



GRAND DUCHY OF LUXEMBOURG  
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa  
Infrastructure Trust Fund

# Multi-homing avancé

## Ingénierie du trafic BGP



# Fournisseur de Services

## Multi-homing

- Les exemples précédents ont traité le trafic entrant en partage de charge
  - préoccupation principale à la périphérie de l'Internet
  - Qu'en est-il du trafic sortant?
- Les FAI Transit s'efforcent d'équilibrer les flux de trafic dans les deux sens
  - Utilisation équilibrée de liaison
  - Essayer et garder la plupart des flux de trafic symétrique
  - Certains FAI Tier 3 essaient et font aussi cela
- Le “Traffic Engineering” original

# Fournisseur de Services Multi-homing

- L'équilibrage du trafic sortant nécessite les informations de routage entrant
  - Solution courante est "table de routage complète"
  - rarement nécessaire
    - Pourquoi utiliser le "maillet de routage" pour tenter de résoudre les problèmes du partage de charge?
  - " Garder les choses simples " est souvent plus facile (et moins cher \$\$\$) que de transporter les N-copies de la table de routage complète

# Fournisseur de Services Multi-homing

## MYTHES!!

### MYTHES courants

1. **Vous avez besoin de la table de routage complète pour multi-homer**
  - Les gens qui vendent les mémoires de routeur voudraient vous faire croire cela
  - N'est vrai que si vous êtes un fournisseur de transit
  - Table de routage complète peut être un obstacle important au multi-homing
2. **Vous avez besoin d'un gros routeur pour multi-homer**
  - Taille de routeur est liée aux débits de données, non pas en exécutant BGP
  - En réalité, pour multihomer, votre routeur doit:
    - Avoir deux interfaces,
    - Être capable de parler BGP pour au moins deux pairs,
    - Être capable de gérer les attributs BGP,
    - Traiter au moins un préfixe
3. **BGP est complexe**
  - Dans de mauvaises mains, oui, il peut être! Gardez les choses simples!

# Multi-homing Fournisseur de services: Quelques stratégies

- Prendre les préfixes dont vous avez besoin pour faciliter l'ingénierie de trafic
  - Regardez les données de NetFlow pour les sites populaires
- Préfixes émis par vos voisins immédiats et leurs voisins vont faire davantage pour faciliter l'équilibrage de la charge que les préfixes de l'ASN de nombreux hops loin
  - Se concentrer sur les destinations locales
- Utiliser le routage par défaut autant que possible
  - Ou utiliser la table de routage complète avec soin

# Fournisseur de Services Multi-homing

- Exemples
  - Un opérateur en amont, un pair local
  - Un opérateur en amont, commutateur local
  - Deux opérateurs en amont, un pair local
  - Trois opérateurs en amont, des bandes passantes inégales de lien
- Exiger BGP et un public ASN
- Les exemples supposent que le réseau local a son propre bloc d'adresse /19

