



GRAND DUCHY OF LUXEMBOURG
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa
Infrastructure Trust Fund

Attributs BGP et sélections de chemin



Attributs BGP

Les "outils" disponibles pour la tâche

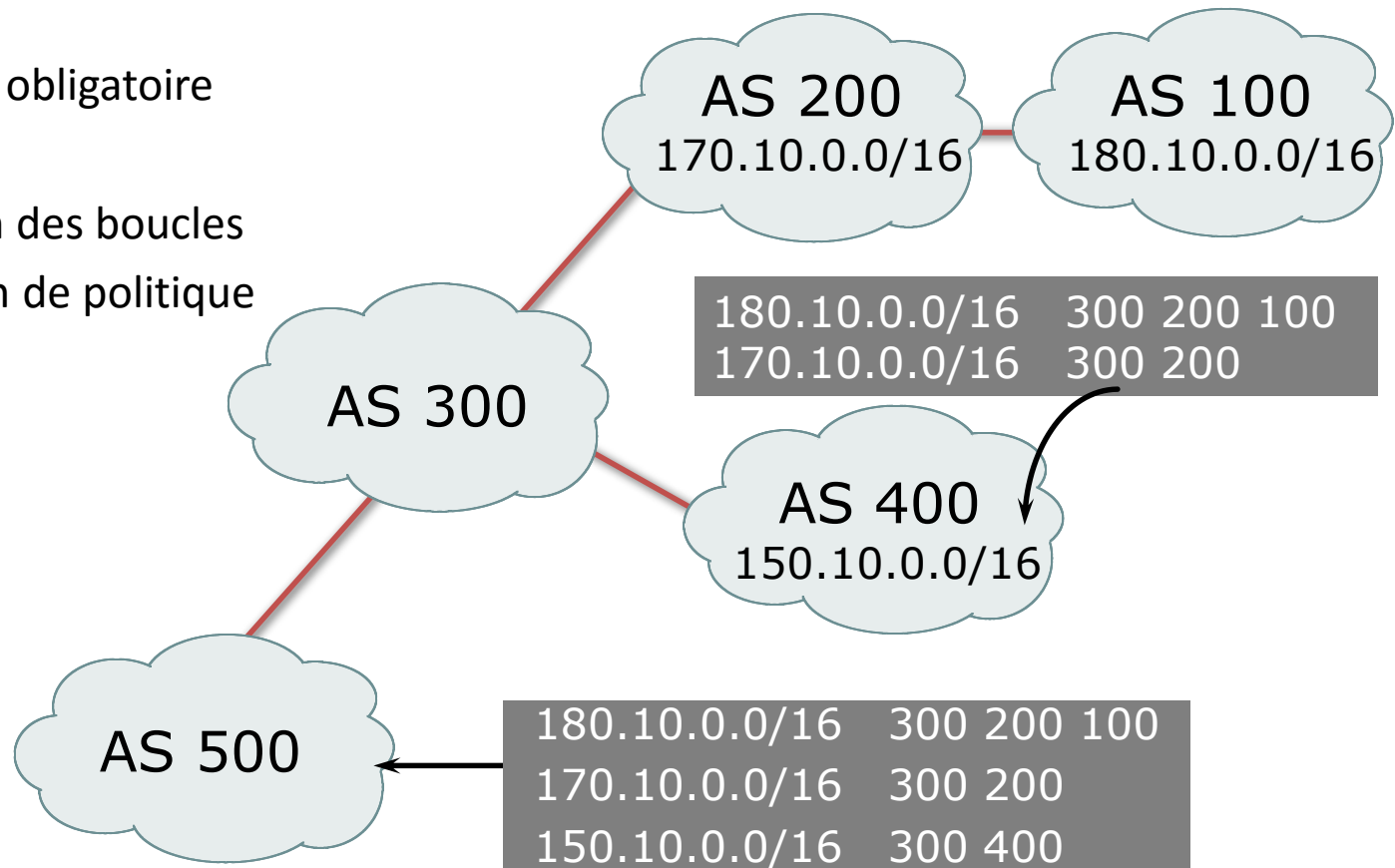
Qu'est-ce qu'un attribut?

...	Next Hop	AS Path	MED	...	...
-----	----------	---------	-----	-----	-----

- Partie d'une mise à jour BGP
- Décrit les caractéristiques de préfixe
- Peut être soit transitif ou non transitif
- Certains sont obligatoires

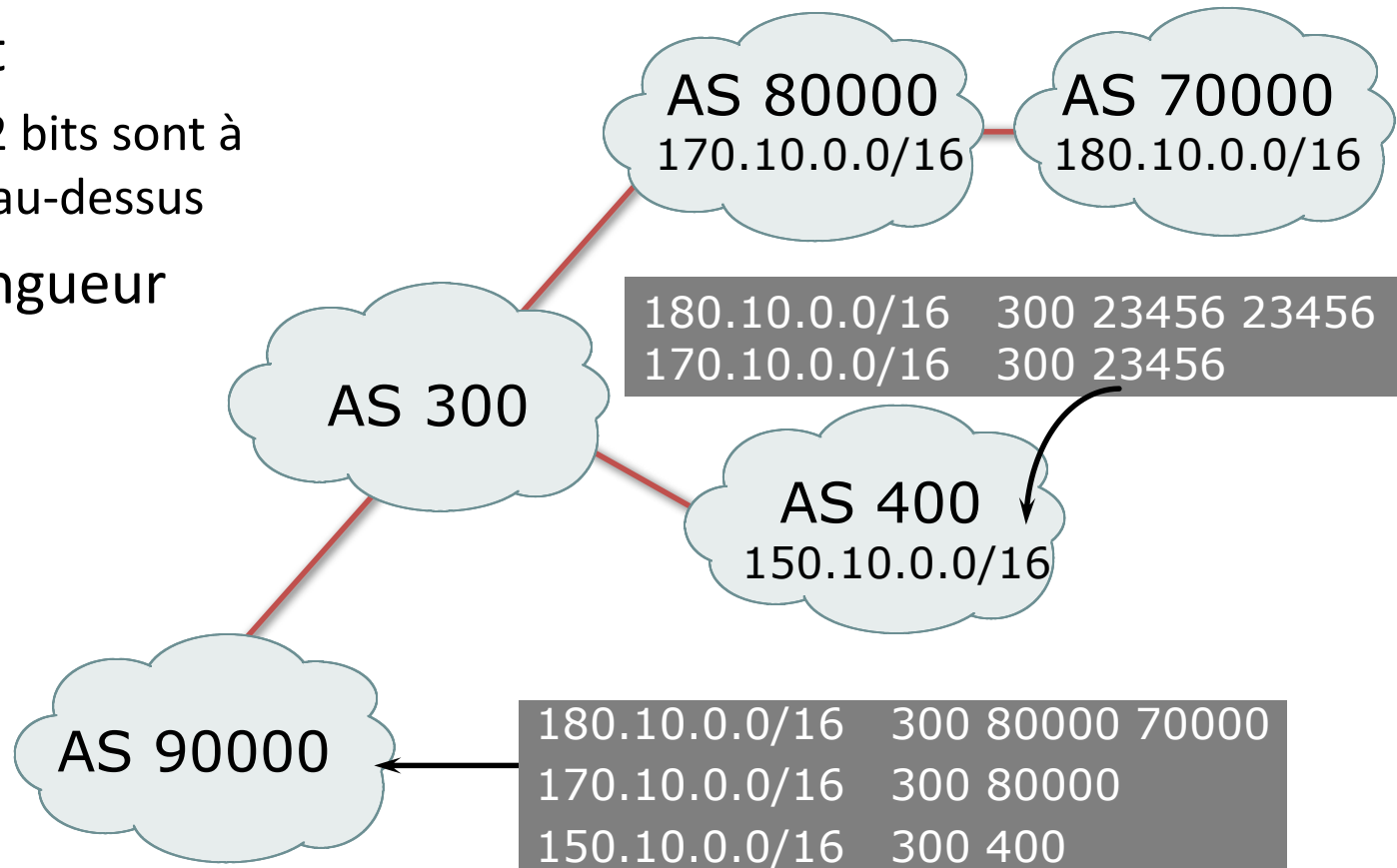
AS-Path

- Séquence d'AS qu'une route a traversé
- Attribut transitif obligatoire
- Utilisé pour:
 - La détection des boucles
 - L'application de politique

