



GRAND DUCHY OF LUXEMBOURG  
Ministry of Foreign Affairs

Directorate for Development Cooperation



European Union Africa  
Infrastructure Trust Fund

# Multi-homing simple



# Pourquoi le multi-homing?

- Redondance
  - Une connexion internet signifie que le réseau dépend du:
    - Routeur local (configuration, logiciels, matériel)
    - Média du réseau étendu (WAN) (défaillance physique, défaillance du transporteur)
    - Fournisseur de services en amont (configuration, logiciels, matériel)

# Pourquoi le multi-homing?

- Fiabilité
  - Applications critiques d'entreprise exigent une disponibilité continue
  - Le manque de redondance implique le manque de fiabilité qui implique la perte de revenu

# Pourquoi le multi-homing?

- Diversité de fournisseur
  - De toute évidence, de nombreuses entreprises exigent la diversité des fournisseurs
  - Une connexion Internet à partir de deux ou plusieurs fournisseurs
    - Avec deux ou plusieurs chemins d'accès au réseau étendu (WAN)
    - Avec deux ou plusieurs points de sortie
    - Avec deux ou plusieurs liaisons internationales
    - **Tout doubler**

# Pourquoi le multi-homing?

- Pas vraiment une raison, mais souvent cité...
- Résultat:
  - Faire le choix entre plusieurs fournisseurs d'accès Internet:
    - Qualité de service
    - Service Offerts
    - Disponibilité

# Pourquoi multi-homing?

- Résumé:
  - Le multi-homing est facile d'exiger comme condition de toute opération
  - Mais qu'est-ce que cela signifie vraiment:
    - En réalité?
    - Pour le réseau?
    - Pour l'Internet?
  - Et comment pouvons-nous le faire?

# Définition du multi-homing

- Plus d'un lien externe au réseau local
  - deux ou plusieurs liens vers le même FAI
  - deux ou plusieurs liens vers différents FAI
- Habituellement **deux** routeurs face externes
  - un routeur donne le lien et fournit la redondance seulement

# multi-homing

- Les scénarios décrits ici s'appliquent tout aussi bien pour les sites finals étant les clients des FAI et des FAI étant les clients d'autres FAI
- Détail d'implémentation peut être différent
  - sites finals → ISP      FAI commande config
  - ISP1 → ISP2      FAI partage config

# Numéro de système autonome (NSA)

- Deux gammes
  - 0-65535 (gamme originale de 16-bit)
  - 65536-4294967295 (gamme 32-bit – RFC4893)
- Usage:
  - 0 and 65535 (Réservé)
  - 1-64495 (Internet public)
  - 64496-64511 (documentation – RFC5398)
  - 64512-65534 (usage privé uniquement)
  - 23456 (représentent la gamme 32 bits dans le monde de 16-bit)
  - 65536-65551 (documentation – RFC5398)
  - 65552-4294967295 (Internet public)
- Représentation de la plage 32-bit range representation spécifié dans RFC5396
  - Définit le “asplain” (format traditionnel) en tant que notation standard

# Numéro de système autonome (NSA)

- NSA sont distribués par les registres Internet régionaux
  - Ils sont également disponibles auprès de fournisseurs de services Internet en amont qui sont membres de l'un des RIR
- Allocations actuelles du NSA 16 bits jusqu'à 61439 ont été apportées aux RIR
  - Autour de 41200 sont visibles sur Internet
- Chaque RIR a également reçu un bloc de 32-bit NSA
  - Sur les 2800 affectations, autour de 2400 sont visibles sur Internet
- Voir [www.iana.org/assignments/as-numbers](http://www.iana.org/assignments/as-numbers)

