



GRAND DUCHY OF LUXEMBOURG
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa
Infrastructure Trust Fund

Introduction à BGP



Border Gateway Protocol

- Protocole de routage utilisé pour échanger des informations de routage entre les différents réseaux
 - Exterior Gateway Protocol
- Décrit dans la RFC 4271
 - RFC4276 donne un rapport d'exécution sur BGP
 - RFC4277 décrit les expériences opérationnelles d'utilisation de BGP
- Le système autonome (AS) est la pierre angulaire de BGP
 - Il est utilisé pour identifier de façon unique les réseaux ayant une politique de routage commune (généralement 1 par entreprise)

BGP

- Path Vector Protocol
- Mises à jour incrémentales
- Beaucoup d'options pour l'application des stratégies
- Classless Inter Domain Routing (CIDR)
- Largement utilisé pour le backbone Internet
- Systèmes autonomes

Path Vector Protocol

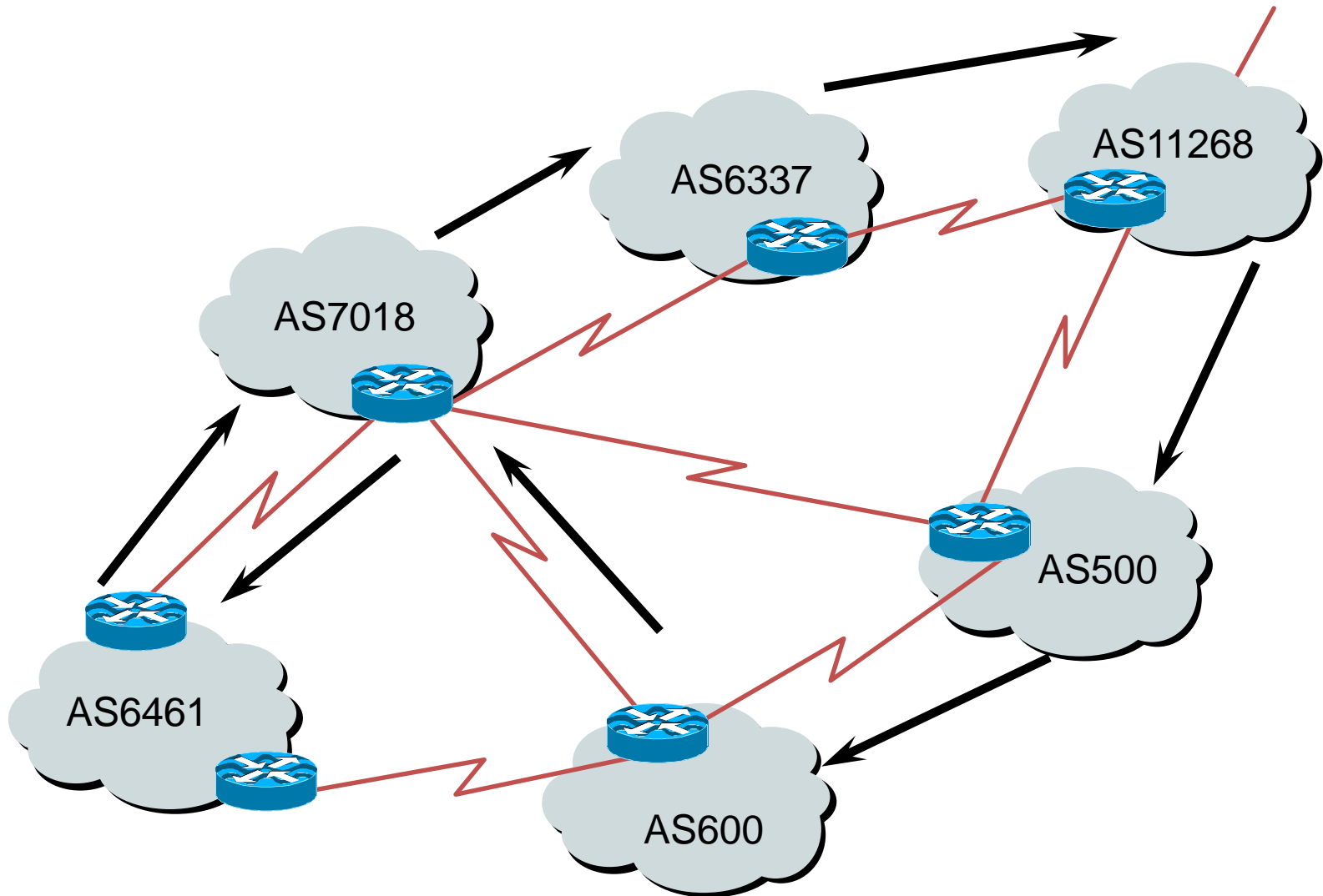
- BGP est classé comme un protocole de routage *path vector* (voir RFC 1322)
 - Un protocole path vector définit un trajet a partir d'une destination et les caractéristiques du PATH qui mène à cette destination.