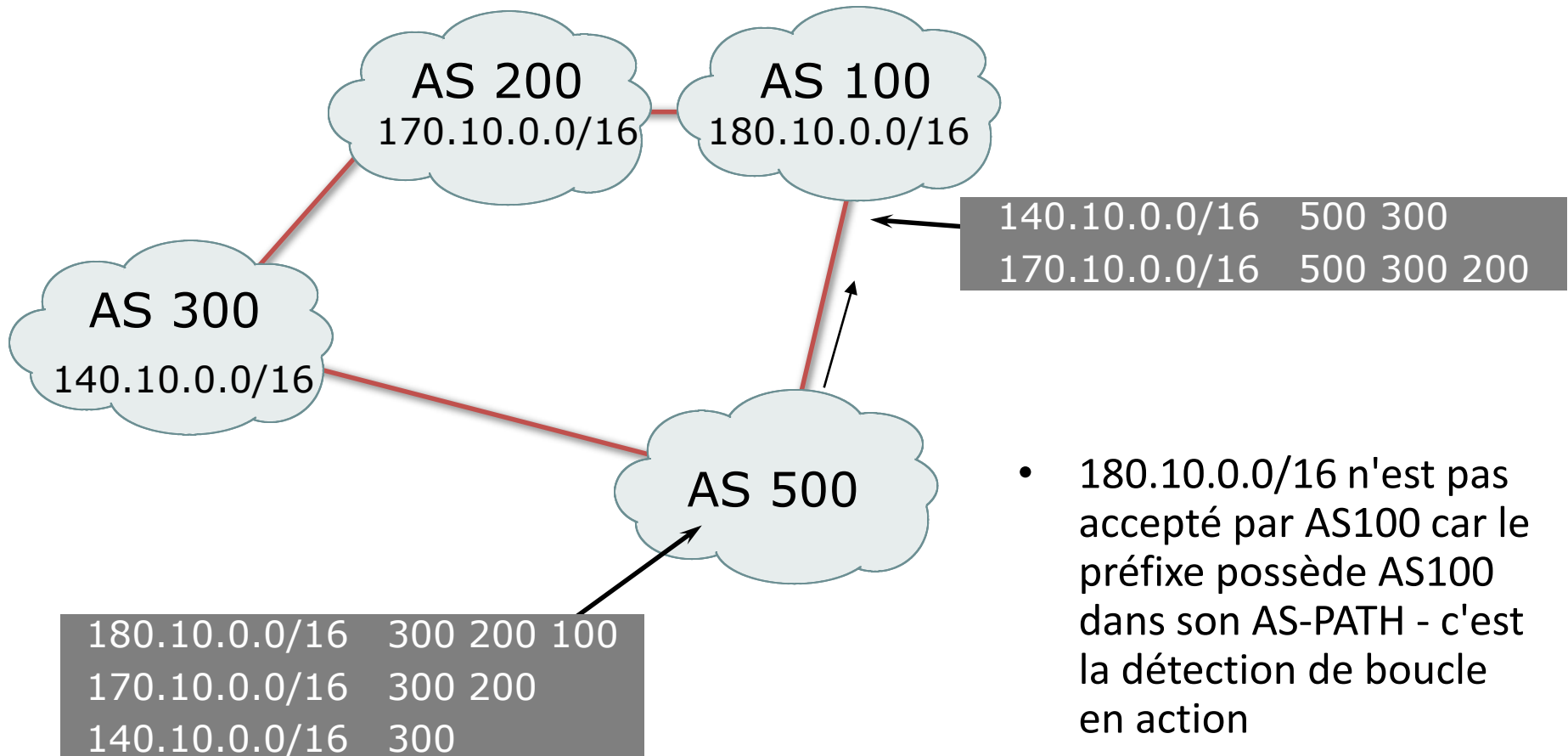


AS-Path (avec des ASN de 16 et 32-bit)

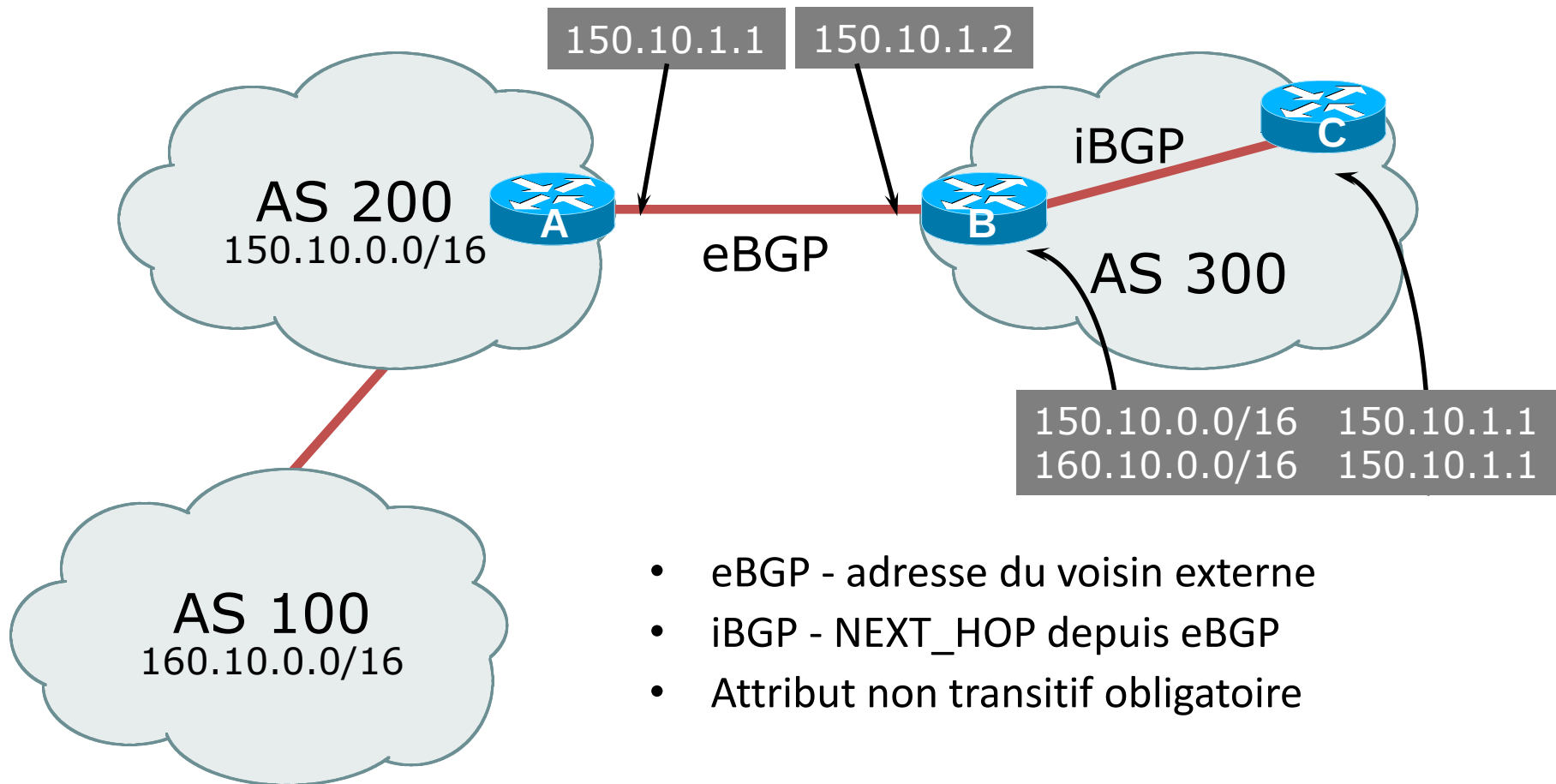
- Internet avec des ASN 16-bit et 32-bit
 - ASN de 32 bits sont à 65536 et au-dessus
- AS-PATH longueur maintenue