# AS-Privé – Application

- Applications

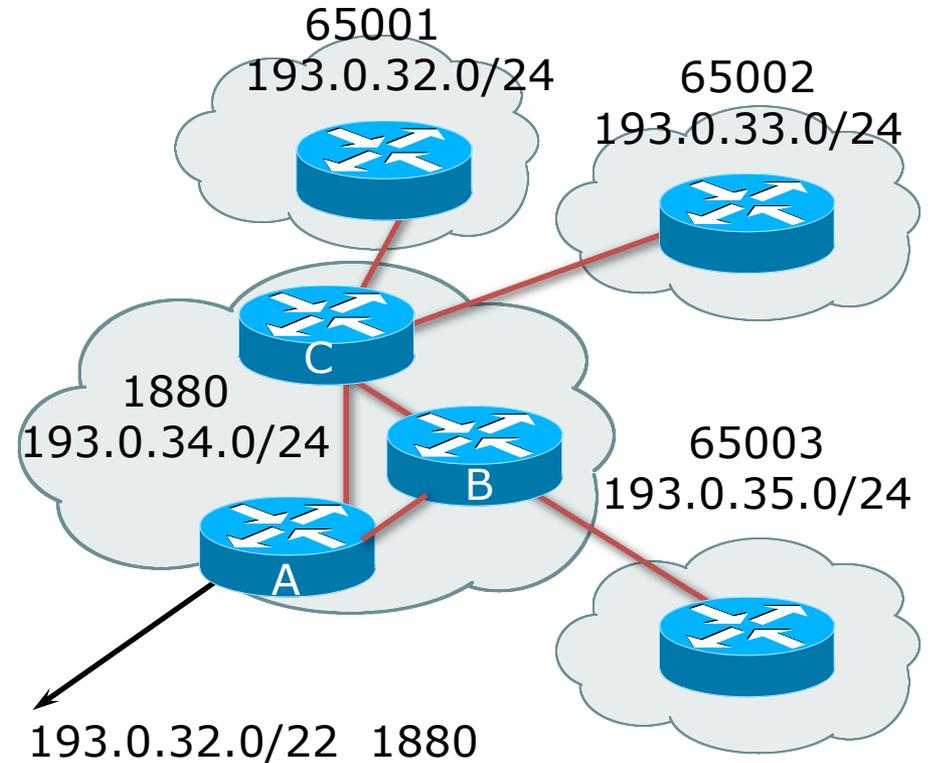
- Un FAI avec les clients multi-homé sur leur backbone (RFC2270)

-ou-

- Un réseau d'entreprise avec plusieurs régions mais également les connexions à l'Internet que dans le noyau

-ou-

- Au sein d'une confédération BGP



# AS-Privé – Suppression

- l'ASN privé DOIT être retiré de tous les préfixes annoncés à l'Internet public
  - Inclure la configuration pour supprimer les ASN privés dans le modèle eBGP
- Comme pour l'espace d'adressage RFC1918, les ASN privés sont destinés à un usage interne
  - Ils ne doivent pas être divulgués à l'Internet public
- Cisco IOS  
**neighbor x.x.x.x remove-private-AS**

# Transit/Appairage/Défaut

- **Transit**
  - Acheminement du trafic sur un réseau
  - Habituellement **pour un montant**
- **Peering**
  - Échanger des informations de routage de source locale et du trafic
  - Habituellement **sans frais**
- **Défaut**
  - Où envoyer le trafic quand il n'y a pas de correspondance explicite dans la table de routage

# Politique de configuration

- Hypothèses
  - préfix-lists sont employés partout
  - facile / mieux / plus rapide que des listes d'accès
- Trois principes DE BASE
  - **prefix-lists** pour filtrer **préfixes**
  - **filter-lists** pour filtrer **ASN**
  - **route-maps** pour appliquer **la politique**
- Route-Map peuvent être utilisées pour le filtrage, mais c'est une configuration plus "avancée"

# Outils de politique

- Local Pref
  - flux de trafic sortant
- Metric (MED)
  - flux de trafic entrant (champ d'application local)
- AS-PATH prepend
  - flux de trafic entrant (champ d'application Internet)
- communautés
  - informations spécifiques inter-fournisseur

# Génération des Préfixes: Hypothèses

- DOIT annoncer le bloc d'adresse attribuée à Internet
- Peut aussi annoncer des sous-préfixes -  
accessibilité n'est pas garantie
- L'allocation minimum actuel est de / 20 à / 24 en  
fonction de la RIR
  - Plusieurs FAI filtrent les blocs RIR sur cette frontière
  - Plusieurs FAI filtrent le reste de l'espace d'adressage  
selon les affectations d'IANA
  - Cette activité est appelée " Netiquette" par certains