12.6.126.0/24 207.126.96.43 1021 0 6461 7018 6337 11268 i



AS Path

Path Vector Protocol



Définitions

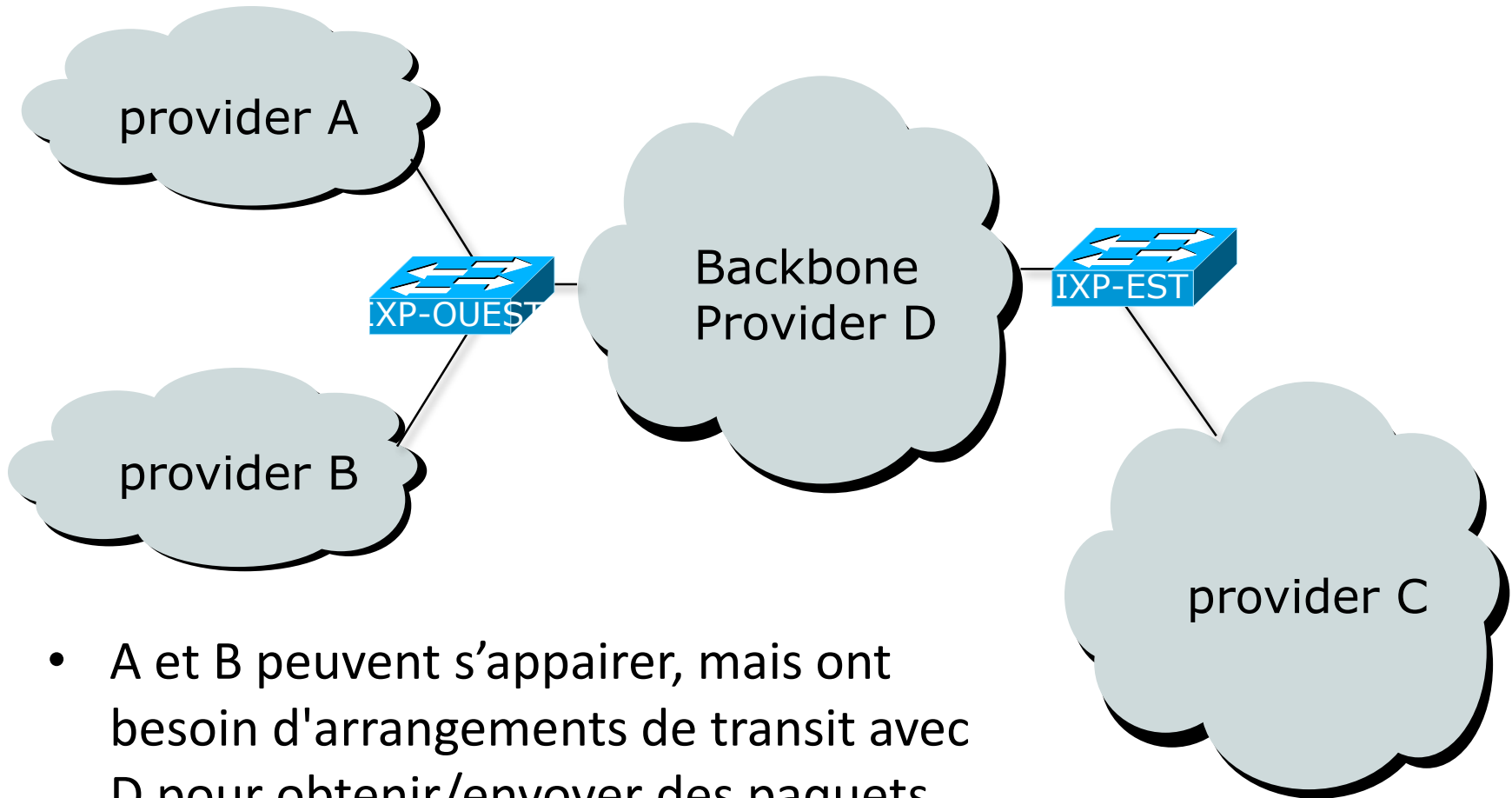
- **Transit** – acheminement de trafic sur un réseau, généralement moyennant un paiement
- **Peering** – échanger des informations de routage et de trafic gratuitement
- **Default** – où envoyer le trafic quand il n'y a pas de correspondance explicite dans la table de routage

Default Free Zone

La Default Free Zone est constituée de routeurs Internet qui ont des informations de routage explicites sur le reste de l'Internet et n'ont donc pas besoin d'utiliser une route par défaut

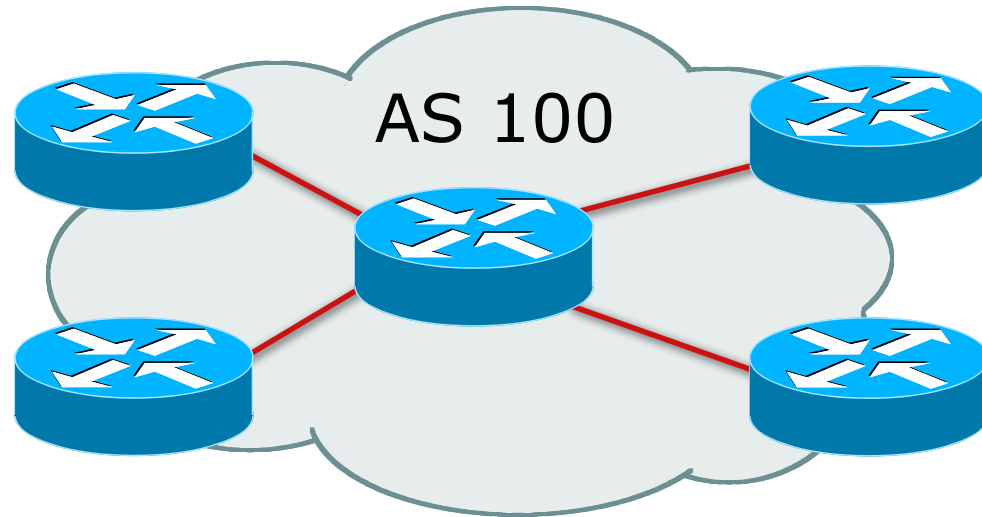
NB: n'est pas liée à l'emplacement d'un ISP dans la hiérarchie