Détection de boucle AS-Path

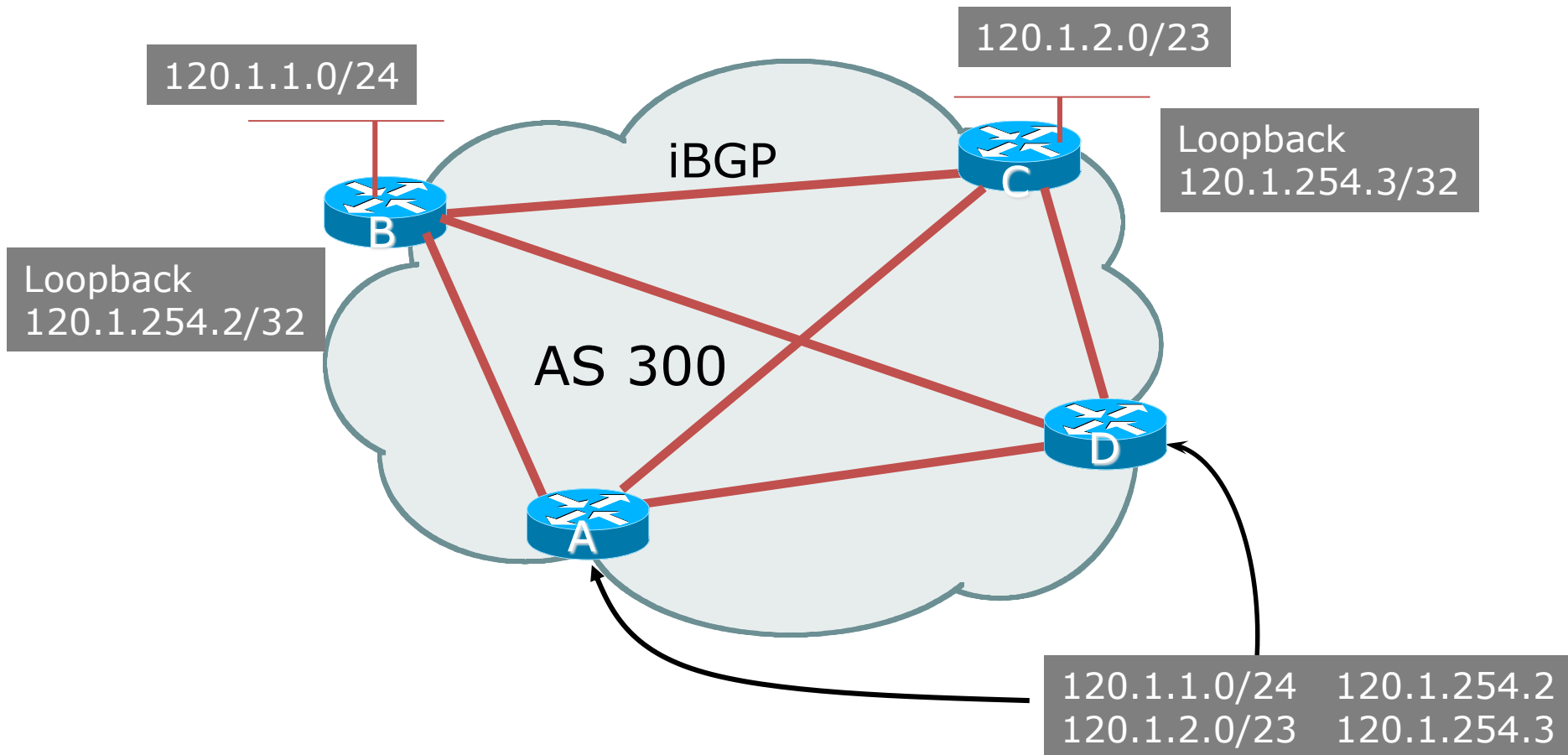


Next Hop



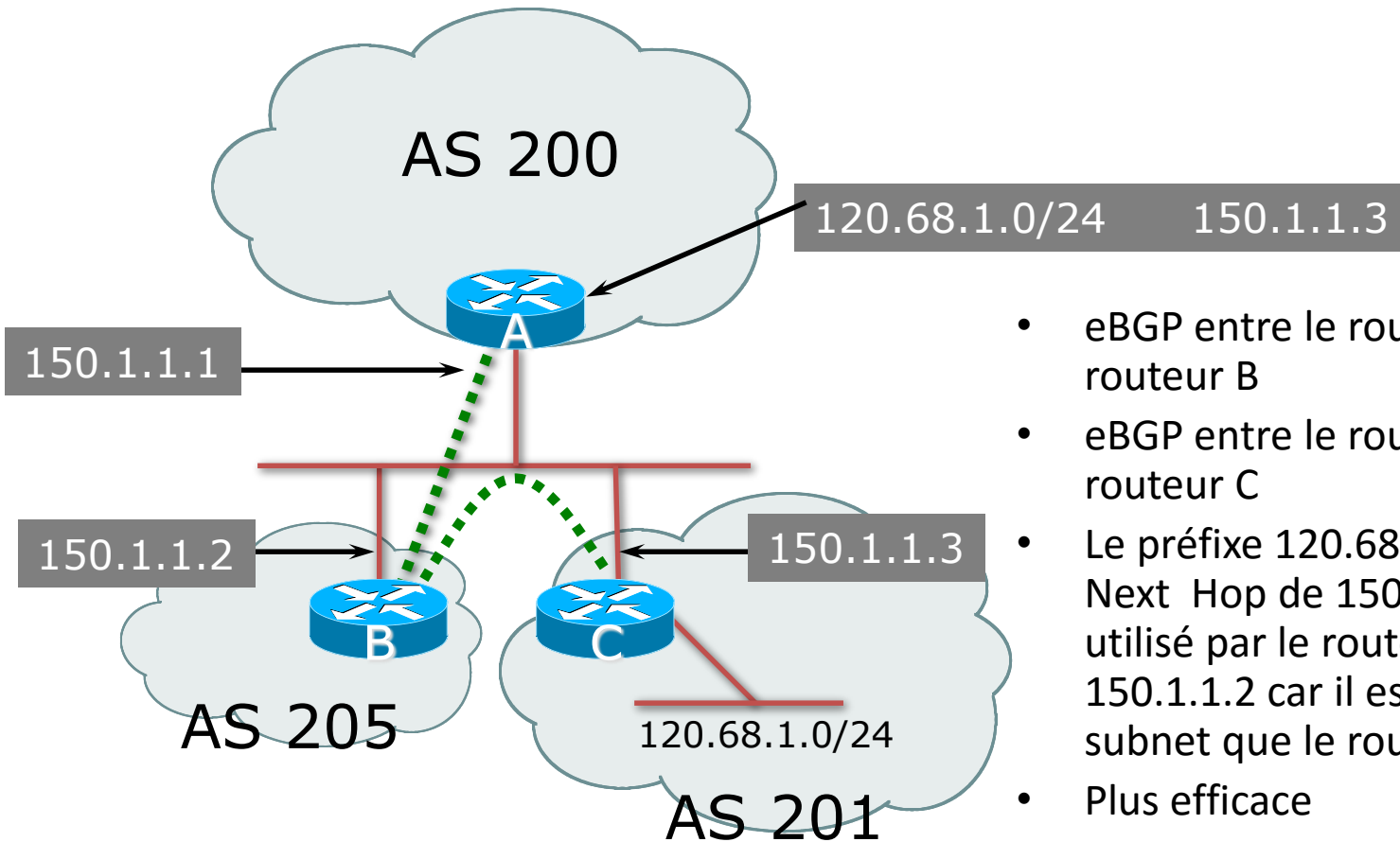
- eBGP - adresse du voisin externe
- iBGP - NEXT_HOP depuis eBGP
- Attribut non transitif obligatoire

iBGP Next Hop



- Next hop est l'adresse loopback du routeur iBGP
- Route look-up récursif

Next Hop tiers



- eBGP entre le routeur A et le routeur B
- eBGP entre le routeur B et le routeur C
- Le préfixe 120.68.1/24 a l'adresse Next Hop de 150.1.1.3 - ceci est utilisé par le routeur A au lieu de 150.1.1.2 car il est sur le même subnet que le routeur B
- Plus efficace
- Aucune configuration supplémentaire n'est nécessaire