# Génération des Préfixes

- Les RIR publient leurs tailles d'allocation minimum par bloc d'adresse de 8
  - AfriNIC: [www.afrinic.net/docs/policies/afpol-v4200407-000.htm](http://www.afrinic.net/docs/policies/afpol-v4200407-000.htm)
  - APNIC: [www.apnic.net/db/min-alloc.html](http://www.apnic.net/db/min-alloc.html)
  - ARIN: [www.arin.net/reference/ip\\_blocks.html](http://www.arin.net/reference/ip_blocks.html)
  - LACNIC: [lacnic.net/en/registro/index.html](http://lacnic.net/en/registro/index.html)
  - RIPE NCC: [www.ripe.net/ripe/docs/smallest-alloc-sizes.html](http://www.ripe.net/ripe/docs/smallest-alloc-sizes.html)
  - Notez que AfriNIC ne publie que sa taille actuelle d'allocation minimale, pas la taille d'allocation pour ses blocs d'adresses
- IANA publie l'espace d'adressage qu'il a assigné aux sites finals et a alloué aux RIR:
  - [www.iana.org/assignments/ipv4-address-space](http://www.iana.org/assignments/ipv4-address-space)
- Plusieurs FAI utilisent cette information publiée à filtrer les préfixes:
  - Ce qui doit être routé (de l'IANA)
  - La taille d'allocation minimale par les RIR

# “Netiquette” liste de préfixe

- Destiné à “punir” les FAI qui polluent la table de routage avec des détails plutôt que d’annoncer les agrégats
- Les impacts légitiment multi-homing particulièrement au bord de l'Internet
- Impact sur les régions où le backbone n'est pas disponible ou coûte \$\$\$, par rapport à la bande passante internationale
- Difficile à maintenir – nécessite une mise à jour lorsque les RIR commencent à allouer de nouveaux blocs d'adresses
- Ne le faites pas à moins que vous compreniez les conséquences et que vous êtes prêt à maintenir à jour la liste
  - Envisagez d'employer Team Cymru ou autre réputé bogon BGP flux de données:
  - [www.team-cymru.org/Services/Bogons/routeserver.html](http://www.team-cymru.org/Services/Bogons/routeserver.html)

# Comment multi-homer

Quelques choix...

# Transits

- Le Fournisseur de transit est un autre système autonome qui est utilisé pour fournir l'accès à d'autres réseaux
  - Peut être seulement locale ou régionale
  - Mais plus habituellement tout l'Internet
- Les fournisseurs de transit doivent être choisis judicieusement:
  - Un seulement
    - pas de redondance
  - Plusieurs
    - plus difficile d'équilibrer la charge
    - aucune économie d'échelle (coûts plus par Mbit/s)
    - difficile de fournir une qualité de service
- **Recommandation: au moins deux, pas plus de trois**

# Erreurs courantes

- FAI s'inscrivent avec plusieurs fournisseurs de transit
  - Beaucoup de petits circuits (coût plus par Mbits/s que les grands)
  - Les prix de transit par Mbit/s diminuent avec l'augmentation de la bande passante achetée
  - Difficile de mettre en œuvre de l'ingénierie de trafic fiables qui n'a pas besoin d'ajustement fin quotidien en fonction des activités des clients
- Pas de diversité
  - Les fournisseurs de transit choisis sont tous atteints avec le même satellite ou le même câble sous-marin
  - Les fournisseurs de transit choisis ont un transit ultérieur ou peu de peering

# Pairs

- Un pair est un autre système autonome avec lequel le réseau local a convenu d'échanger les routes et le trafic d'origine locale
- Pair privé
  - Liaison privée entre deux fournisseurs afin de l'interconnexion
- Pair public
  - Point d'échange Internet, où les fournisseurs se rencontrent et décident librement avec qui ils vont s'interconnecter
- **Recommandation: les pairs, autant que possible!**