Exemple de Peering et de Transit



- A et B peuvent s'appairer, mais ont besoin d'arrangements de transit avec D pour obtenir/envoyer des paquets de/vers C

Systeme autonome (AS)



- Regroupement de réseaux avec la même politique de routage
- Protocole de routage unique
- Habituellement en propriété, confiance et contrôle administratif unique
- Identifié par un entier 32 bits unique (ASN)

Numéro de système autonome (ASN)

- Deux plages
 - 0-65535 (original 16-bit range)
 - 65536-4294967295 (32-bit range – RFC4893)
- Usage:
 - 0 and 65535 (réservé)
 - 1-64495 (Internet public)
 - 64496-64511 (documentation – RFC5398)
 - 64512-65534 (usage privé uniquement)
 - 23456 (représente un AS32 bits si non géré)
 - 65536-65551 (documentation – RFC5398)
 - 65552-4294967295 (Internet public)
- Représentation de plage 32 bits spécifiée dans RFC5396
 - Définit “asplain” (format traditionnel) comme notation standard

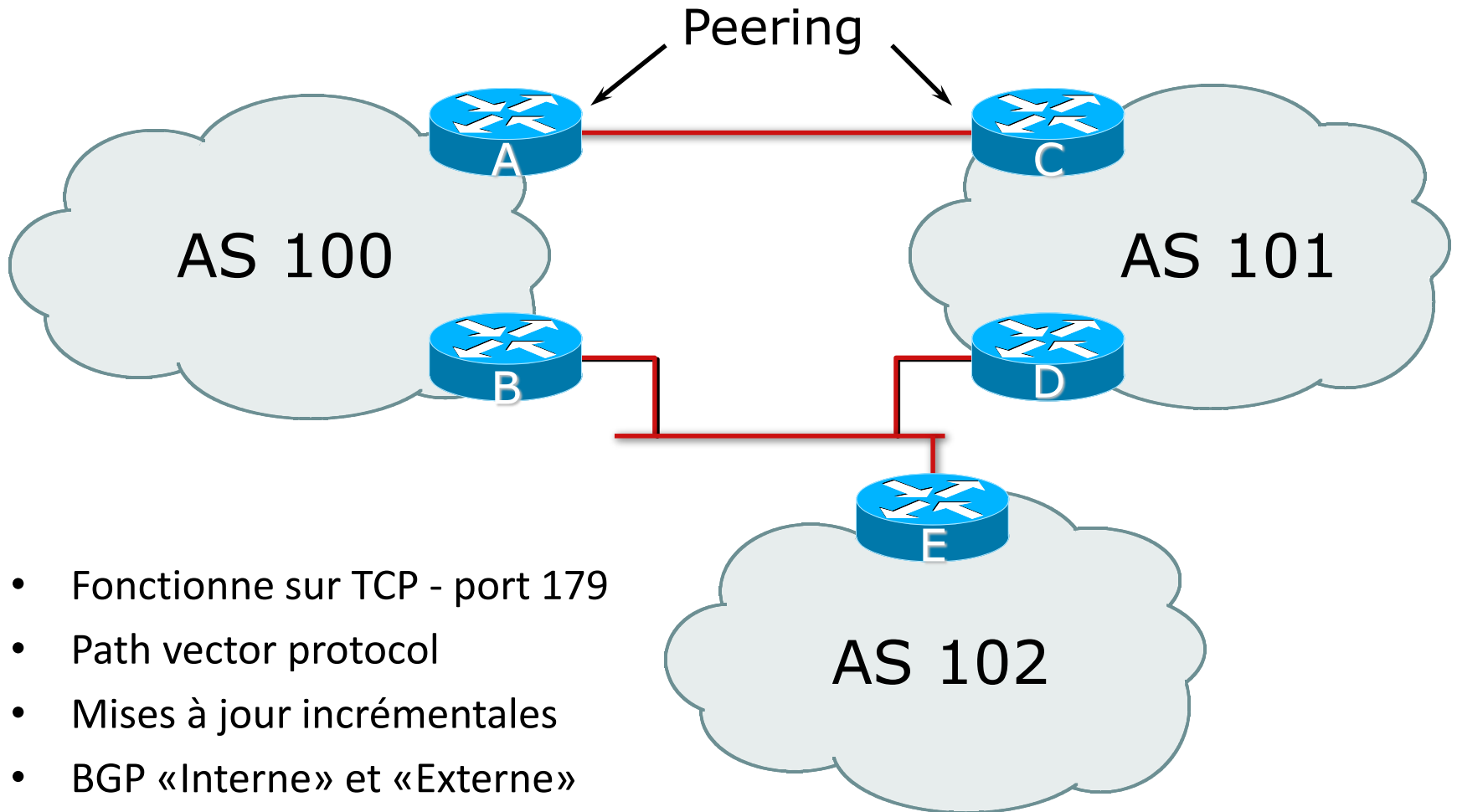
Numéro de système autonome (ASN)