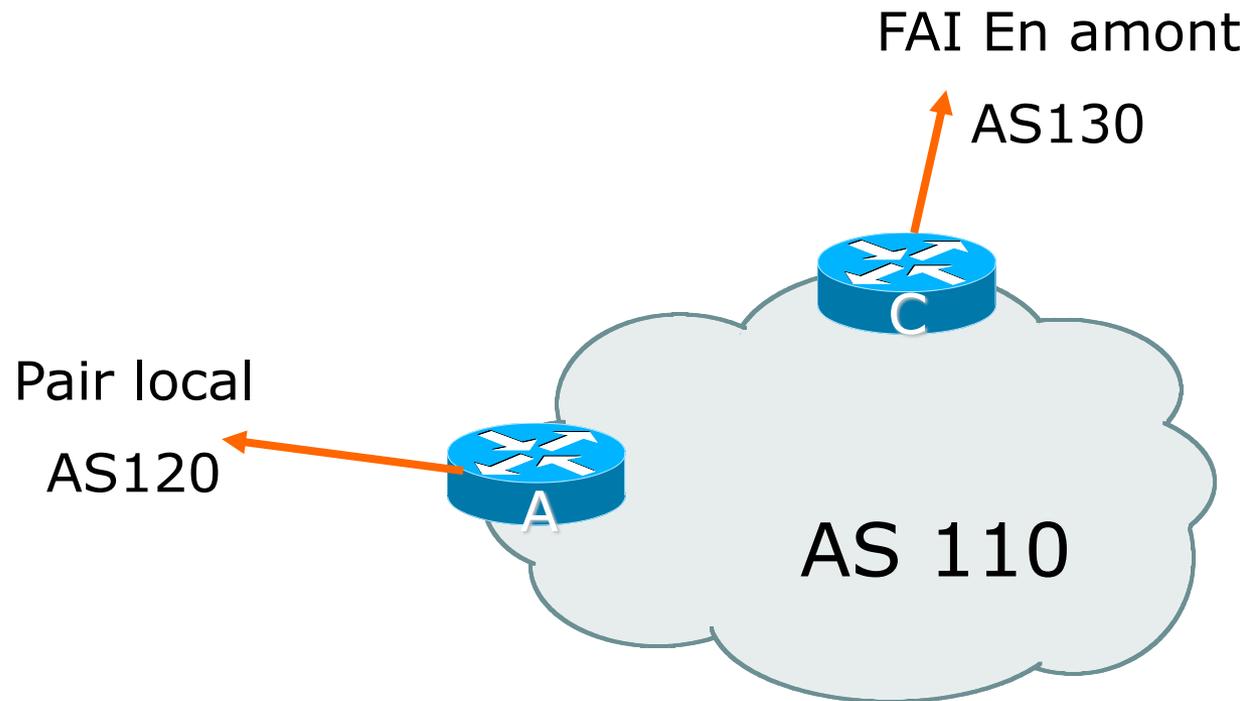
# Fournisseur de Services Multi-homing

Un opérateur en amont, un pair local

# Un opérateur en amont, un pair local

- Situation très courante dans de nombreuses régions de l'Internet
- Se connecter à un fournisseur de transit en amont pour voir “Internet”
- Se connecter à la concurrence locale afin que le trafic local reste local
  - Evite de dépenser de l'argent précieux sur les coûts de transit en amont pour le trafic local

# Un opérateur en amont, un pair local



# Un opérateur en amont, un pair local

- Annoncer /19 agrégat sur chaque lien
- Accepter route par défaut que de l'opérateur en amont
  - Soit 0.0.0.0 / 0 ou un réseau qui peut être utilisé par défaut
- Accepter toutes les routes qui proviennent du pair local

# Un opérateur en amont, un pair local

- Routeur A Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.2 remote-as 120
```

```
neighbor 122.102.10.2 prefix-list my-block out
```

```
neighbor 122.102.10.2 prefix-list AS120-peer in
```

```
!
```

```
ip prefix-list AS120-peer permit 122.5.16.0/19
```

```
ip prefix-list AS120-peer permit 121.240.0.0/20
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0 250
```

préfixe filtrés  
entrant

# Un opérateur en amont, un pair local

- Routeur A – Configuration Alternative

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.2 remote-as 120
```

```
neighbor 122.102.10.2 prefix-list my-block
```

```
neighbor 122.102.10.2 filter-list 10 in
```

```
!
```

```
ip as-path access-list 10 permit ^(120_)+$
```

```
!
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0
```

AS filtre  
chemin - plus

de

confiance”

# Un opérateur en amont, un pair local

- Routeur C Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.1 remote-as 130
```

```
neighbor 122.102.10.1 prefix-list default in
```

```
neighbor 122.102.10.1 prefix-list my-block out
```

```
!
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
ip prefix-list default permit 0.0.0.0/0
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0
```

# Un opérateur en amont, un pair local

- Deux configurations possibles pour Routeur A
  - Filter-lists assume que les pairs savent ce qu'ils font
  - Prefix-list de maintenance plus élevé, mais plus sûr
  - Certains FAI utilisent **les deux**
- Le trafic local va vers et du pair local, tout le reste va vers l'opérateur en amont

# Mis à part:

## Recommandations de configuration

- Pairs privés
  - Les peering des FAI échangent des préfixes qu'ils génèrent
  - Parfois, ils échangent des préfixes des ASN voisins aussi
- Sachez que le routeur pair privé eBGP devrait transporter que les préfixes que vous voulez que le pair privé reçoive
  - **Sinon, ils pourraient vous indiquer une route par défaut et transiter involontairement votre backbone**

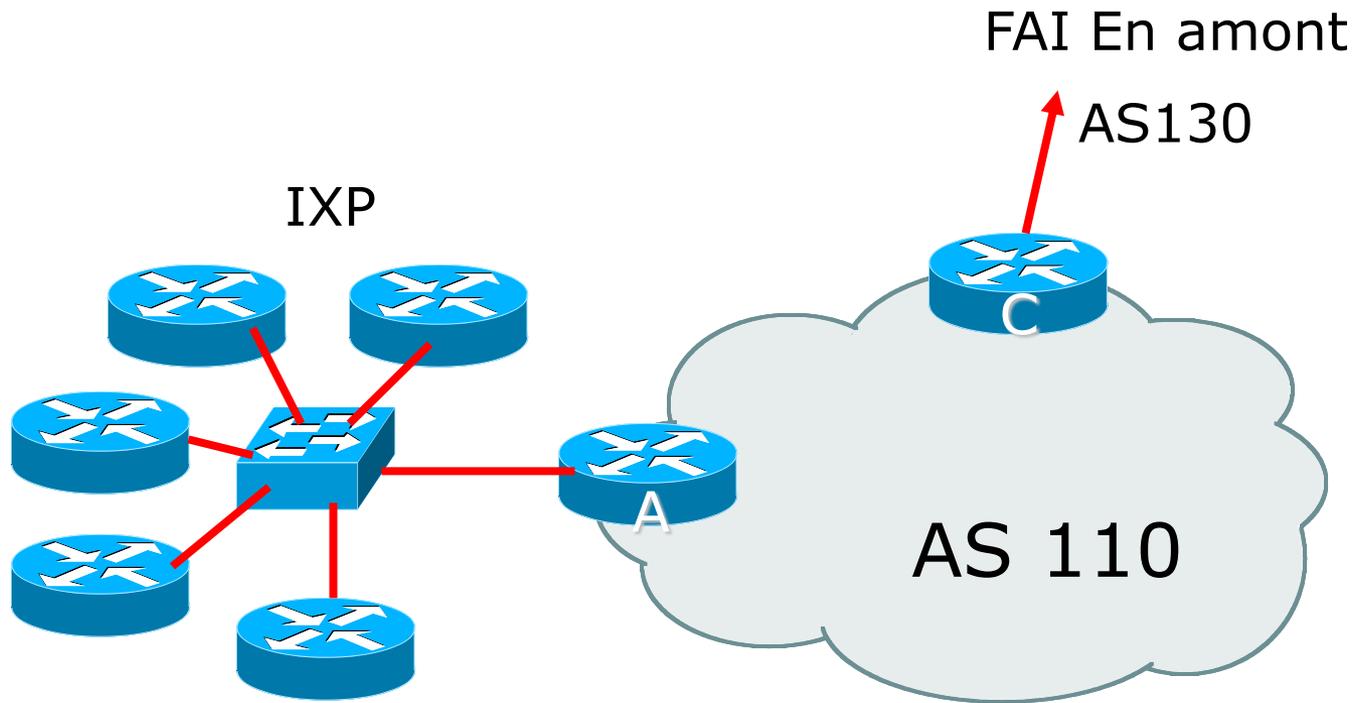
# Fournisseur de Services Multi-homing

Un opérateur en amont,  
commutateur local

# Un opérateur en amont, commutateur local

- Situation très courante dans de nombreuses régions de l'Internet
- Se connecter à un fournisseur de transit en amont pour voir “Internet”
- Se connecter au point d'échange d'Internet local afin que le trafic local reste local
  - Evite de dépenser de l'argent précieux sur les coûts de transit en amont pour le trafic local
- Cet exemple est une version mise à l'échelle de la précédente

# Un opérateur en amont, commutateur local



## Un opérateur en amont, commutateur local

- Annoncer /19 agrégat à chaque voisin AS
- Accepter la route par défaut que de l'opérateur en amont
  - Soit 0.0.0.0 / 0 ou un réseau qui peut être utilisé par défaut
- Accepter toutes les routes générées par les peers de l'IXP

# Un opérateur en amont, commutateur local

- Routeur A Configuration