# Erreurs courantes

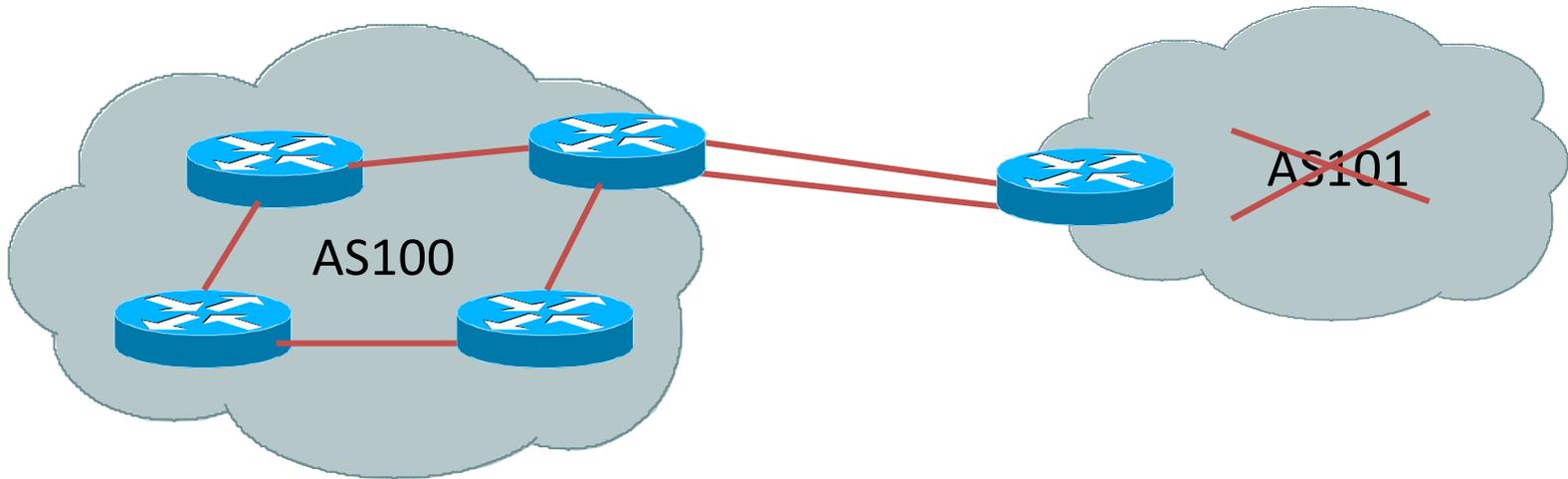
- Confondre un fournisseur de transit et un point de peering public sans frais
- Ne travaillant assez pour obtenir autant de peering que possible
  - Physiquement, près d'un point d'appairage (IXP) mais n'est pas présent à cet IXP
  - (Transit est parfois moins cher que peering!!)
- Ignorer/ éviter les concurrents

Potentiellement de précieux partenaire de peering pour donner aux clients une meilleure expérience

# Scénarios de Multi-homing

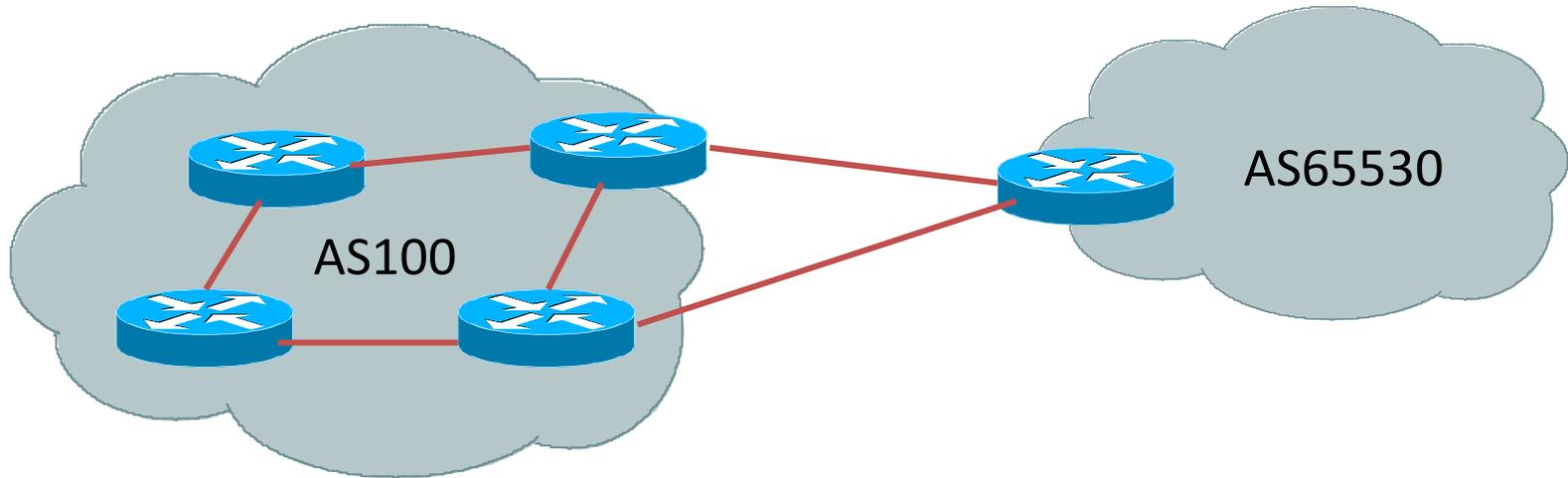
- Stub network
- Multi-homed stub network
- Multi-homed network
- Sessions Multiples vers d'autres AS