- Les ASN sont distribués par les Registres Internet Régionaux
 - Ils sont également disponibles auprès des ISP en amont qui sont membres de l'un des RIR
- Les allocations ASN actuelles de 16-bit jusqu'à 61439 ont été faites pour les RIR
 - Environ 41200 sont visibles sur Internet
- Chaque RIR a également reçu un bloc d'ASN 32-bit
 - Sur les 2800 affectations, autour de 2400 sont visibles sur Internet
- Cf www.iana.org/assignments/as-numbers

Configurer BGP dans Cisco IOS

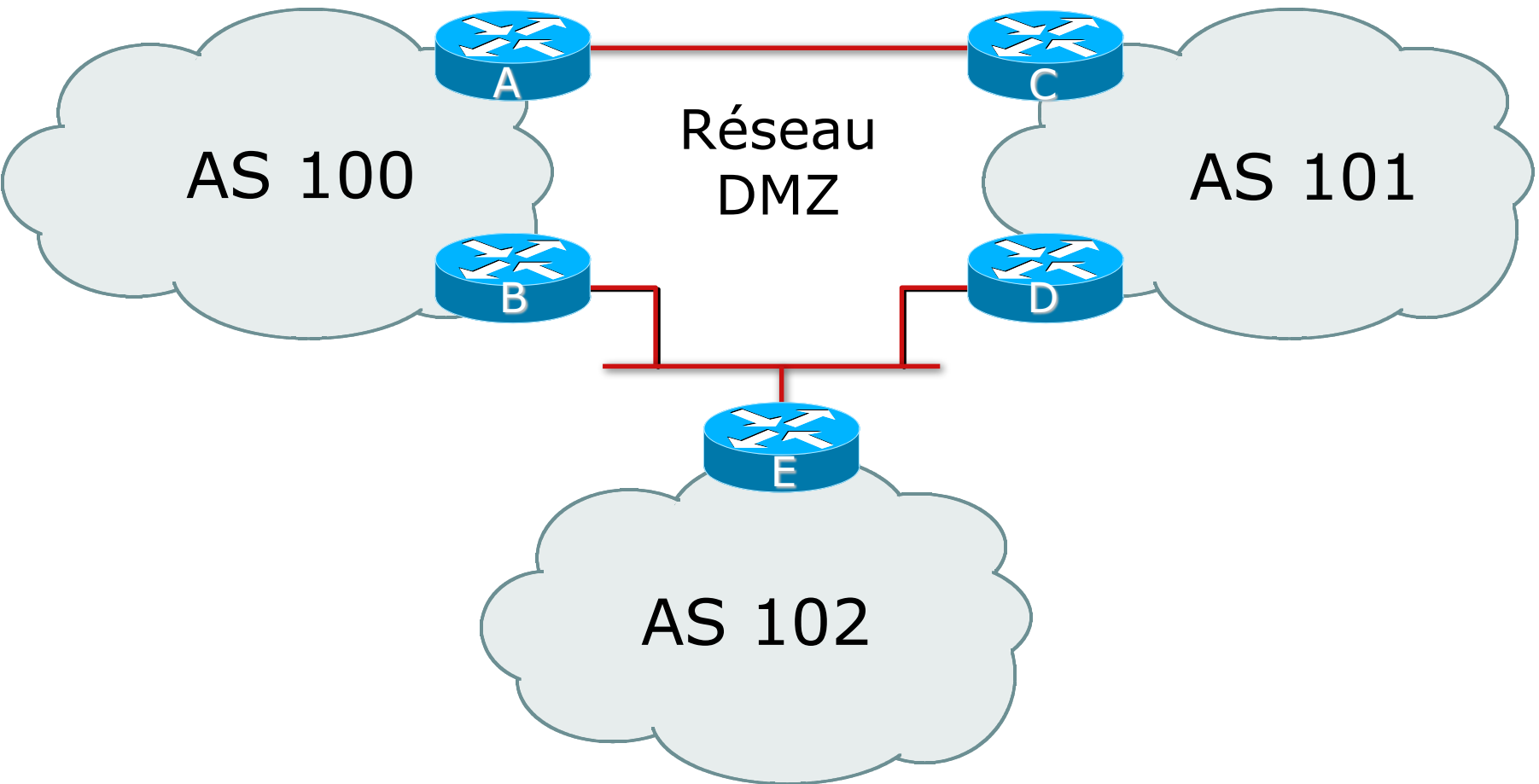
- Cette commande active BGP dans Cisco IOS:
`router bgp 100`
- Pour les ASN > 65535, le numéro d'AS peut être entré en notation "plain" ou "dot " :
`router bgp 131076`
ou
`router bgp 2.4`
- IOS affiche les ASN en notation "plain" par défaut
 - La notation "Dot" est facultative:
`router bgp 2.4`
`bgp asnotation dot`

Notions de base BGP



- Fonctionne sur TCP - port 179
- Path vector protocol
- Mises à jour incrémentales
- BGP «Interne» et «Externe»

Zone de démarcation (DMZ)



