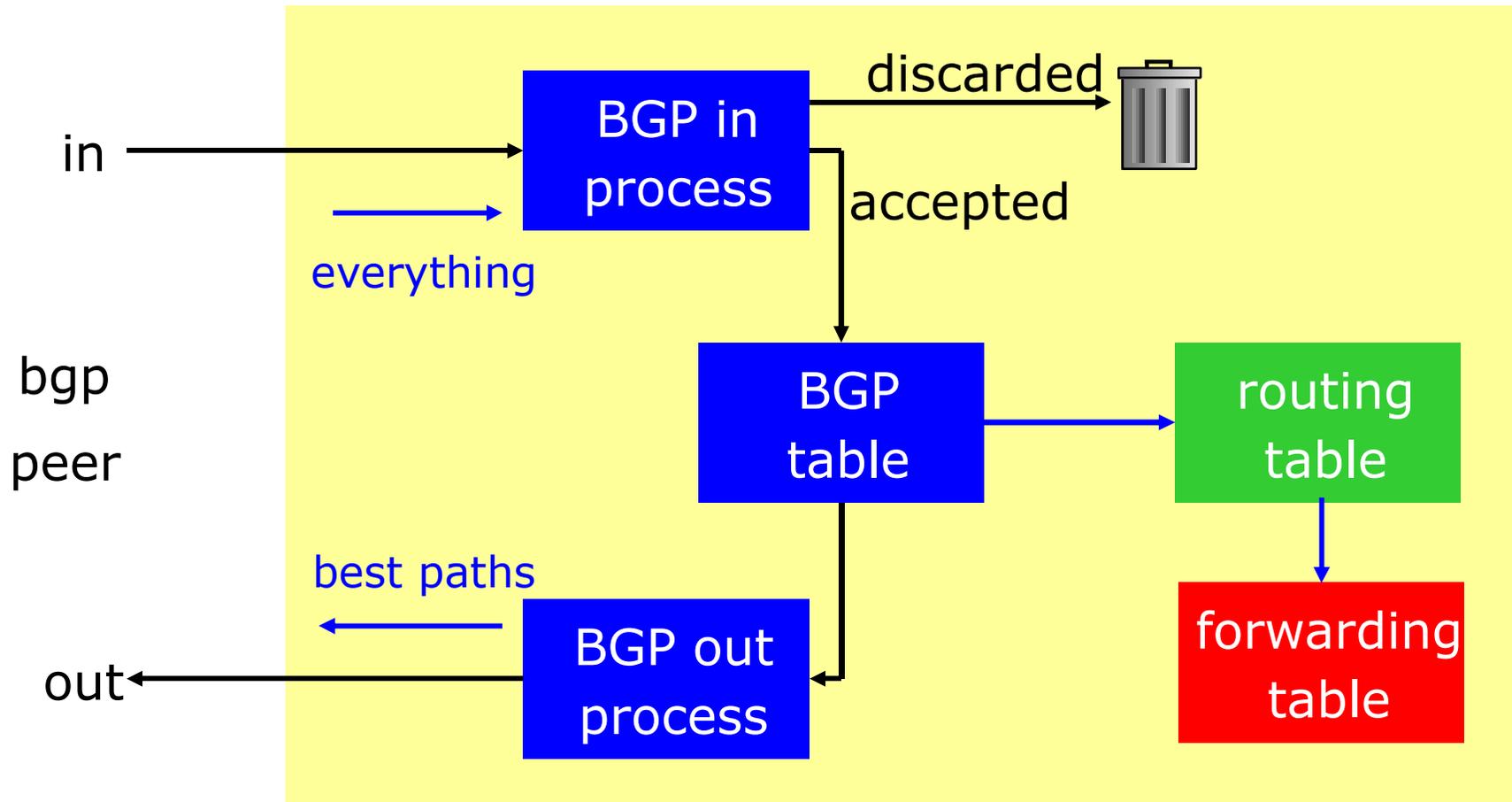
Fonctionnement général

- ❑ BGP apprend plusieurs chemins via les routeurs "peers" qui communiquent via internal et external BGP.
- ❑ Le meilleur chemin est installé dans la table de routage (RIB)
- ❑ Les meilleurs chemins sont exportés vers les routeurs voisins supportant "external BGP"
- ❑ La politique de routage est appliquées en influençant la selection du meilleur chemin.