# Stub Network



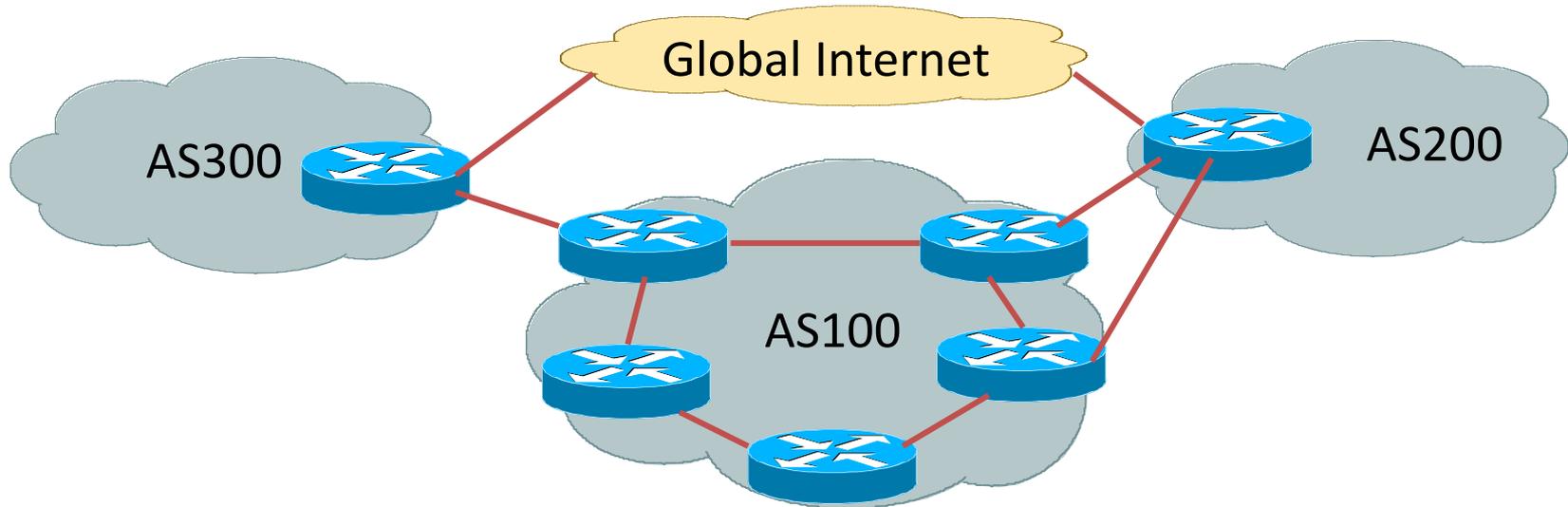
- Pas besoin de BGP
- Pointer la route statique par défaut vers le FAI
- Le FAI en amont fait l'annonce du réseau
- Politique confinée à l'intérieur des politiques du FAI en amont

# Multi-homed stub network



- Utiliser BGP (non IGP ou statique) pour l'annonce
- Utiliser un AS privé (ASN > 64511)
- Le FAI en amont fait l'annonce sur Internet
- Politique confinée à l'intérieur des politiques du FAI en amont