- DMZ est le lien ou réseau partagé entre les AS

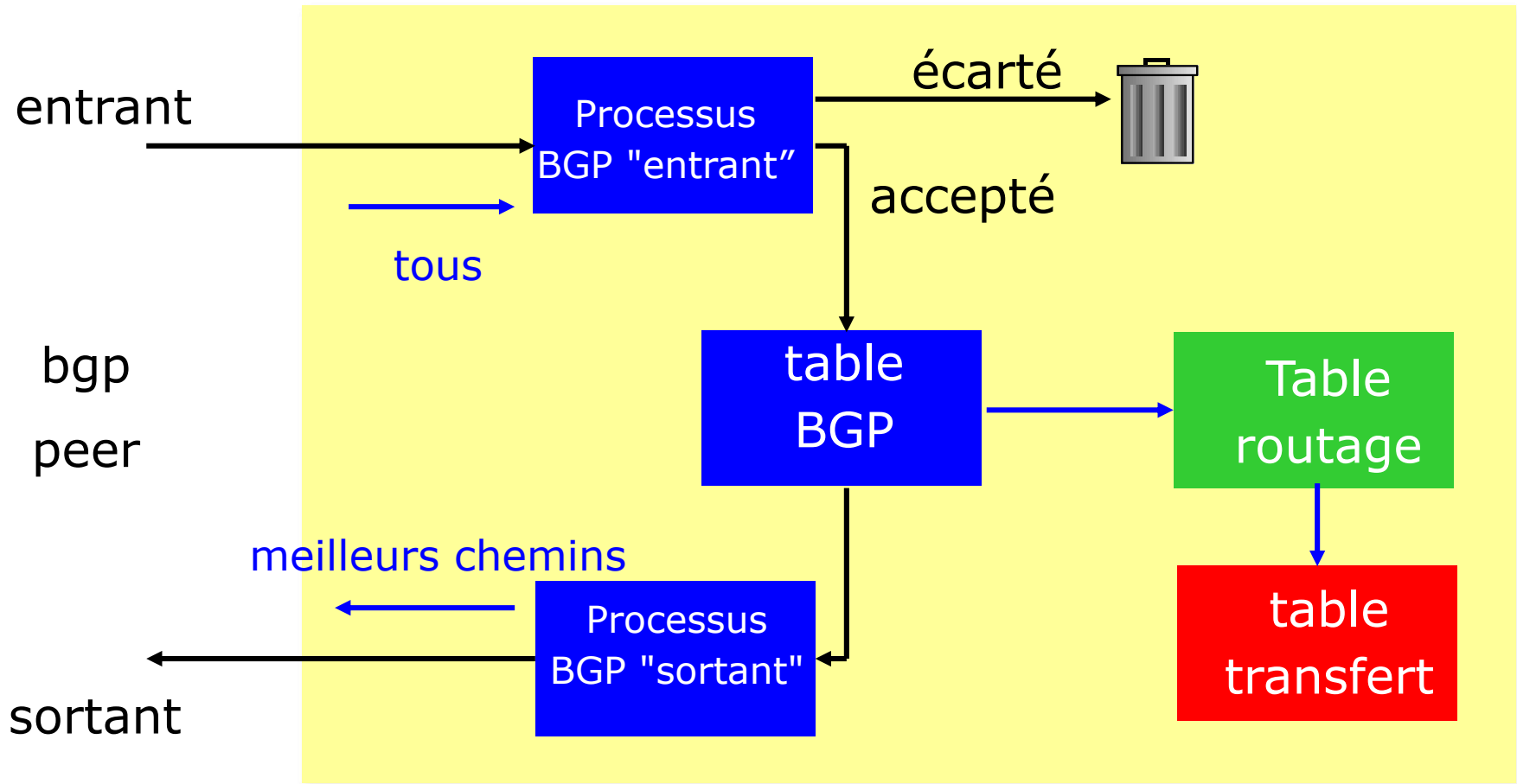
Opération générale BGP

- Apprend de multiples chemins par l'intermédiaire de speakers BGP internes et externes
- Choisit le meilleur chemin et l'installe dans la table de routage (RIB)
- Le meilleur chemin est envoyé aux voisins BGP externes
- Les stratégies sont appliquées en influant sur le choix du meilleur chemin

Construire la table de Transfert (Forwarding Table)

- Processus BGP “entrant”
 - reçoit les informations de chemin des pairs
 - les résultats de la sélection de chemin BGP placés dans la table BGP
 - “Meilleur chemin” marqué
- Processus BGP “sortant”
 - annonce le “meilleur chemin” aux pairs
- Meilleur chemin stocké dans la table de routage (RIB)
- Les meilleurs chemins dans le RIB sont ajoutés a la forwarding table (FIB) si:
 - préfixe et longueur de préfixe sont uniques
 - “distance administrative” la plus faible

Construire la table de Transfert (Forwarding Table)