Best Practice : Next Hop

- IOS de Cisco est, par défaut, pour que le next-hop externe se propage sans changement aux pairs iBGP
 - Cela signifie que les IGP doivent porter les next-hop externes
 - L'oublier signifie que le réseau externe est invisible
 - Avec des pairs eBGP nombreux, c'est une charge supplémentaire inutile sur les IGP
- La meilleure pratique ISP est de changer le next-hop externe à être celui du routeur local

```
neighbor x.x.x.x next-hop-self
```

Next Hop (Résumé)

- L'IGP devrait porter le chemin aux next-hop
- Route look-up récursif
- Dissocie BGP de la topologie physique réelle
- Utilise “next-hop-self” pour les next-hops externes
- Permet IGP de prendre une décision de transfert intelligente

Origine

- Transmet l'origine du préfixe
- **Attribut** historique
 - Utilisé en transition d'EGP à BGP
- Attribut transitif et obligatoire
- Influence la sélection du meilleur chemin
- Trois valeurs: IGP, EGP, incomplète
 - IGP - généré par déclaration de réseau BGP
 - EGP – généré par EGP
 - incomplète – redistribué depuis un autre protocole de routage

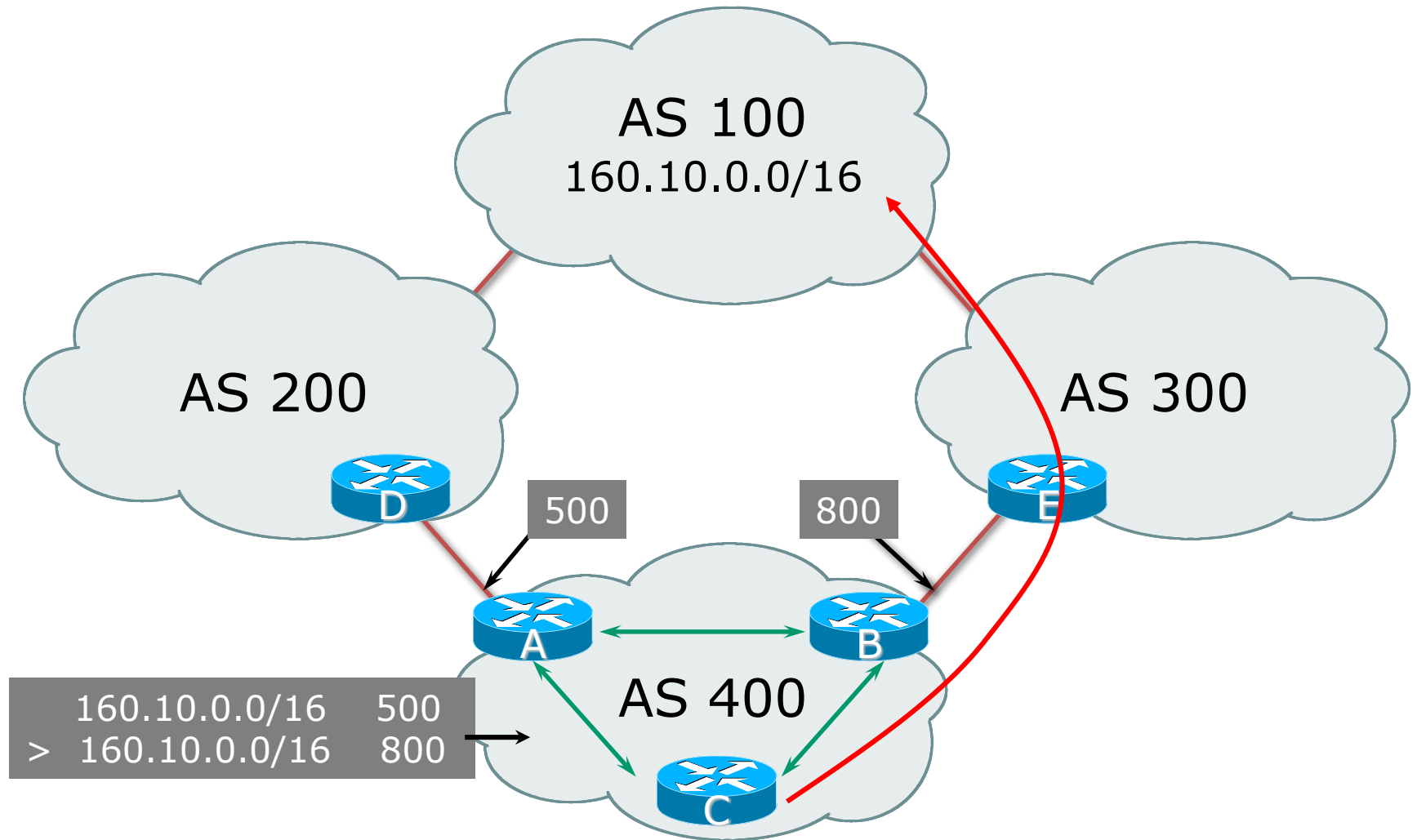
Agrégateur

- Transmet l'adresse IP du routeur ou du haut-parleur BGP créant l'itinéraire agrégat
- Attribut facultatif et transitif
- Utile pour le débogage
- N'influence pas le choix du meilleur chemin
- Créer un agrégat en utilisant " aggregate-address " définit l'attribut de l'agrégateur:

```
router bgp 100
```

```
aggregate-address 100.1.0.0 255.255.0.0
```