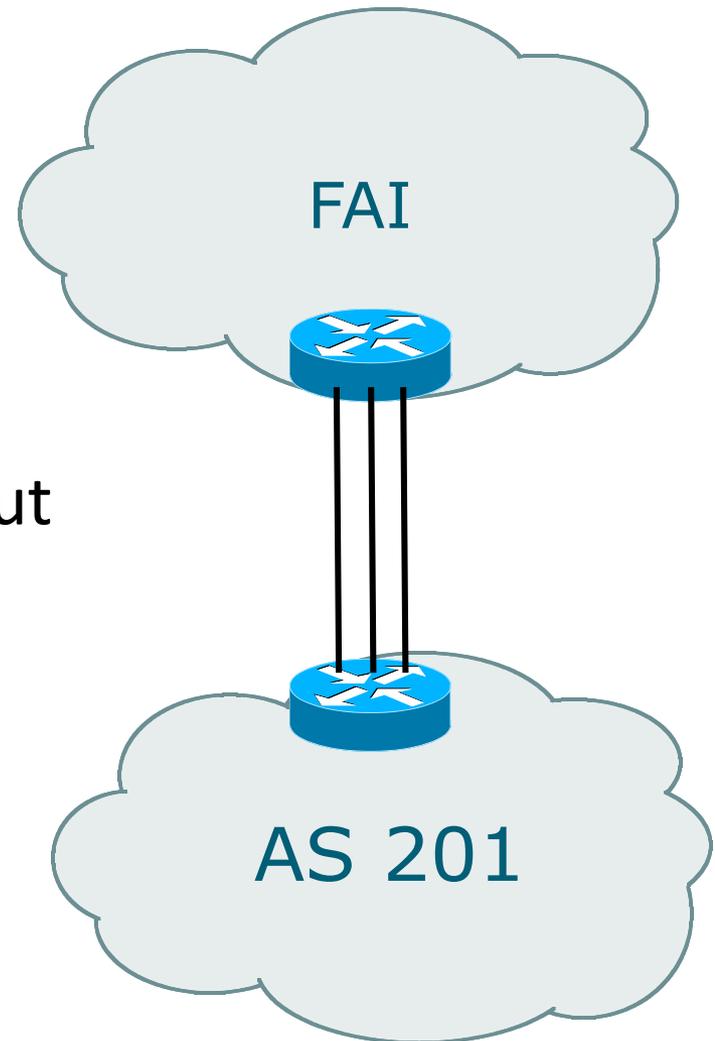
# Multi-homed network



- Plusieurs situations possibles
  - plusieurs sessions au même FAI
  - une sessions secondaire de backup uniquement
  - load balancing entre primaire et secondaire
  - utiliser sélectivement les différents FAI

# plusieurs sessions à un FAI

- Plusieurs options
  - ebgp à sauts multiples
  - bgp trajets multiples
  - cef partage de charge
  - bgp manipulation d'attribut