```
interface fastethernet 0/0
  description Exchange Point LAN
  ip address 120.5.10.1 mask 255.255.255.224
!
router bgp 110
  neighbor ixp-peers peer-group
  neighbor ixp-peers prefix-list my-block out
  neighbor ixp-peers remove-private-AS
  neighbor ixp-peers send-community
  neighbor ixp-peers route-map set-local-pref in
```

...diapositive suivante

# Un opérateur en amont, commutateur local

```
neighbor 120.5.10.2 remote-as 100
neighbor 120.5.10.2 peer-group ixp-peers
neighbor 120.5.10.2 prefix-list peer100 in
neighbor 120.5.10.3 remote-as 101
neighbor 120.5.10.3 peer-group ixp-peers
neighbor 120.5.10.3 prefix-list peer101 in
neighbor 120.5.10.4 remote-as 102
neighbor 120.5.10.4 peer-group ixp-peers
neighbor 120.5.10.4 prefix-list peer102 in
neighbor 120.5.10.5 remote-as 103
neighbor 120.5.10.5 peer-group ixp-peers
neighbor 120.5.10.5 prefix-list peer103 in
...diapositive suivante
```

# Un opérateur en amont, commutateur local

!

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
ip prefix-list peer100 permit 122.0.0.0/19
```

```
ip prefix-list peer101 permit 122.30.0.0/19
```

```
ip prefix-list peer102 permit 122.12.0.0/19
```

```
ip prefix-list peer103 permit 122.18.128.0/19
```

!

```
route-map set-local-pref permit 10
```

```
  set local-preference 150
```

!

# Un opérateur en amont, commutateur local

- Notez que le routeur A ne génère pas l'agrégat pour AS110
  - Si le routeur A devient déconnecté du backbone, alors l'agrégat n'est plus annoncé à l'IX
  - Basculement de BGP fonctionne comme prévu
- Notez que plan de route d'arrivée qui définit la préférence locale supérieure au défaut
  - Ceci est un rappel visuel que le meilleur chemin BGP pour le trafic local sera dans l'IXP

# Un opérateur en amont, commutateur local

- Routeur C Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.1 remote-as 130
```