Local Pref



Local Pref

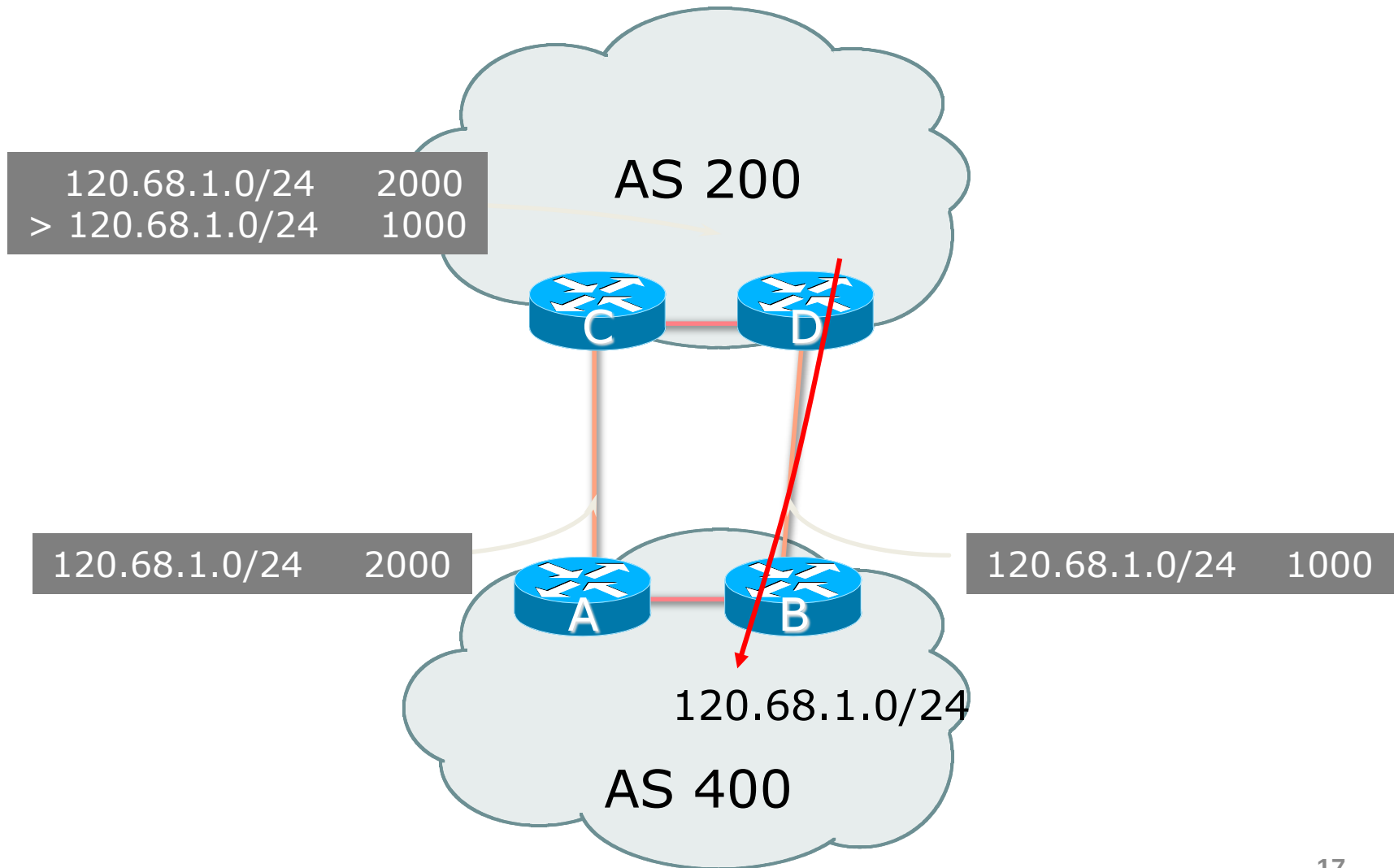
- Attribut non-transitif et optionnel
- Local à un seul AS
 - Préférence locale par défaut est 100 (IOS)
- Utilisé pour influencer la sélection de chemin BGP
 - détermine le meilleur chemin pour le trafic *sortant*
- Le chemin avec le plus de préférence locale gagne

Local Pref

- Configuration du routeur B:

```
router bgp 400
  neighbor 120.5.1.1 remote-as 300
  neighbor 120.5.1.1 route-map local-pref in
!
route-map local-pref permit 10
  match ip address prefix-list MATCH
  set local-preference 800
route-map local-pref permit 20
!
ip prefix-list MATCH permit 160.10.0.0/16
```


Multi-Exit Discriminator - MED



Discriminateur Multi-Sortie

- Inter-AS - Attribut non-transitif et optionnel
- Utilisé pour donner la préférence relative des points d'entrée
 - détermine le meilleur chemin pour le trafic entrant
- Comparables si les chemins sont du même AS
 - **bgp always-compare-med** permet de comparer les MEDs de différents AS
- Le chemin avec le plus faible MED gagne
- L'absence d'attribut MED implique une valeur MED de **zéro** (RFC4271)

MED & métrique IGP

- la métrique IGP peut être convoyé comme MED
 - **définir metric-type internal** dans route-map
 - permet à BGP d'annoncer une MED qui correspond aux valeurs métriques IGP
 - les modifications sont suivies (et de nouveau annoncées si nécessaire) tous les 600s
 - **bgp dynamic-med-interval <secs>**

MED

- Configuration du routeur B:

```
router bgp 400
  neighbor 120.5.1.1 remote-as 200
  neighbor 120.5.1.1 route-map set-med out
!
route-map set-med permit 10
  match ip address prefix-list MATCH
  set metric 1000
route-map set-med permit 20
!
ip prefix-list MATCH permit 120.68.1.0/24
```

Weight

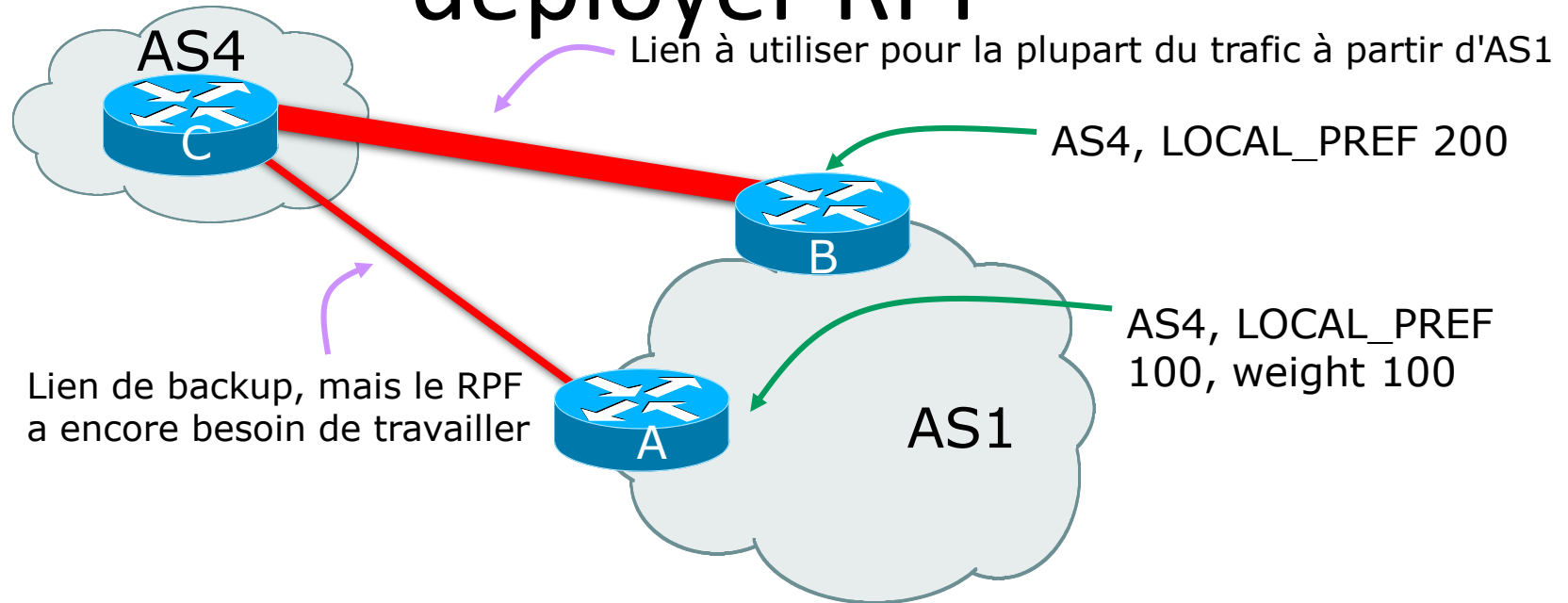
- Pas vraiment un attribut - local au routeur
- Le "Weight" le plus élevé gagne
- Appliqué à toutes les routes depuis un voisin

```
neighbor 120.5.7.1 weight 100
```