# Plusieurs Sessions à un AS

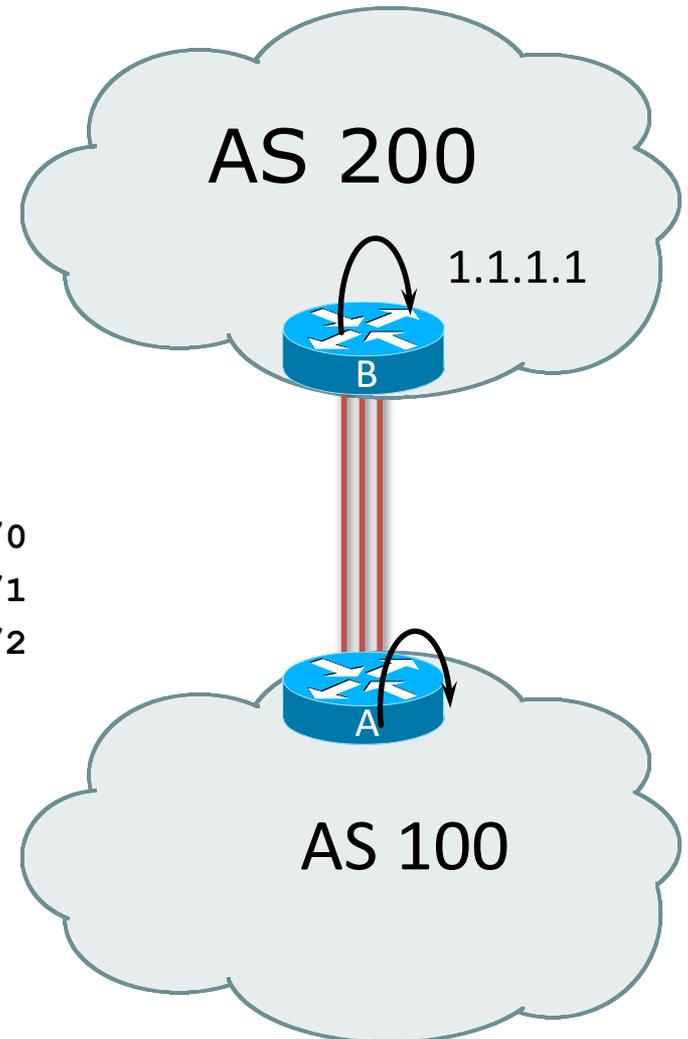
## – eBGP à sauts multiples

- Utiliser eBGP à sauts multiples
  - Exécuter eBGP entre les adresses de loopback
  - Les préfixes eBGP ont appris avec l'adresse de loopback comme le prochain saut

- Cisco IOS

```
router bgp 100m
  neighbor 1.1.1.1 remote-as 200
  neighbor 1.1.1.1 ebgp-multihop 2
  !
  ip route 1.1.1.1 255.255.255.255 serial 1/0
  ip route 1.1.1.1 255.255.255.255 serial 1/1
  ip route 1.1.1.1 255.255.255.255 serial 1/2
```

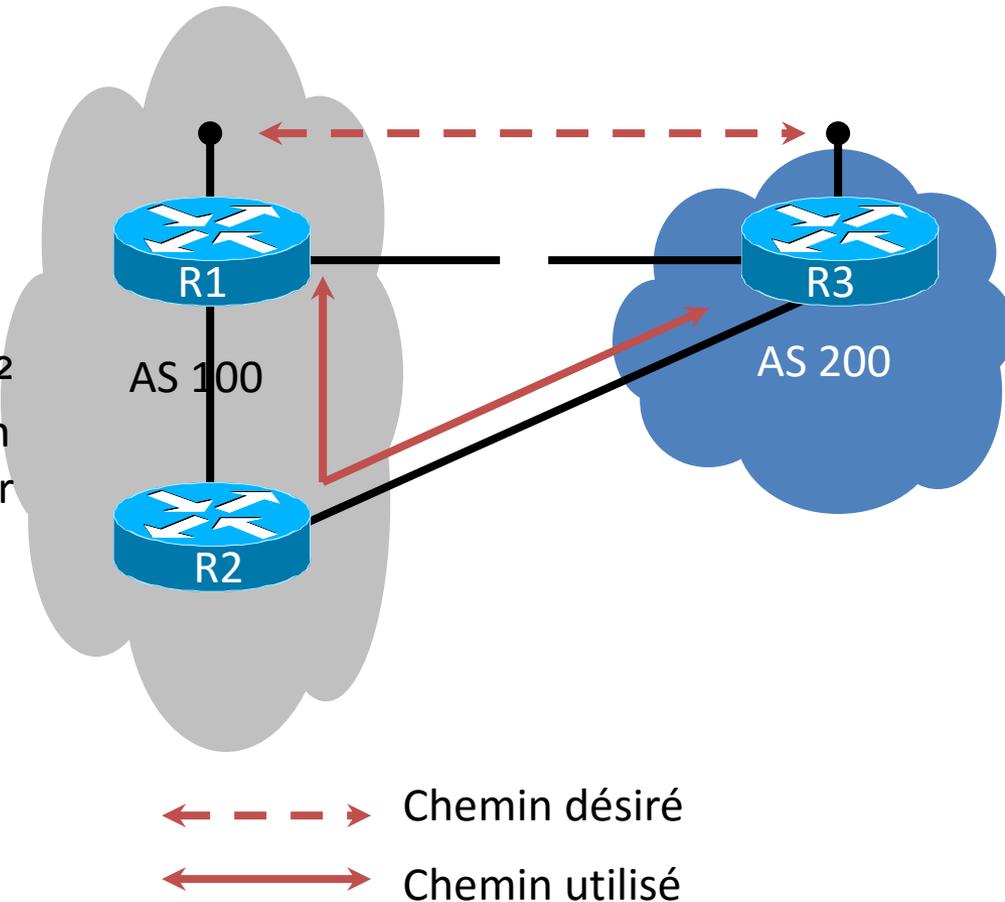
- Erreur fréquente consiste à pointer la route vers l'adresse IP distante plutôt que lien spécifique