```
neighbor 122.102.10.1 prefix-list default in
```

```
neighbor 122.102.10.1 prefix-list my-block out
```

```
!
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
ip prefix-list default permit 0.0.0.0/0
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0
```

# Un opérateur en amont, commutateur local

- Notez la configuration du routeur A
  - Prefix-list de maintenance plus élevé, mais plus sûr
  - Aucune génération de l'agrégat AS110
- Le trafic local va vers et du IXP local, tout le reste va vers l'opérateur en amont

# Mis à part:

## Recommandations de la configuration de IXP

- Pairs IXP
  - Les peering des FAI à l'IXP échange les préfixes qu'ils génèrent
  - Parfois, ils échangent des préfixes des ASN voisins aussi
- Sachez que le routeur de frontière d'IXP ne doit porter que les préfixes que vous voulez les pairs d'IXP reçoivent et les destinations vous voulez qu'ils soient en mesure d'atteindre
  - Sinon, ils pourraient vous indiquer une route par défaut et transiter involontairement votre backbone
- Si le routeur IXP est à IX, et éloigné de votre backbone
  - Ne générez pas votre bloc d'adresse à votre routeur d'IXP

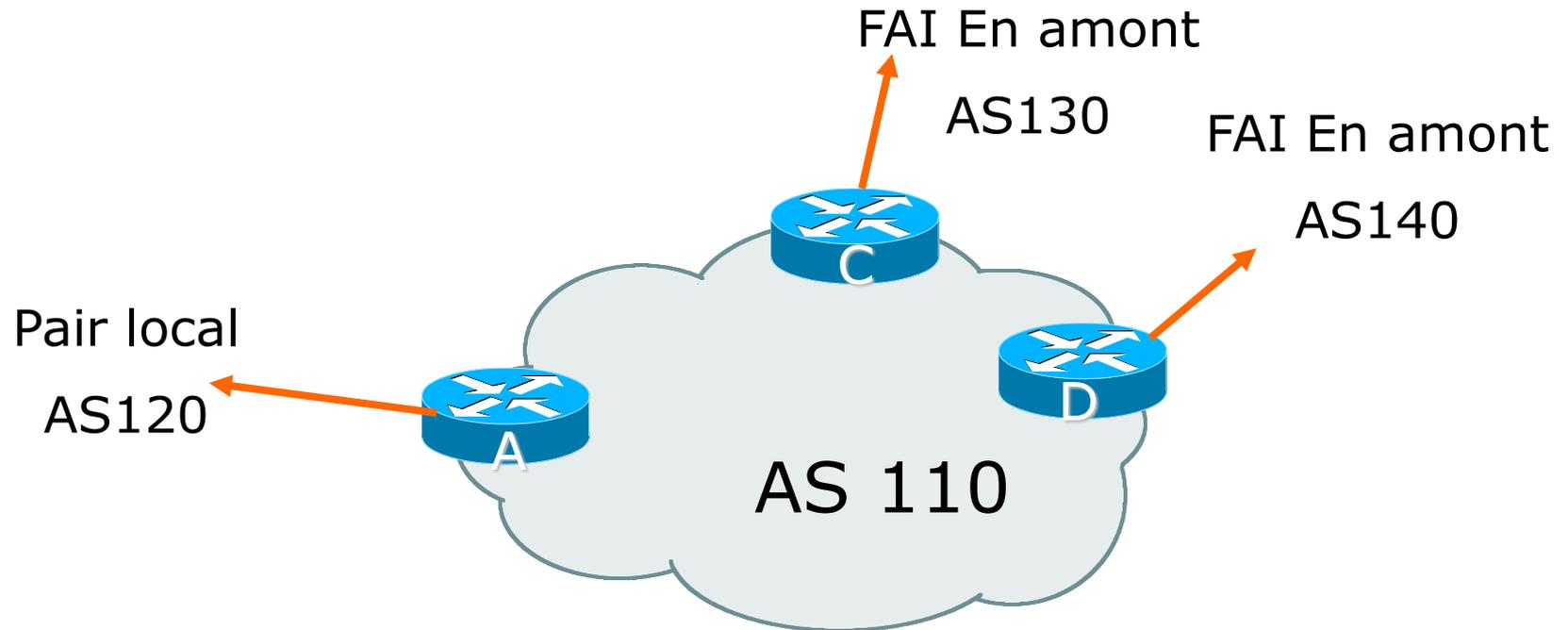
# Fournisseur de Services Multi-homing

Deux opérateurs en amont, un pair  
local

# Deux opérateurs en amont, un pair local

- Se connecter aux fournisseurs de transit en amont pour voir “Internet”
  - Fournit une redondance externe et de la diversité
    - la raison pour multi-homer
- Se connecter au pair local afin que le trafic local reste local
  - Économiser de dépenser de l'argent précieux sur les coûts de transit en amont pour le trafic local

# Deux opérateurs en amont, un pair local



# Deux opérateurs en amont, un pair local

- Annoncer /19 agrégat sur chaque lien
- Accepter la route par défaut que des opérateurs en amont
  - Soit 0.0.0.0 / 0 ou un réseau qui peut être utilisé par défaut
- Accepter toutes les routes générées par le pair local
- Notez la séparation du Routeur C et D
  - Routeur de bordure unique signifie pas de redondance
- Routeur A
  - Même configuration de routage comme dans l'exemple avec un opérateur en amont et un pair local

# Deux opérateurs en amont, un pair local

- Routeur C Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.1 remote-as 130
```

```
neighbor 122.102.10.1 prefix-list default in
```

```
neighbor 122.102.10.1 prefix-list my-block out
```

```
!
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
ip prefix-list default permit 0.0.0.0/0
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0
```

# Deux opérateurs en amont, un pair local

- Routeur D Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.5 remote-as 140
```

```
neighbor 122.102.10.5 prefix-list default in
```

```
neighbor 122.102.10.5 prefix-list my-block out
```

```
!
```

```
ip prefix-list my-block permit 121.10.0.0/19
```

```
ip prefix-list default permit 0.0.0.0/0
```

```
!
```

```
ip route 121.10.0.0 255.255.224.0 null0
```

# Deux opérateurs en amont, un pair local

- C'est la configuration simple pour le Routeur C et D
- Le trafic aux deux opérateurs en amont prendra la sortie la plus proche
  - Routeurs peu coûteux requis
  - Ce n'est pas utile en pratique surtout pour les liaisons internationales
  - Partage de charge doit être meilleur

# Deux opérateurs en amont, un pair local

- Bonnes options de configuration:
  - Accepter la Full table des deux opérateurs en amont
    - **Coûteux et inutile!**
  - Accepter la route par défaut d'un opérateur en amont et certaines routes des autres opérateurs en amont
    - **Le chemin à parcourir!**

# Deux opérateurs en amont, un pair local

## Routes complète

- Routeur C Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.1 remote-as 130
```

```
neighbor 122.102.10.1 prefix-list rfc1918-deny in
```

```
neighbor 122.102.10.1 prefix-list my-block out
```

```
neighbor 122.102.10.1 route-map AS130-loadshare in
```