- Weight affecté à des routes en fonction du filtre

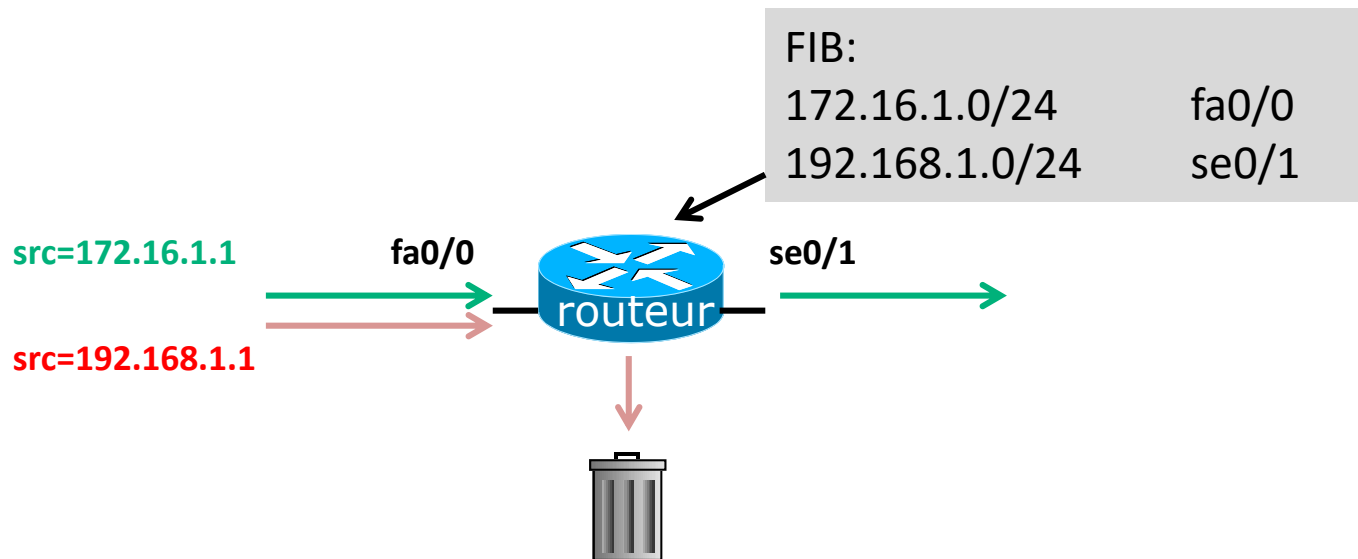
```
neighbor 120.5.7.3 filter-list 3 weight 50
```

Weight - utilisé pour aider à déployer RPF



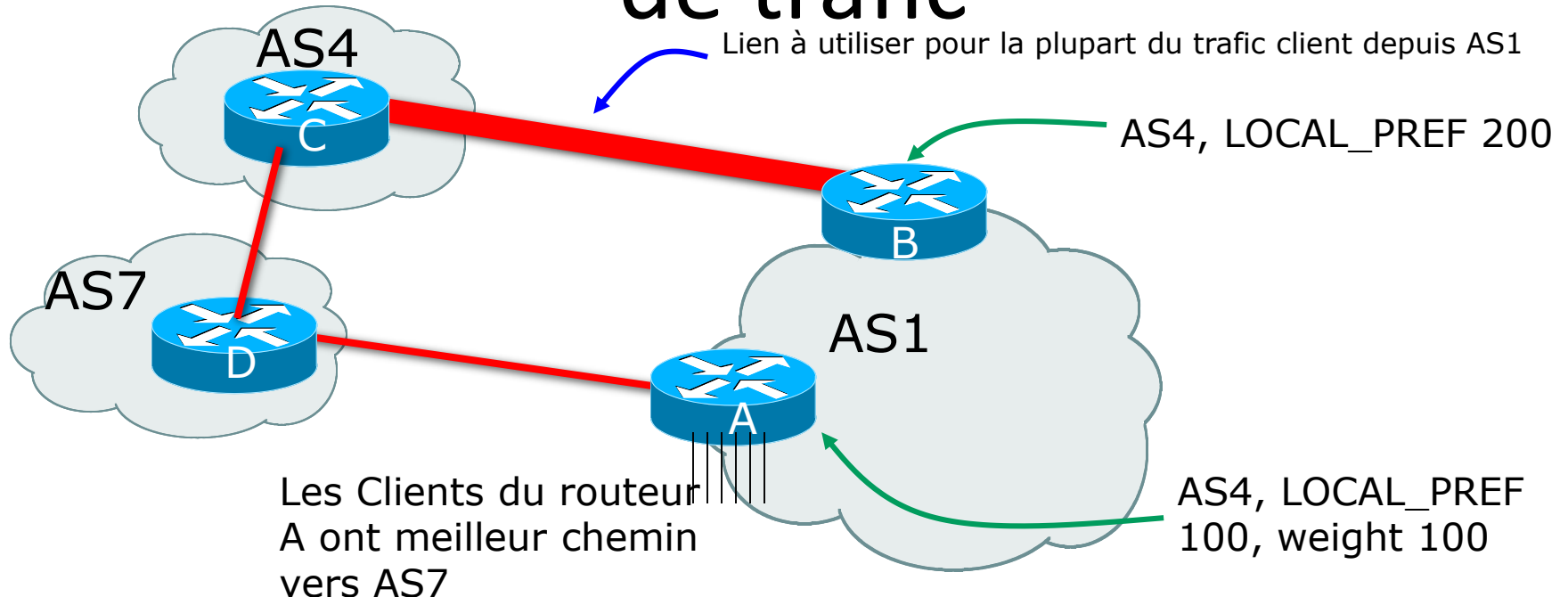
- Meilleur chemin à destination d'AS4 à partir d'AS1 est toujours via B en raison de local-pref
- Mais les paquets arrivant à A depuis AS4 à travers le lien direct C à A vont passer le contrôle du RPF car cette voie a une priorité en raison du weight défini
 - Si weight n'est pas défini, le meilleur chemin de retour vers AS4 serait via B, et la vérification RPF échouerait

En outre: qu'est-ce que uRPF?



- Le routeur compare l'adresse source du paquet entrant avec entrée FIB
 - Si l'interface d'entrée FIB correspond à l'interface entrante, le paquet est transmis
 - Si l'interface d'entrée FIB ne correspond pas à l'interface entrante, le paquet est abandonné