Création de la table de Forwarding

- BGP “in” fonctionnement
 - Reçois les informations chemin du “peers”
 - Resultat de la selection des chemins est inséré dans la table BGP.
 - “best path” est taggué
- BGP “out” fonctionnement
 - announce les informations des “meilleurs routes” aux peers
- Les meilleures routes sont installées dans la table de routage (RIB)
- Les meilleurs routes sont installées dans la table de forwarding (FIB):
 - Les prefixes et longueur de préfixe sont uniques
 - Plus petite “Distance administrative”

Création de la table de Forwarding

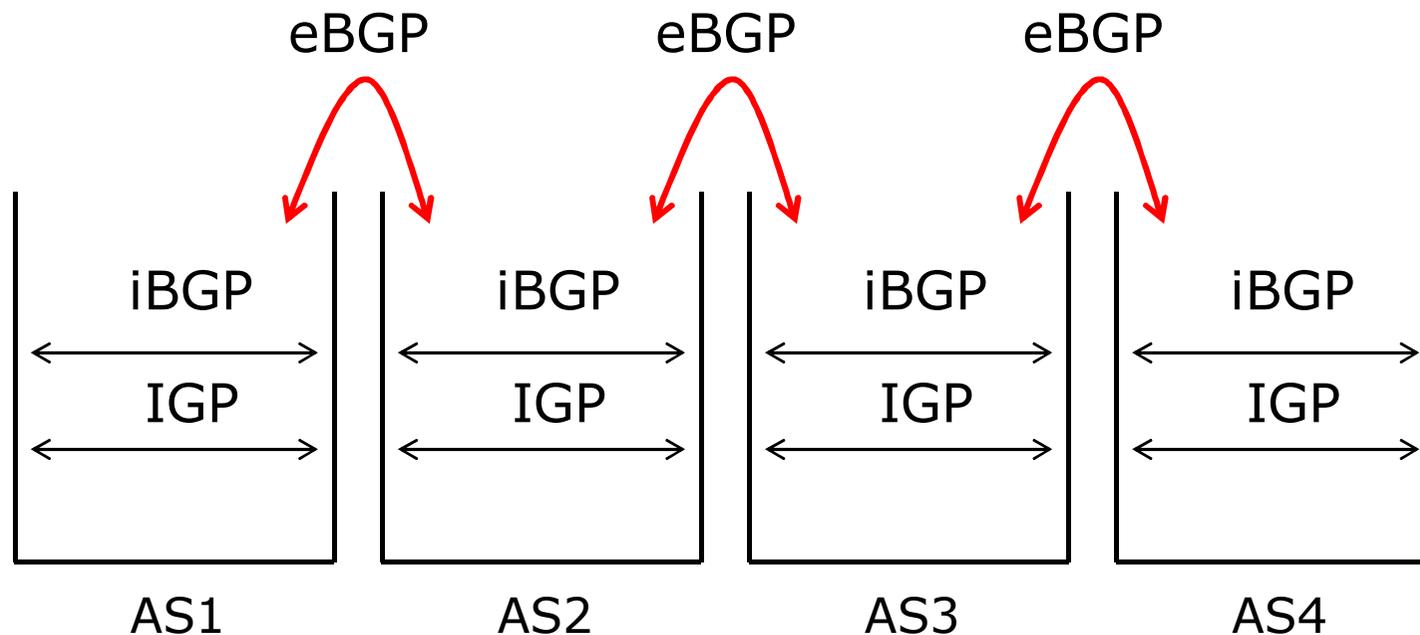


eBGP & iBGP

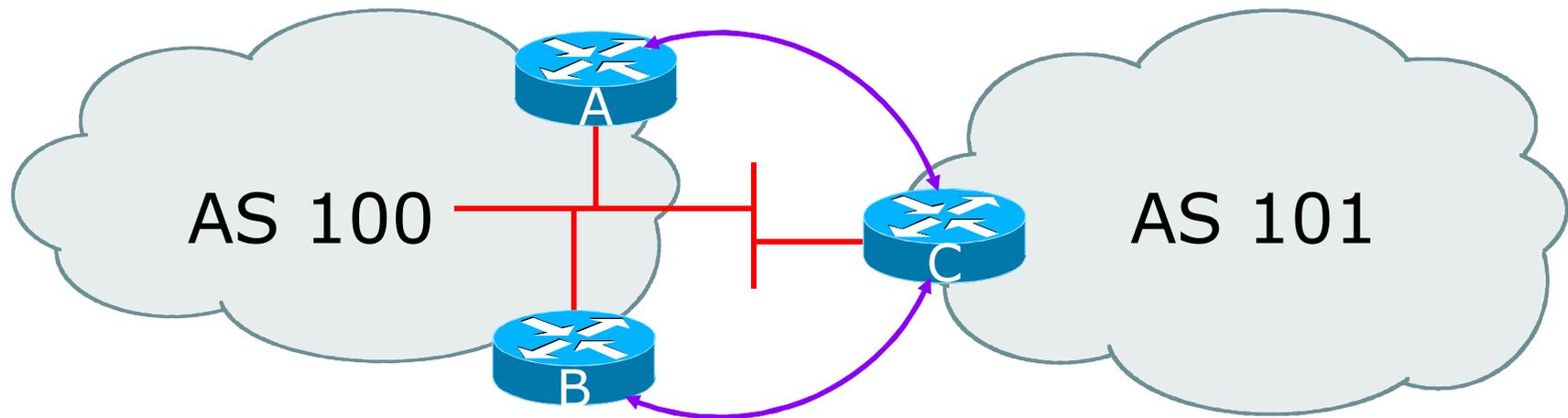
- BGP utilise en interne (iBGP) en externe (eBGP)
- iBGP pour le transport
 - Quelques/Tous les préfixes Internet à travers le backbone de l'ISP
 - Préfixe des clients des ISP
- eBGP utilisé
 - Echanger les préfixes avec les autres AS
 - Implemente la politique de routage

BGP/IGP modèle d'utilisée par les ISP

□ representation



BGP Externe Peering (eBGP)



- ❑ Entre des homologues BGP dans différents AS
- ❑ Les homologues doivent directement connectés
- ❑ **Ne Jamais** utiliser un IGP entre deux peers eBGP

Configurer BGP Externe

Router A dans AS100

```
interface ethernet 5/0
  ip address 102.102.10.2 255.255.255.240
!
router bgp 100
  network 100.100.8.0 mask 255.255.252.0
  neighbor 102.102.10.1 remote-as 101
  neighbor 102.102.10.1 prefix-list RouterC in
  neighbor 102.102.10.1 prefix-list RouterC out
!
```

ip address on
ethernet interface

Local ASN

Remote ASN

ip address of Router
C ethernet interface

Inbound and
outbound filters

Configurer BGP Externe

Router C in AS101

```
interface ethernet 1/0/0
  ip address 102.102.10.1 255.255.255.240
!
router bgp 101
  network 100.100.64.0 mask 255.255.248.0
  neighbor 102.102.10.2 remote-as 100
  neighbor 102.102.10.2 prefix-list RouterA in
  neighbor 102.102.10.2 prefix-list RouterA out
!
```