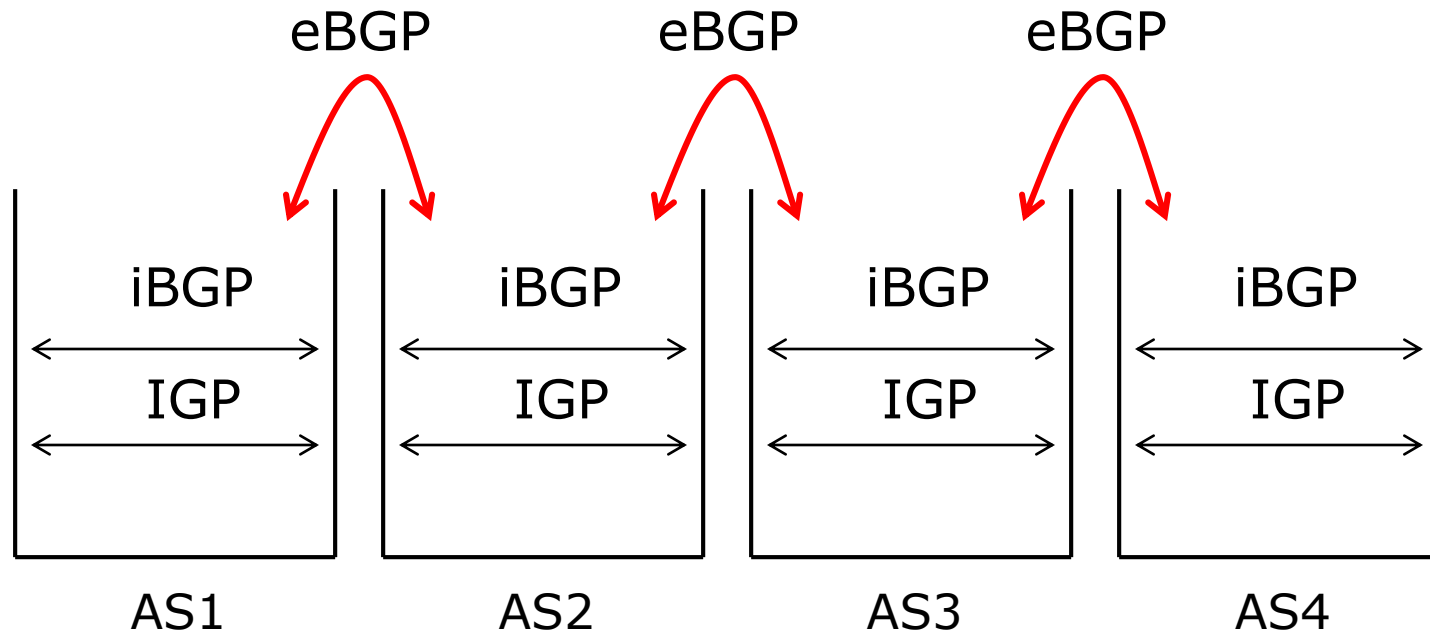


eBGP & iBGP

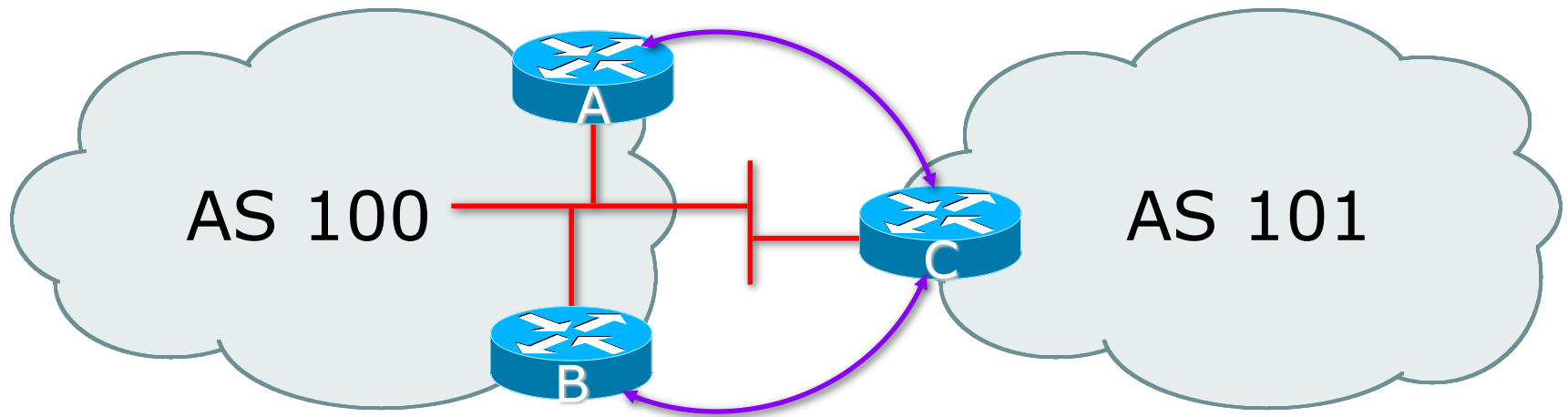
- BGP utilisé en interne (iBGP) et externe (eBGP)
- iBGP utilisé pour transporter
 - Certains / tous les préfixes Internet à travers le backbone ISP
 - les préfixes des clients ISP
- eBGP utilisé pour
 - Echanger des préfixes avec d'autres AS
 - Mettre en œuvre la politique de routage

Modèle BGP/IGP utilisé dans les réseaux ISP

- Représentation du modèle



External BGP Peering (eBGP)



- Entre speakers BGP de différents AS
- Doivent être directement connectés
- **Ne jamais** exécuter un IGP entre pairs eBGP