Weight – Utilisé pour la politique de trafic

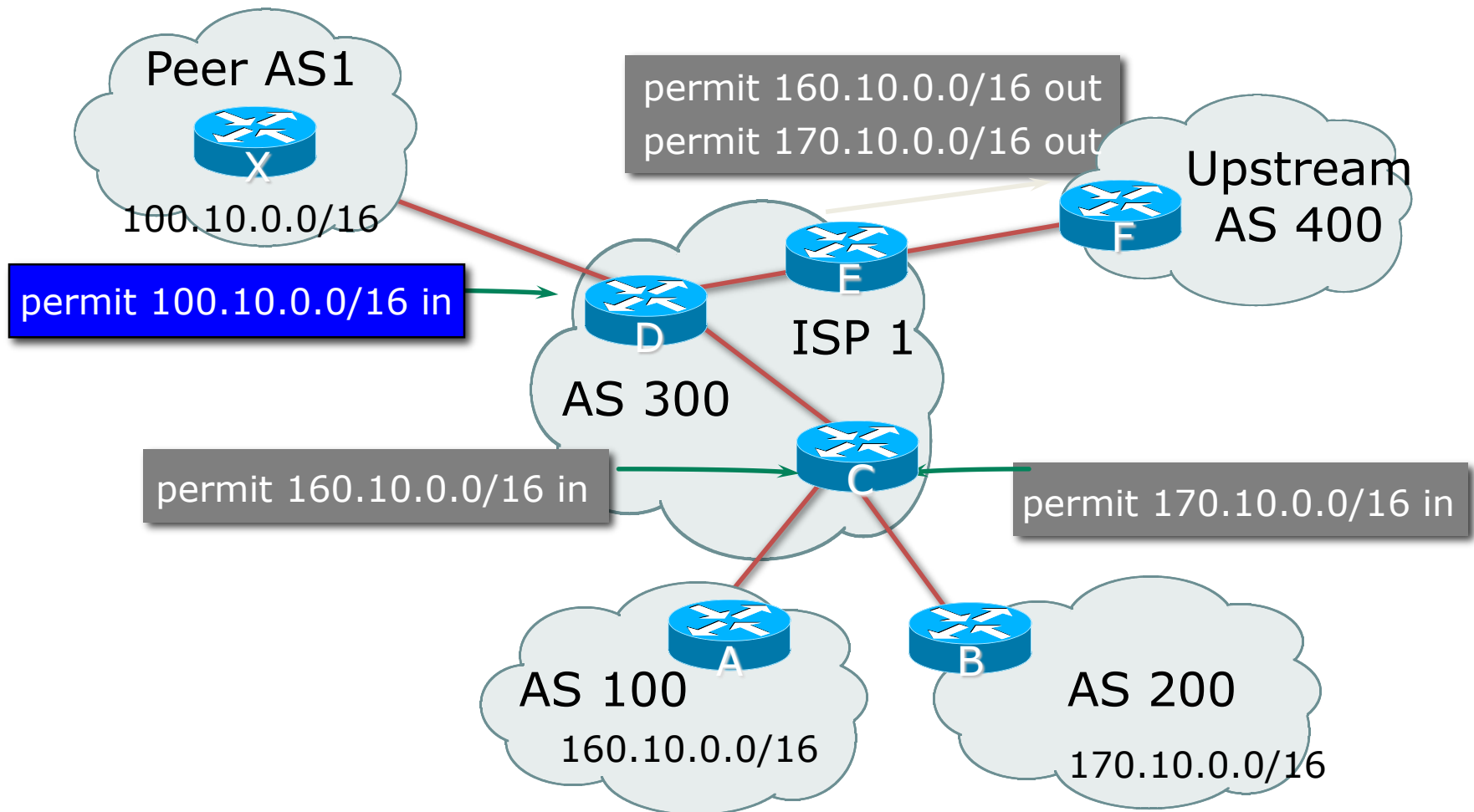


- Meilleur chemin à destination d'AS4 à partir d'AS1 est toujours via B en raison de local-pref
- Mais les clients directement connectés au routeur A utilisent le lien vers AS7 comme meilleur chemin sortant en raison du weight élevé appliqué aux routes entendues depuis AS7
 - Si le lien de A à D tombe en panne, les clients du routeur A voient le meilleur chemin via le routeur B et AS4

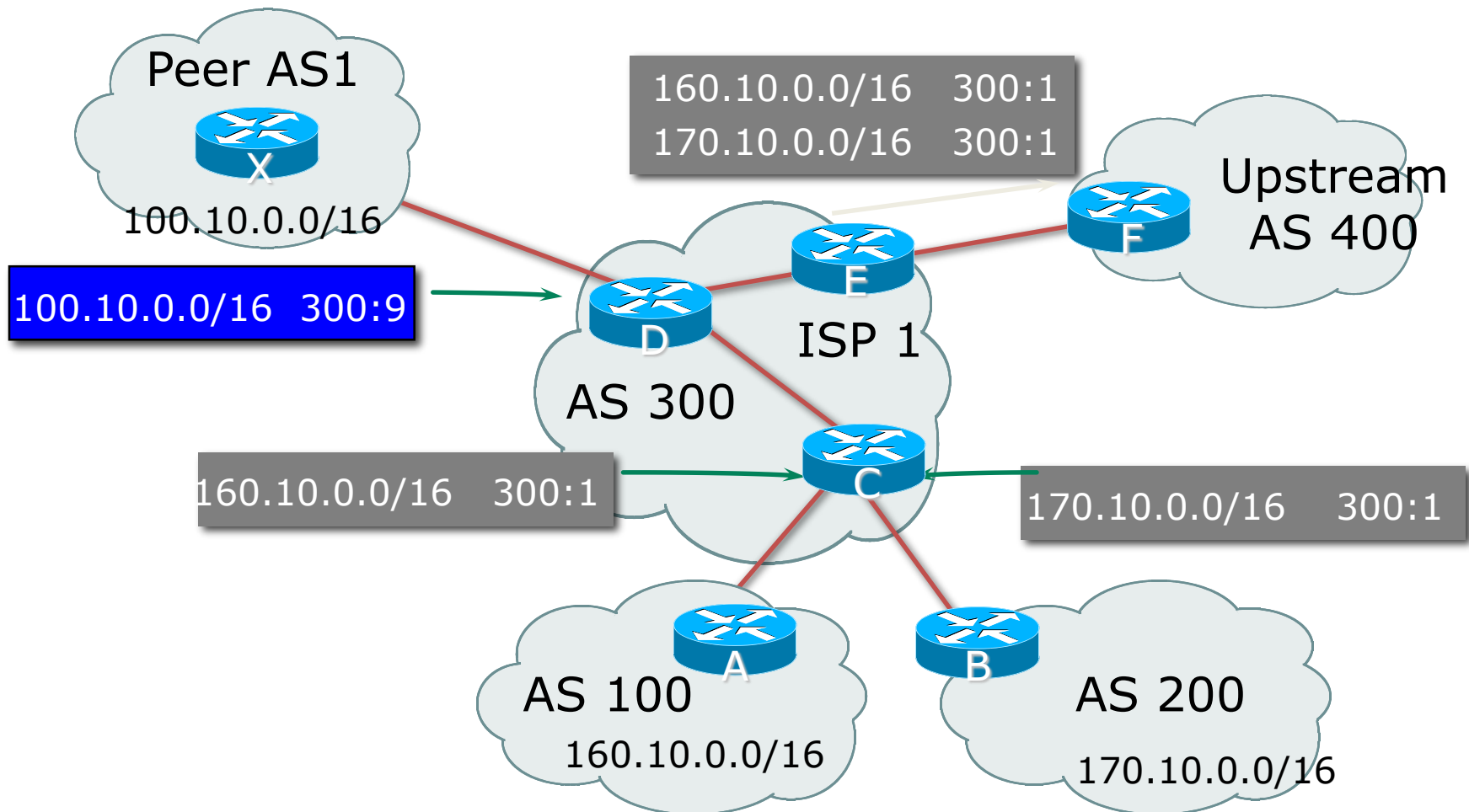
Communauté

- Les communautés sont décrites dans le RFC1997
 - Attribut transitif et optionnel
- Entier 32 bits
 - Représenté sous forme de deux entiers 16 bits (RFC1998)
 - Format commun est <local-ASN>:xx
 - 0:0 à 0:65535 et 65535:0 à 65535:65535 sont réservés
- Utilisé pour grouper des destinations
 - Chaque destination peut être membre de plusieurs communautés
- Très utile dans l'application des politiques au sein et entre les AS

Exemple de communauté (avant)



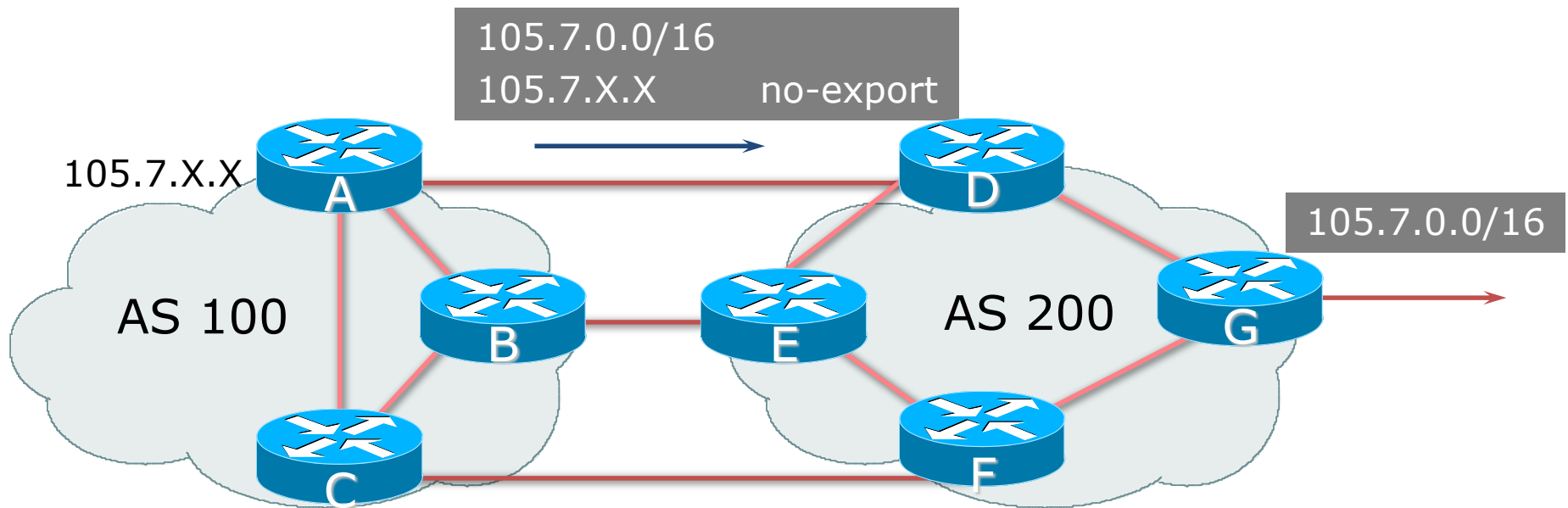
Exemple de communauté (après)