!

```
ip prefix-list my-block permit 121.10.0.0/19
```

! See [www.cymru.com/Documents/bogon-list.html](http://www.cymru.com/Documents/bogon-list.html)

! ...for "RFC1918 and friends" list

...diapositive suivante

Permettre à tous les  
préfixes en dehors  
de RFC1918 et les  
amis

# Deux opérateurs en amont, un pair local Routes complète

```
ip route 121.10.0.0 255.255.224.0 null0
!  
ip as-path access-list 10 permit ^(130_)+$  
ip as-path access-list 10 permit ^(130_)+_[0-9]+$  
!  
route-map AS130-loadshare permit 10  
  match ip as-path 10  
  set local-preference 120  
!  
route-map AS130-loadshare permit 20  
  set local-preference 80  
!
```

# Deux opérateurs en amont, un pair local

## Routes complète

- Routeur D Configuration

```
router bgp 110
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.5 remote-as 140
```

```
neighbor 122.102.10.5 prefix-list rfc1918-deny in
```

```
neighbor 122.102.10.5 prefix-list my-block out
```

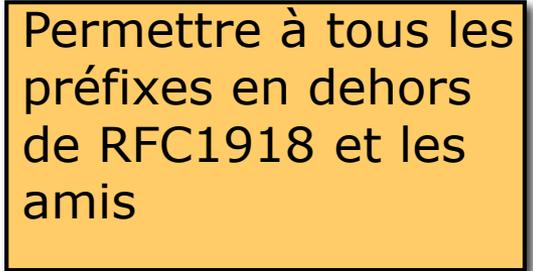
!

```
ip prefix-list my-block permit 121.10.0.0/19
```

! See [www.cymru.com/Documents/bogon-list.html](http://www.cymru.com/Documents/bogon-list.html)

! ...for "RFC1918 and friends" list

Permettre à tous les  
préfixes en dehors  
de RFC1918 et les  
amis



# Deux opérateurs en amont, un pair local

## Routes complète

- Routeur C configuration:
  - Accepter les routes de AS130
  - Tag préfixes émis par AS130 et AS130 avoisinant AS avec une préférence locale 120
    - Le trafic vers les AS va aller sur le lien AS130
  - Préfixes restants marqués avec la préférence locale de 80
    - Le trafic vers tous les autres AS va aller sur le lien AS140
- Routeur D même configuration que C Routeur sans le plan de route

# Deux opérateurs en amont, un pair local

## Routes complète

- Routes complètes des opérateurs en amont
  - Cher – besoin de beaucoup de mémoire et de CPU
  - Besoin de jouer avec les préférences
  - L'exemple précédent n'est qu'un exemple - la vraie vie aura besoin de meilleur réglage fin!
  - L'exemple précédent ne considère pas le trafic entrant – voir plus haut dans la présentation des exemples

# Deux opérateurs en amont, un pair local

## Routes partielles: Stratégie

- Demander un opérateur en amont pour une route par défaut
  - Facile de générer le défaut vers un voisin BGP
- Demander à un autre opérateur en amont d'une table de routage complète
  - Puis filtrer cette table de routage basée sur le voisinage ASN
  - E.g. vouloir le trafic des voisins pour aller sur le lien de l'ASN
  - La plupart de ce que l'opérateur en amont envoie est jeté
  - Plus facile que de demander l'opérateur en amont de mettre en place des filtres personnalisés BGP pour vous

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routeur C Configuration

```
router bgp 110
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 prefix-list rfc1918-nodef-deny in
  neighbor 122.102.10.1 prefix-list my-block out
  neighbor 122.102.10.1 filter-list 10 in
  neighbor 122.102.10.1 route-map tag-default-low in
```

!

...diapositive suivante

Permettre à tous les préfixes et par défaut, refuser RFC1918 et les amis

AS la liste de filtre filtre des préfixes basés sur l'origine ASN

# Deux opérateurs en amont, un pair local

## Routes partielles

```
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
!
ip as-path access-list 10 permit ^(130_)+$
ip as-path access-list 10 permit ^(130_)+_[0-9]+$
!
route-map tag-default-low permit 10
  match ip address prefix-list default
  set local-preference 80
!
route-map tag-default-low permit 20
!
```

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routeur D Configuration

```
router bgp 110
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.5 remote-as 140
  neighbor 122.102.10.5 prefix-list default in
  neighbor 122.102.10.5 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routeur C configuration:
  - Accepter les routes pleines de AS130
    - (ou les amener à envoyer moins)
  - Filtrer les ASN donc seul les AS130 et AS130 au voisinage de AS sont acceptés
  - Permettre par défaut, et définir les préférences local 80
  - Le trafic vers les AS va aller sur le lien AS130
  - Le trafic vers tous les autres AS va aller sur le lien AS140
  - Si AS140 lien échoue, sauvegarde via AS130 - et vice-versa

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routeur C IGP Configuration

```
router ospf 110
default-information originate metric 30
passive-interface Serial 0/0
!
ip route 0.0.0.0 0.0.0.0 serial 0/0 254
```

- Routeur D IGP Configuration