Configurer un BGP externe

Routeur 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
!
```

adresse IP sur interface Ethernet

ASN local

ASN distant

adresse ip de l'interface Ethernet du routeur C

Filtres entrants et sortants

Configurer un BGP externe

Routeur C dans 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
!
```

adresse IP sur interface Ethernet

ASN locale

ASN distant

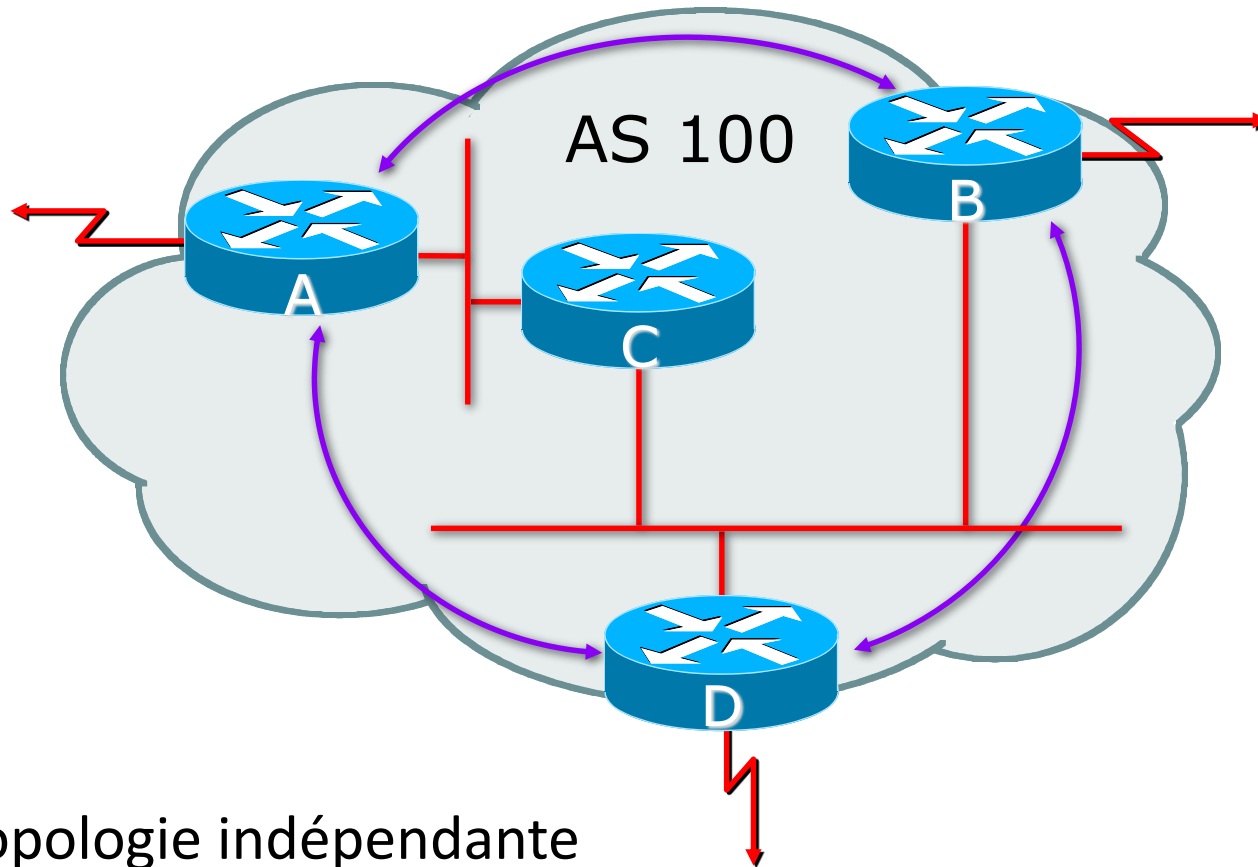
adresse ip de l'interface Ethernet du routeur A

Filtres entrants et sortants

BGP interne (iBGP)

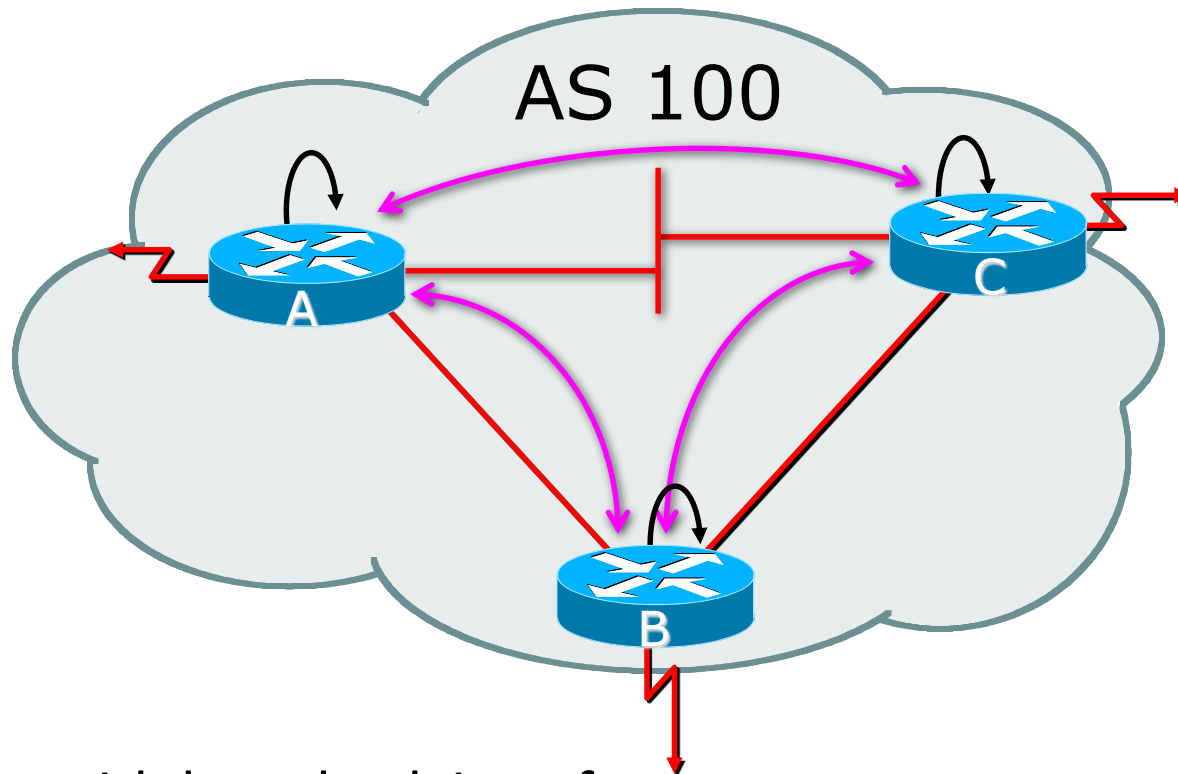
- BGP peer dans le même AS
- Pas requis d'être directement connecté
 - IGP prend soin de la connectivité des speakers inter-BGP
- Les speakers iBGP doivent être entièrement maillés:
 - Ils génèrent des réseaux connectés
 - Ils transmettent des préfixes appris de l'extérieur de l'ASN
 - Ils ne transmettent pas des préfixes appris d'autres voisins iBGP