ip address on
ethernet interface

Local ASN

Remote ASN

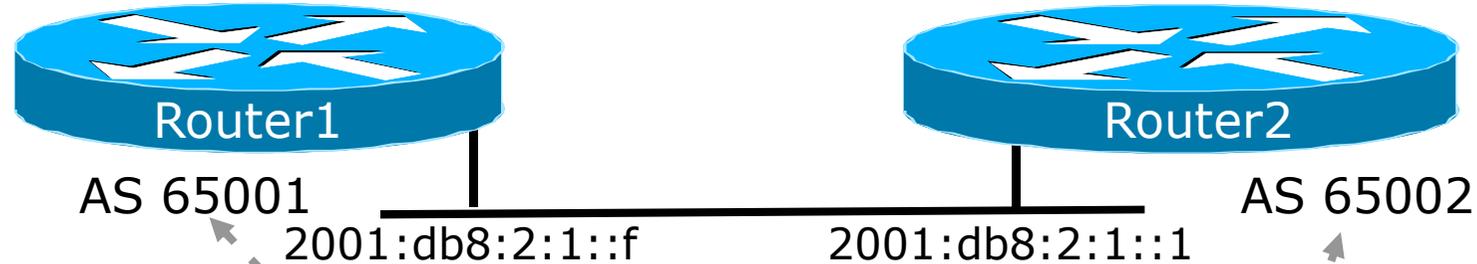
ip address of Router
A ethernet interface

Inbound and
outbound filters

Multi-Protocol BGP pour IPv6 – RFC2545

- Extensions spécifiques IPv6
 - Adresses “Scoped”: Next-hop contient une adresse globale IPv6 et/ou dans certain cas une adresse “link-local”
 - Le NEXT_HOP et NLRI sont définis comme les adresses et préfixe IPv6
 - Address Family Information (AFI) = 2 (IPv6)
 - Sub-AFI = 1 (NLRI est utilisé pour l’unicast)
 - Sub-AFI = 2 (NLRI est utilisé pour verifier l’adresse multicast)
 - Sub-AFI = 3 (NLRI est utilisé pour verifier l’adresse unicast et multicast)
 - Sub-AFI = 4 (label)