```
router ospf 110
default-information originate metric 10
passive-interface Serial 0/0
!
ip route 0.0.0.0 0.0.0.0 serial 0/0 254
```

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routes partielles des opérateurs en amont
  - Utiliser OSPF pour déterminer le chemin d'accès sortant
  - Le défaut du routeur D a métrique 10 – chemin sortant principal
  - Le défaut du routeur C a métrique 30 – chemin sortant de sauvegarde
  - Interface série tombe en panne, statique par défaut est retirée de la table de routage, OSPF par défaut retiré

# Deux opérateurs en amont, un pair local

## Routes partielles

- Routes partielles des opérateurs en amont
  - Pas cher – porte seulement les routes nécessaires pour le partage de charge
  - Besoin de filtrer sur les chemins AS
  - L'exemple précédent n'est qu'un exemple - la vraie vie aura besoin de meilleur réglage fin!
  - L'exemple précédent ne considère pas le trafic entrant – voir plus haut dans la présentation des exemples

# Mis à part:

## Recommandation de configuration

- Lors de la distribution interne de la route par défaut par iBGP ou OSPF/ISIS
  - Assurez-vous que les routeurs connectés à des pairs privées ou à des IXP **ne transportent pas** la route par défaut
  - Sinon, ils pourraient vous indiquer une route par défaut et transiter involontairement votre backbone
  - Simple fix pour les routeurs privé Peer/IXP:

```
ip route 0.0.0.0 0.0.0.0 null0
```

# Fournisseur de Services Multi-homing

Trois opérateurs en amont, des  
bandes passantes inégales

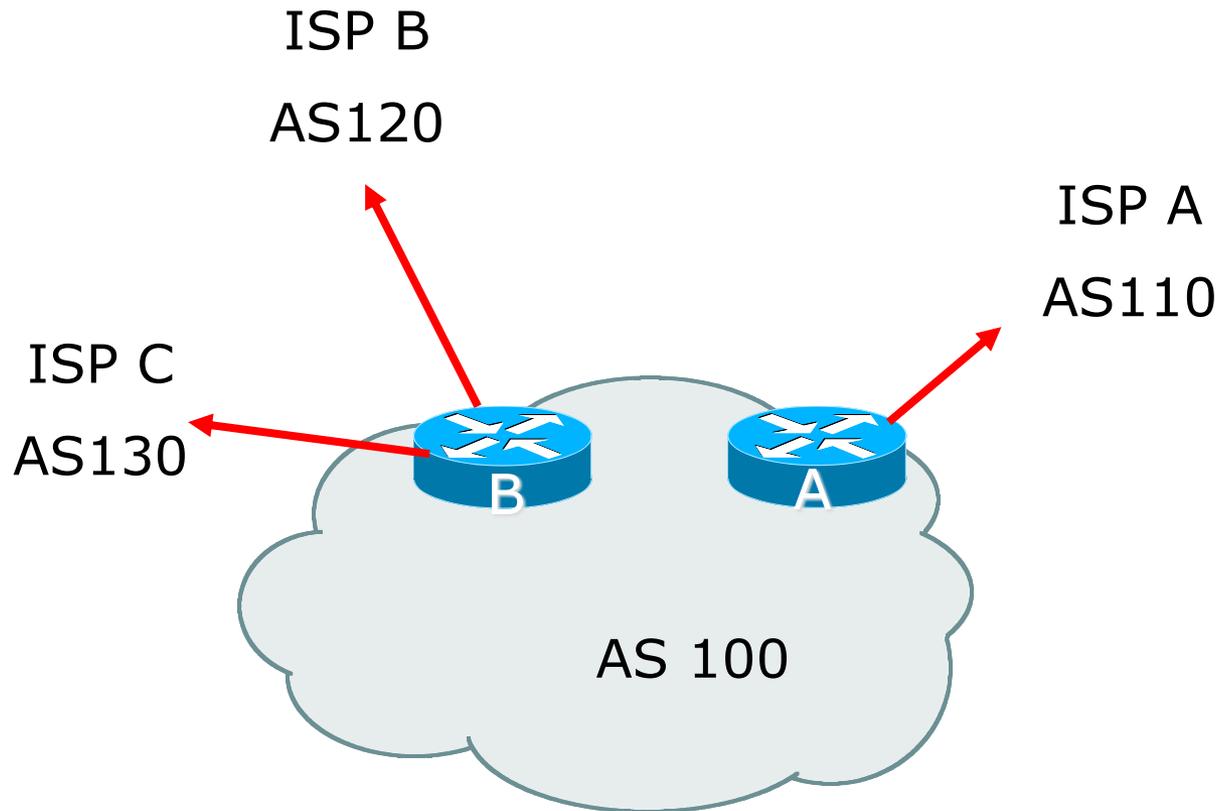
# Trois opérateurs en amont, des bandes passantes inégales

- Système Autonome a trois opérateurs en amont
  - 16Mbps pour ISP A
  - 8Mbps pour ISP B
  - 4Mbps pour ISP C
- Quelle est la stratégie ici?
  - Une option est une table complète de chacun
    - 3x 400k préfixes  $\Rightarrow$  1200k chemins
  - Autre option est la table partielle et des défauts de chacun
    - Comment??

# Stratégie

- Deux routeurs externes (donne redondance des routeurs)
  - Pas besoin de trois routeurs pour cela
- Connecter la plus grande bande passante à un routeur
  - La plupart du trafic entrant et sortant iront ici
- Connecter les deux autres liens vers le deuxième routeur
  - Offre une capacité maximale de sauvegarde si le lien principal tombe en panne
- Utiliser le plus grand lien comme défaut
  - La plupart du trafic entrant et sortant iront ici
- Faire de l'ingénierie du trafic sur les deux plus petits liens
  - Focus sur les besoins du trafic régional

# Diagramme



- Routeur A a un circuit de 16Mbps pour ISP A
- Routeur B a des circuits de 8Mbps et de 4Mbps aux FAI B&C

# Stratégie d'équilibrage de charge sortant

- Disponible BGP se nourrit des fournisseurs de transit:
  - Table complète
  - Préfixes des clients et défaut
  - Route par défaut
- Voici les options courantes sur Internet aujourd'hui
  - Très rare pour n'importe quel fournisseur d'offrir quelque chose de différent
  - Très rare que n'importe quel fournisseur de personnaliser le flux BGP pour un client

# Stratégie d'équilibrage de charge sortant

- Accepter qu'une route par défaut du fournisseur avec **la plus grande** connectivité, ISP A
  - Parce que la plupart du trafic va utiliser ce lien
- Si FAI A ne fournira pas un défaut:
  - Toujours exécuter BGP avec eux, mais abandonner tous les préfixes
  - Pointer la route statique par défaut sur le lien en amont
  - Distribuer le défaut dans l'IGP
- Demander la table complète des FAI B & C
  - La plupart de ceci est jeté
  - ( “Défaut plus clients” ne suffit pas)

# Stratégie d'équilibrage de charge sortant

- Comment décider de ce qu'il faut conserver et ce qu'il faut abandonner des FAI B & C?
  - Plus de trafic utilisera le lien du FAI A — et nous devons donc trouver un sous-ensemble bien/utile
- Abandonner préfixes transitant par les FAI mondiaux de transit
  - FAI de transit mondiaux apparaissent généralement dans la plupart des AS-PATHs non-locaux ou régionaux
- Abandonner les préfixes avec ASN du FAI A dans le chemin
  - Semble plus raisonnable pour le trafic d'aller à ces destinations via le lien du FAI A

# Stratégie d'équilibrage de charge sortant

- Global Transit FAI comprennent:  
209 CenturyLink 3549 Niveau 3  
701 VerizonBusiness3356 Niveau 3  
1239 Sprint 3561 Savvis  
1668 AOL TDN 7018 AT&T  
2914 NTT America

# FAI B peering entrant AS-PATH filtre