Peering BGP interne (iBGP)



- Topologie indépendante
- Chaque speaker iBGP doit peering avec chacun des autres routeurs iBGP dans l'AS

Peering entre interfaces de Loopback



- Peer with loop-back interface
 - L'interface de loop-back ne tombe pas - jamais!
- Ne laissez pas une session iBGP dépendre de l'état d'une interface unique ou de la topologie physique

Configurer un BGP interne

Routeur 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 interface loopback

ASN local

ASN local

adresse ip de l'interface loopback du routeur B

Configurer un BGP interne

Routeur 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 interface loopback

ASN local

ASN local

adresse ip de l'interface loopback du routeur A

Insérer des préfixes dans BGP

- Deux façons d'insérer des préfixes dans BGP

- Commande

- redistribute static**

- Commande

- Network x.x.x.x**

Insérer des préfixes dans BGP – redistribute static

- Exemple de configuration:

```
router bgp 100
  redistribute static
  ip route 102.10.32.0 255.255.254.0 serial10
```

- Une route statique doit exister avant que la commande « redistribute » fonctionne
- Force l'origine à être “incomplète”
- Soins requis!

Insérer des préfixes dans BGP – redistribute static

- Soins requis avec redistribute!
 - `redistribute<routing-protocol>` signifie que tout ce qui est dans le `<routing-protocol>` sera transféré dans BGP
 - Ne s'ajustera pas s'il n'est pas contrôlé
 - Préférable d'éviter autant que possible
 - **redistribute** normalement utilisé avec des “route-maps” et sous contrôle administratif serré

Insérer des préfixes dans BGP – commande network

- 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
```

- Un itinéraire correspondant doit exister dans la table de routage avant que le réseau soit annoncé
- Force l'origine à être “IGP”

Configurer une agrégation

- Trois façons de configurer l'agrégation
 - `redistribute static`
 - `adresse-agrégats`
 - `network`

Configurer une agrégation

- Exemple de configuration:

```
router bgp 100
```

```
redistribute static
```

```
ip route 102.10.0.0 255.255.0.0 null0 250
```

- une route statique vers “null0” est appelée une route « pull up »
 - Paquets envoyés ici uniquement s'il n'y a plus de correspondance spécifique dans la table de routage
 - distance de 250 assure que ceci est le dernier recours statique
 - soins requis - voir plus haut!

Configurer une agrégation – 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
```

- Un itinéraire correspondant doit exister dans la table de routage avant que le réseau soit annoncé
- Moyen le plus simple et le meilleur de générer un agrégat

Configurer une agrégation – commande aggregate-address

- Exemple de configuration:

```
routeur bgp 100
```

```
network 102.10.32.0 mask 255.255.252.0
```

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

- Nécessite un préfixe plus spécifique dans le tableau BGP avant que l'agrégat soit annoncé
- Mot-clé uniquement de résumé (summary-only keyword)
 - Mot-clé facultatif qui garantit que seul le résumé est annoncé si un préfixe plus spécifique existe dans la table de routage

Résumé

BGP neighbour status

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

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

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

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

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

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

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

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

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

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

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

Version BGP

Mises à jour
envoyées et
reçues

Mises à jour en attente

Résumé

Table BGP

```
Router6>sh ip bgp
```

```
BGP table version is 30, local router ID is 10.0.15.246
```

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

```
          r RIB-failure, S Stale
```

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

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

```
...
```

Résumé

- BGP4 – protocole path vector
- iBGP / eBGP
- iBGP stable – sessions avec loopbacks
- annoncer préfixes et agrégats

Reconnaissance et attribution

Cette présentation contient des informations initialement développées et maintenues par les organisations et individu suivants et prévues pour le projet AXIS de l'Union africaine

Cisco ISP/IXP Workshops

Philip Smith: - pfsinoz@gmail.com

 **APNIC** www.apnic.net



GRAND DUCHY OF LUXEMBOURG
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa
Infrastructure Trust Fund

Introduction à BGP

Fin