Communautés “well known”

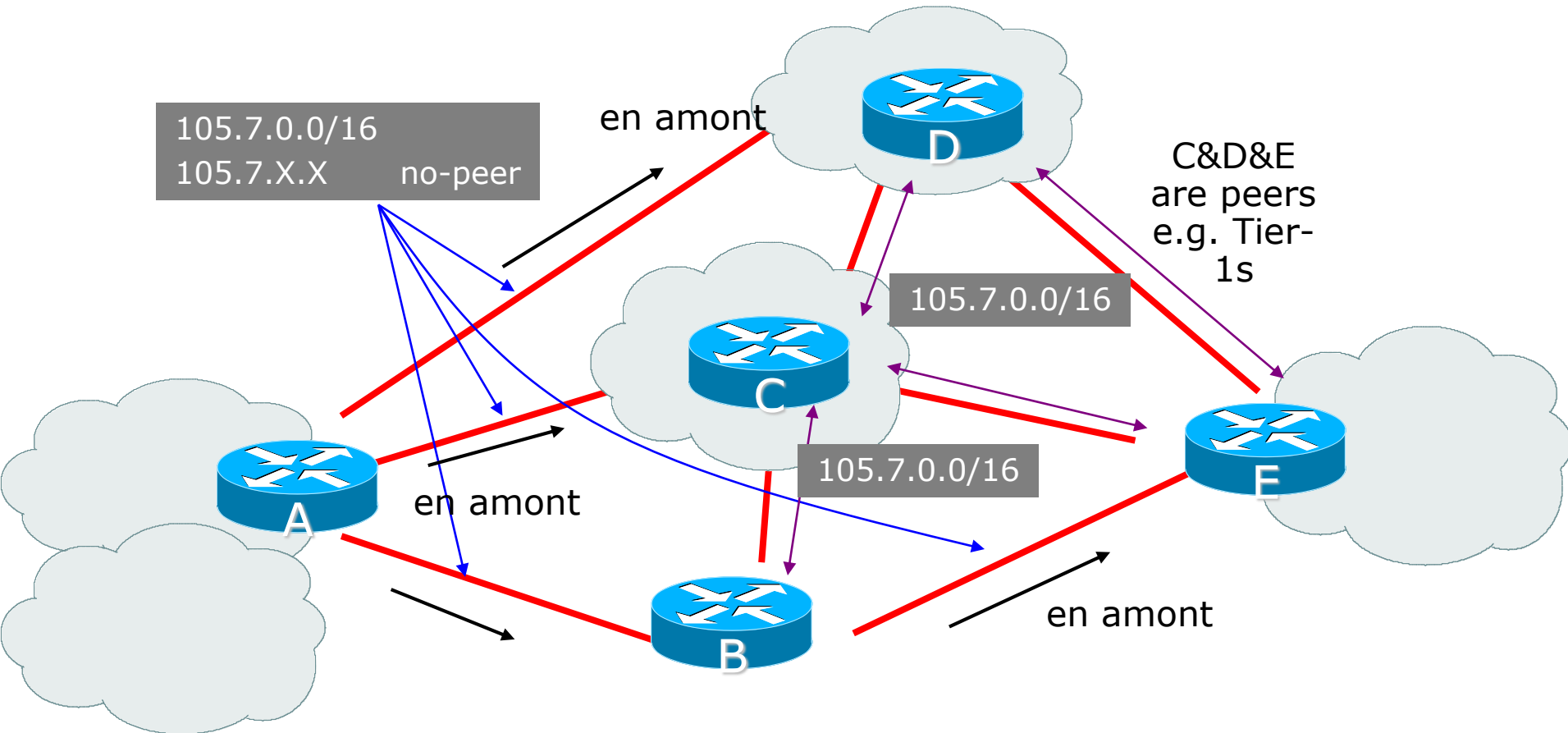
- Plusieurs communautés bien connues
 - www.iana.org/assignments/bgp-well-known-communities
- no-export 65535:65281
 - ne faire de publicité à aucun pair eBGP
- no-advertise 65535:65282
 - ne faire de publicité à aucun pair BGP (qu’ils soit iBGP ou eBGP)
- no-export-subconfed 65535:65283
 - ne pas faire de publicité à l'extérieur d'un AS local (utilisé uniquement avec les confédérations)
- no-peer 65535:65284
 - ne pas faire de publicité à des pairs bi-latéraux (RFC3765)

Communauté No-Export



- AS100 annonce agrégats et subprefixes
 - L'intention est d'améliorer le partage de charge par des fuites de subprefixes
- Subprefixes marqués avec **no-export** community
- Routeur G dans AS200 n'annonce pas des prefixes avec **no-export** community défini

Communauté No-peer



- Les sous-préfixes marqués de **no-peer** community ne sont pas envoyés aux pairs bilatéraux
 - Ils ne sont envoyés qu'aux fournisseurs en amont

Qu'en est-il des ASN 4 octets?

- Les communautés sont largement utilisées pour l'encodage de politique de routage ISP
 - Attribut 32 bits
- Le format RFC1998 est désormais pratique "standard"
 - ASN: nombre
- Bon pour les ASN 2 octets, mais ceux de 4 octets ne peuvent pas être encodés
- Solutions:
 - Utilisez un " ASN privé " pour les 16 premiers bits
 - Attendez que <http://datatracker.ietf.org/doc/draft-ietf-idr-as4octet-extcomm-generic-subtype/> soit mis en œuvre

Résumé

Attributs en action

```
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
...					

Algorithme de sélection de chemin BGP

Pourquoi est-ce le meilleur chemin?

Algorithme de sélection de chemin BGP pour IOS de Cisco: Première partie

1. Ne tenez pas compte du chemin s'il n'y a aucune route vers next hop
2. Ne pas considérer un chemin iBGP s'il n'est pas synchronisé (Cisco IOS)
3. Weight le plus élevé (du local au routeur)
4. Local Pref la plus élevée (globale au sein d'un AS)
5. Préférez une route d'origine locale
6. Le plus court AS PATH

Algorithme de sélection de chemin BGP pour IOS de Cisco: Deuxième partie

7. Code d'origine le plus bas
 - IGP < EGP < incomplete
8. Discriminateur Multi-Sortie (MED) le plus bas
 - Si **bgp deterministic-med**, commander les chemins par numéro d'AS avant de comparer
 - Si **bgp always-compare-med**, alors comparez tous les chemins
 - Sinon MED envisagé que si les chemins sont du même AS (défaut)

Algorithme de sélection de chemin BGP pour IOS de Cisco: Troisième partie

9. Préférez chemin eBGP sur chemin iBGP
10. Chemin avec IGP métrique le plus bas vers next-hop
11. Pour les chemins eBGP:
 - Si trajets multiples activés, installez N chemins parallèles dans la table de transfert
 - Si router-id est le même, passez à l'étape suivante
 - Si router-id n'est pas le même, sélectionnez la voie la plus ancienne

Algorithme de sélection de chemin BGP pour IOS de Cisco: Quatrième partie

12.router-id le plus bas (originator-id pour les routes réfléchies)

13.Cluster-list la plus courte

- Le client doit être conscient des attributs réflecteur de route!

14.neighbour address le plus bas

Attributs BGP et sélection de chemin

Ateliers de formation ISP

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





GRAND DUCHY OF LUXEMBOURG
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa
Infrastructure Trust Fund

Attributs BGP et sélection de chemin

Fin