```
ip as-path access-list 1 deny _209_  
ip as-path access-list 1 deny _701_  
ip as-path access-list 1 deny _1239_  
ip as-path access-list 1 deny _3356_  
ip as-path access-list 1 deny _3549_  
ip as-path access-list 1 deny _3561_  
ip as-path access-list 1 deny _2914_  
ip as-path access-list 1 deny _7018_  
!  
ip as-path access-list 1 deny _ISPA_ ←  
ip as-path access-list 1 deny _ISPC_  
!  
ip as-path access-list 1 permit _ISPB$  
ip as-path access-list 1 permit _ISPB_[0-9]+$  
ip as-path access-list 1 permit _ISPB_[0-9]+_[0-9]+$  
ip as-path access-list 1 permit _ISPB_[0-9]+_[0-9]+_[0-9]+$  
ip as-path access-list 1 deny .*
```

Pas besoin de l' ISPA  
et des préfixes ISPC  
via ISPB

# Stratégie d'équilibrage de charge sortant:

## Configuration de ISP B peering

- Partie 1: Suppression des préfixes des FAI Global Transit
  - Cela peut être affinée si le volume du trafic n'est pas suffisant
  - (Plusieurs préfixes entrant = plus de trafic sortant)
- Partie 2: Suppression des préfixes qui transitent par le réseau du FAI A & C
- Partie 3: Permettre les préfixes du FAI B, leurs voisins BGP, et leurs voisins, et leurs voisins
  - Plus de clauses d'autorisation AS\_PATH, plus de préfixes entrant, plus de trafic sortant
  - Trop de préfixes signifie plus de trafic sortant que le lien vers FAI B peut gérer

# Stratégie d'équilibrage de charge sortant

- Filtre AS-PATH semblable peut être construit pour le peering de ISP C BGP
- Si les mêmes préfixes sont entendus à la fois du FAI B et C, puis établir la proximité de leur origine ASN à l'ISP B ou C
  - e.g. FAI B peut être au Japon, avec l'ASN voisin en Europe, encore FAI C pourrait être en Europe
  - Transit à l'ASN via FAI C est plus logique dans ce cas

# Stratégie d'équilibrage de charge entrant

- Le plus grand lien sortant doit annoncer simplement l'agrégat
- Les autres liens devraient annoncer:
  - a) L'agrégat avec AS-PATH prepend
  - b) Les sous-préfixes d'agrégats, choisis selon les volumes de trafic pour ces sous-préfixes et selon les services sur ces sous-préfixes
- Exemple:
  - Le lien au FAI B pourrait être employé juste pour des clients du haut débit/réseau commuté — numéroter ainsi tels clients sur un sous-préfixe contigu
  - Le lien au FAI C pourrait être employé juste pour les clients commerciaux de ligne louée — numérotez ainsi tous tels clients sur un sous-préfixe contigu

# Routeur A: Configuration eBGP Exemple

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote 110
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list aggregate out
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list aggregate permit 100.10.0.0/19
!
```