Session MP-BGP

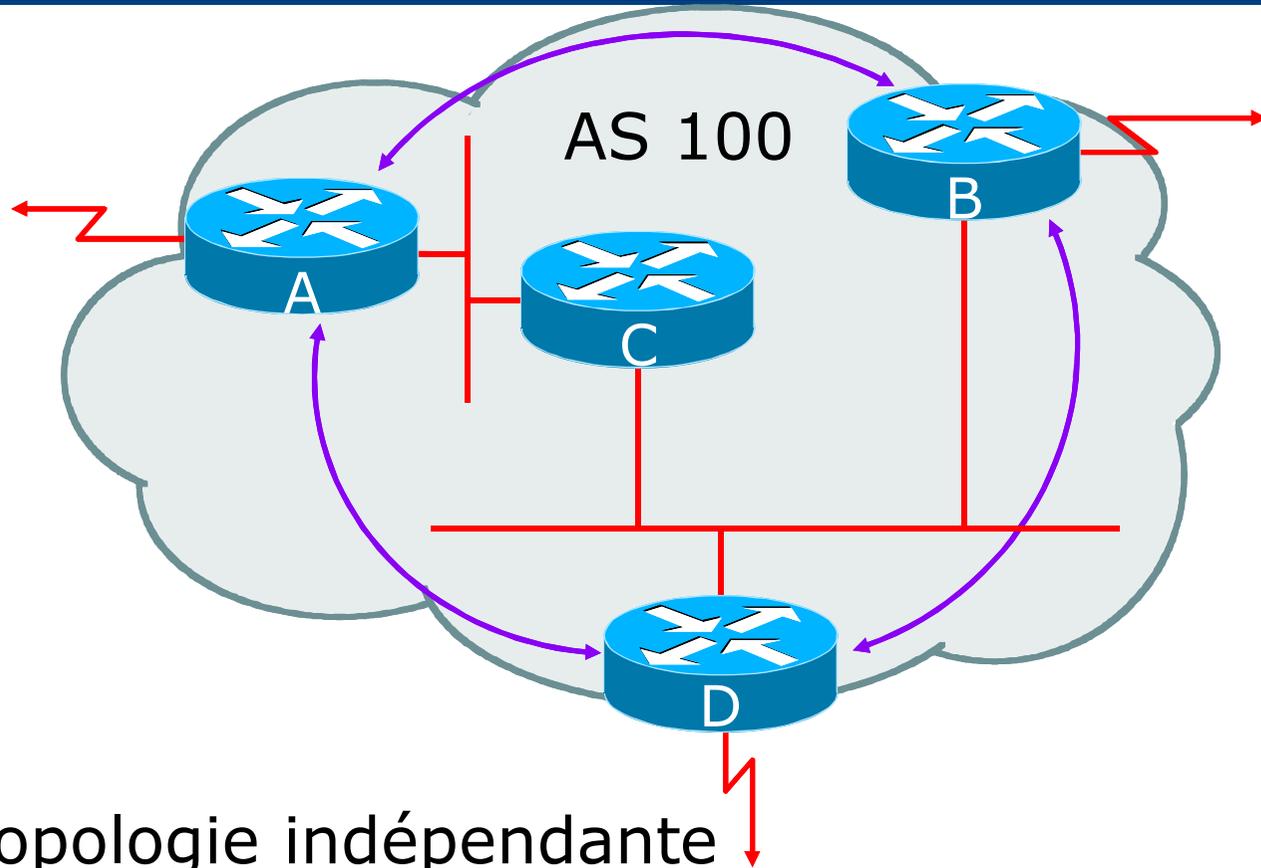


```
Router1#  
interface Ethernet0  
  ipv6 address 2001:db8:2:1::f/64  
!  
router bgp 65001  
  bgp router-id 10.10.10.1  
  no bgp default ipv4-unicast  
  neighbor 2001:db8:2:1::1 remote-as 65002  
  address-family ipv6  
    neighbor 2001:db8:2:1::1 activate  
    neighbor 2001:db8:2:1::1 prefix-list bgp65002in in  
    neighbor 2001:db8:2:1::1 prefix-list bgp65002out out  
  exit-address-family
```

Internal BGP (iBGP)

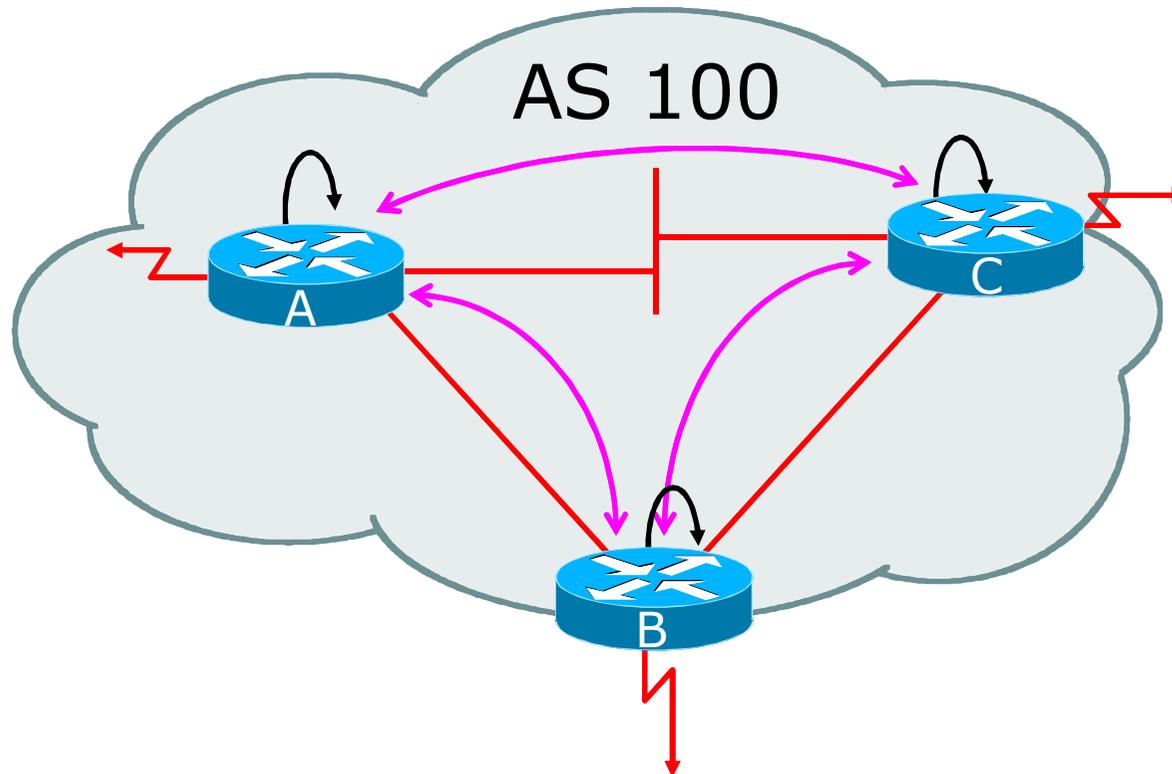
- ❑ Homologues BGP dans le même AS
- ❑ Ne requierent pas une connectivité directe
 - IGP utilisé pour faire communiquer les homologues BGP
- ❑ Homologues iBGP doivent être en “full meshed” :
 - Les routes apparaissent comme étant connectés
 - Diffuse les préfixes appris de l'extérieur d'autres AS
 - Ne diffuse pas les préfixes appris des autres routeurs homologues IBGP

BGP interne Peering (iBGP)



- ❑ Topologie indépendante
- ❑ Tous les homologues IBGP dans un AS doivent communiquer entre eux.

Interfaces Loopback et “peering”



- ❑ Communique avec l'interface loop-back
 - interface Loop-back est toujours UP – ever!
- ❑ Pour éviter que la session iBGP dépende de l'état d'une seule interface

Configuration de BGP interne

Router A dans AS100

```
interface loopback 0
  ip address 105.3.7.1 255.255.255.255
!
router bgp 100
  network 100.100.1.0
  neighbor 105.3.7.2 remote-as 100
  neighbor 105.3.7.2 update-source loopback0
  neighbor 105.3.7.3 remote-as 100
  neighbor 105.3.7.3 update-source loopback0
!
```

adresse ip sur
loopback interface

ASN Local

ASN Local

Adresse IP du Routeur B
loopback interface

Configuring Internal BGP

Router B dans AS100

```
interface loopback 0
  ip address 105.3.7.2 255.255.255.255
!
router bgp 100
  network 100.100.1.0
  neighbor 105.3.7.1 remote-as 100
  neighbor 105.3.7.1 update-source loopback0
  neighbor 105.3.7.3 remote-as 100
  neighbor 105.3.7.3 update-source loopback0
!
```

Adresse IP sur
l'interface loopback

ASN Local

ASN Local

ip address of Router
A loopback interface

Insertion de préfixes dans BGP

- Deux méthodes pour insérer un préfixe dans BGPs
 - `redistribute static`
 - `network command`

Inserting prefixes into BGP – redistribute static

- Exemple de configuration:

```
router bgp 100
```

```
  redistribute static
```

```
  ip route 102.10.32.0 255.255.254.0 serial0
```

- La route statique doit exister pour que la commande fonctionne
- Force la valeur “origin” à “incomplete”
- Utiliser avec prudence!