# Routeur B: Configuration eBGP Exemple

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor 120.103.1.1 remote 120
  neighbor 120.103.1.1 filter-list 1 in
  neighbor 120.103.1.1 prefix-list ISP-B out
  neighbor 120.103.1.1 route-map to-ISP-B out
  neighbor 121.105.2.1 remote 130
  neighbor 121.105.2.1 filter-list 2 in
  neighbor 121.105.2.1 prefix-list ISP-C out
  neighbor 121.105.2.1 route-map to-ISP-C out
!
ip prefix-list aggregate permit 100.10.0.0/19
!
```

...diapositive suivante

# Routeur B: Configuration eBGP

## Exemple

```
ip prefix-list ISP-B permit 100.10.0.0/19
ip prefix-list ISP-B permit 100.10.0.0/21
!
ip prefix-list ISP-C permit 100.10.0.0/19
ip prefix-list ISP-C permit 100.10.28.0/22
!
route-map to-ISP-B permit 10
  match ip address prefix-list aggregate
  set as-path prepend 100
!
route-map to-ISP-B permit 20
!
route-map to-ISP-C permit 10
  match ip address prefix-list aggregate
  set as-path prepend 100 100
!
route-map to-ISP-C permit 20
```

/21 pour FAI B  
"clients du  
réseau  
commuté"

/22 pour FAI C  
"clients biz"

e.g. Unique  
prepend sur le  
lien de FAI B

e.g. Double  
prepend sur le  
lien du FAI C

# Qu'en est-il des liens de backup?

- Nous avons:
  - Une Route par défaut du FAI A par eBGP
  - La plupart du temps la table complète filtrée des FAI B & C
- Stratégie:
  - Générer une route par défaut par OSPF sur le routeur A (avec métrique 10) - lien vers FAI A
  - Générer une route par défaut par OSPF sur le routeur B (avec métrique 30) - liens vers des FAI B & C
  - Plus sur le Routeur B:
    - Route statique par défaut à l'ISP B avec la distance 240
    - Route statique par défaut à l'ISP C avec la distance 245
  - Lorsque le lien tombe, la route statique est retirée

# Backup

- Etat d'équilibre

(tous les liens connectés et actifs):

- Route par défaut est au routeur A - 10 métrique OSPF
- (Parce que par défaut appris par eBGP  $\Rightarrow$  défaut est dans RIB  $\Rightarrow$  OSPF provient par défaut)
- route par défaut de backup est au routeur B - OSPF métrique 20
- Les préfixes d'eBGP ont appris des opérateurs en amont distribués par l'iBGP tout au long du bacbone
- (Le défaut peut être filtré dans l'iBGP pour éviter “l'erreur d'échec de RIB” )

# Sauvegarde sortant : exemples d'échec

- Lien vers FAI A interrompu, vers FAI B&C connecté:
  - Route par défaut est au routeur B - OSPF métrique 20
  - (le défaut d'eBGP disparu de RIB, donc OSPF sur le routeur A retire le défaut)
- Ci-dessus est vraie si lien vers B ou C est interrompu aussi
- Le lien vers les FAI B et C interrompu, lien vers FAI A connecté:
  - Route par défaut est au routeur A - 10 métrique OSPF
  - (les défaut statiques sur le routeur B retirés de RIB, donc OSPF sur le routeur B retire le défaut)

# Autres considérations

- La Route par défaut ne doit pas être propagé à des dispositifs de terminaison non-transit pairs et des clients
- on a Rarement le besoin de transporter la route par défaut dans iBGP
  - Mieux pour filtrer par défaut dans iBGP la maille des peerings
- Porter encore les autres préfixes eBGP dans la maille iBGP
  - Sinon les routeurs suivront les règles de route par défaut résultant du trafic sous-optimale
  - Pas un gros problème parce que ne portant pas la table complète

# Routeur A: Configuration iBGP Exemple

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor ibgp-peers peer-group
  neighbor ibgp-peers remote-as 100
  neighbor ibgp-peers prefix-list ibgp-filter out
  neighbor 100.10.0.2 peer-group ibgp-peers
  neighbor 100.10.0.3 peer-group ibgp-peers
!
ip prefix-list ibgp-filter deny 0.0.0.0/0
ip prefix-list ibgp-filter permit 0.0.0.0/0 le 32
!
```

# Trois opérateurs en amont, des bandes passantes inégales:

## Résumé

- Exemple basé sur de nombreuses topologies multi-homing/équilibre de charge de travail déployées
- De nombreuses variantes possibles - celui-ci est:
  - Facile à régler
  - Lumière sur les ressources des routeurs périphériques
  - Lumière sur l'infrastructure du routeur de backbone
  - Table clairsemée de BGP  $\Rightarrow$  convergence plus rapide

# Reconnaissance et attribution

**Cette présentation contient des contenus et des informations initialement développés et gérés par les organisations / personnes suivantes et fournie pour le projet AXIS de l'Union africaine**

**Cisco ISP/IXP Workshops**

**Philip Smith: - [pfsinoz@gmail.com](mailto:pfsinoz@gmail.com)**





GRAND DUCHY OF LUXEMBOURG  
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa  
Infrastructure Trust Fund

# Multi-homing avancé

Fin