Insertion de prefixes dans BGP – redistribute static

- Utiliser avec beaucoup de prudence !
 - `redistribute <routing-protocol>` indique que toutes les routes du `<routing-protocol>` sera transfer
 - Pas evolutif Will not scale if uncontrolled
 - A eviter si possible
 - **redistribute** doit être utilisé conjointement avec des “route-maps” avec beaucoup de controle

Insertion des préfixes dans BGP – network command

- Exemple de configuration

```
router bgp 100
  network 102.10.32.0 mask 255.255.254.0
  ip route 102.10.32.0 255.255.254.0 serial0
```

- Une route correspondante dans la table de route est nécessaire pour que le préfixe soit annoncé par BGP.
- Force l'attribut origin à "IGP"

Configuration Aggregation

- Trois methodes pour configurer les aggregation de route :
 - `redistribute static`
 - `aggregate-address`
 - `network command`

Configuration d'Aggregation

□ Exemple de configuration :

```
router bgp 100
  redistribute static
  ip route 102.10.0.0 255.255.0.0 null0 250
```

□ Route statique vers “null0” is called a pull up route

- Packets only sent here if there is no more specific match in the routing table
- Distance of 250 ensures this is last resort static
- Care required – see previously!

Configuration Aggregation– Commande “Network”

- Exemple de configuration

```
router bgp 100
  network 102.10.0.0 mask 255.255.0.0
  ip route 102.10.0.0 255.255.0.0 null0 250
```

- Une route correspondante doit exister dans la table de routage pour que le réseau soit annoncé
- Methode recommandée pour générer des route agrégée

Configuration Aggregation– Commande “aggregate-address”

- Exemple de configuration:

```
router bgp 100
```

```
network 102.10.32.0 mask 255.255.252.0
```

```
aggregate-address 102.10.0.0 255.255.0.0 [summary-only]
```

- L’existence d’un préfixe plus spécifique dans la table BGP est nécessaire pour que la route agrégée soit annoncée
- L’option “summary-only”
 - Assure que seule la route agrégée est annoncée si un préfixe plus spécifique existe dans la table de routage

Resumé

BGP neighbour status

```
Router6>sh ip bgp sum
```

```
BGP router identifier 10.0.15.246, local AS number 10
```

```
BGP table version is 16, main routing table version 16
```

```
7 network entries using 819 bytes of memory
```

```
14 path entries using 728 bytes of memory
```

```
2/1 BGP path/bestpath attribute entries using 248 bytes of memory
```

```
0 BGP route-map cache entries using 0 bytes of memory
```

```
0 BGP filter-list cache entries using 0 bytes of memory
```

```
BGP using 1795 total bytes of memory
```

```
BGP activity 7/0 prefixes, 14/0 paths, scan interval 60 secs
```

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

BGP Version

Updates sent
and received

Updates waiting

Resumé

Table BGP

```
Router6>sh ip bgp
```

```
BGP table version is 16, local router ID is 10.0.15.246
```

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

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

```
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i 10.0.0.0/26	10.0.15.241	0	100	0	i
*>i 10.0.0.64/26	10.0.15.242	0	100	0	i
*>i 10.0.0.128/26	10.0.15.243	0	100	0	i
*>i 10.0.0.192/26	10.0.15.244	0	100	0	i
*>i 10.0.1.0/26	10.0.15.245	0	100	0	i
*> 10.0.1.64/26	0.0.0.0	0		32768	i
*>i 10.0.1.128/26	10.0.15.247	0	100	0	i
*>i 10.0.1.192/26	10.0.15.248	0	100	0	i
*>i 10.0.2.0/26	10.0.15.249	0	100	0	i
*>i 10.0.2.64/26	10.0.15.250	0	100	0	i

```
...
```

Resumé

- BGP4 – path vector protocol
- iBGP versus eBGP
- Design iBGP stable – peer with loopbacks
- Announce des prefixes & aggregates

Introduction à BGP



AFNOG 2014