# Plusieurs Sessions à un AS

## – ebgp à sauts multiples

- Une sérieuse mise en garde de eBGP-à sauts multiples:
  - R1 et R3 sont des pairs eBGP qui sont des bouclages d'appairage
  - configuré avec:  
`neighbor x.x.x.x ebgp-multihop 2`
  - Si le lien entre R1 et R3 tombe en panne la session pourrait s'établir par l'intermédiaire de R2
- Se produit généralement lors du routage de bouclage distant est dynamique plutôt que statique pointant sur un lien



# Plusieurs Sessions à un FAI

## – ebgp à sauts multiples

- Essayez d'éviter l'utilisation de eBGP-à sauts multiples, sauf si:
  - Il est absolument nécessaire –ou–
  - Partage de charge entre plusieurs liens
- De nombreux FAI découragent son utilisation, par exemple:

Nous allons exécuter eBGP à sauts multiples, mais ne prend pas en charge comme l'offre standard parce que les clients ont généralement du mal à le gérer en raison de:

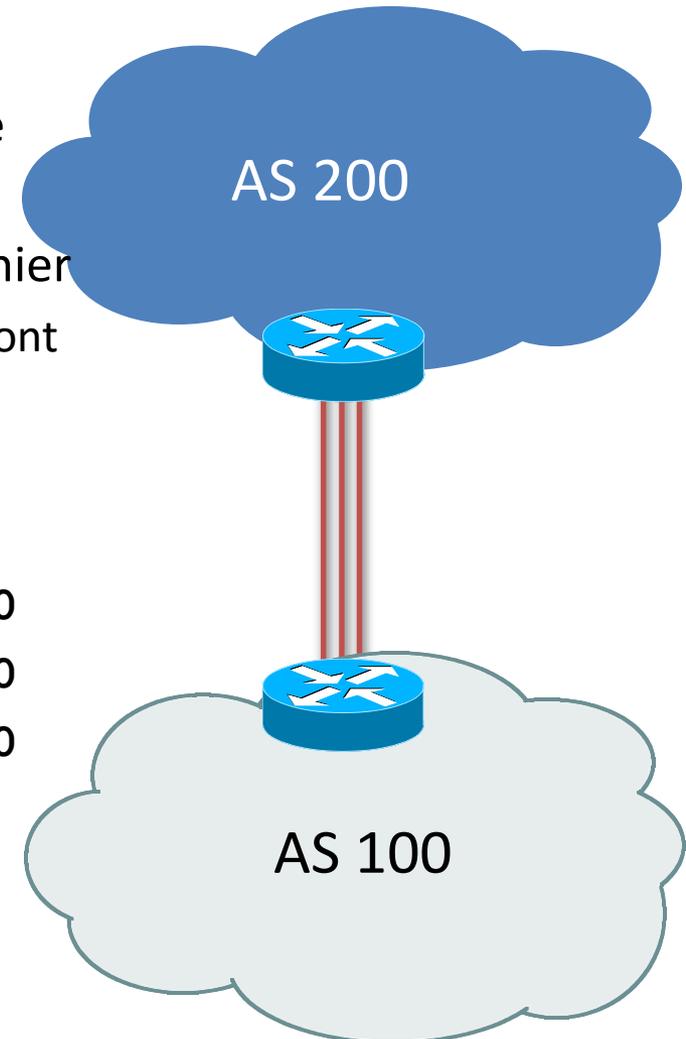
- boucles de routage
- l'impossibilité de réaliser que les problèmes de stabilité de session BGP sont généralement des problèmes de connectivité en raison entre leur CPE et leur interlocuteur BGP

# Plusieurs Sessions à un AS

## – ebgp à trajets multiples

- Trois sessions BGP exigés
- Limite de la plate-forme sur le nombre de chemins (peut être aussi peu que 6)
- Flux BGP complet rend cela difficile à manier
  - 3 copies de la Table de routage Internet vont dans le FIB

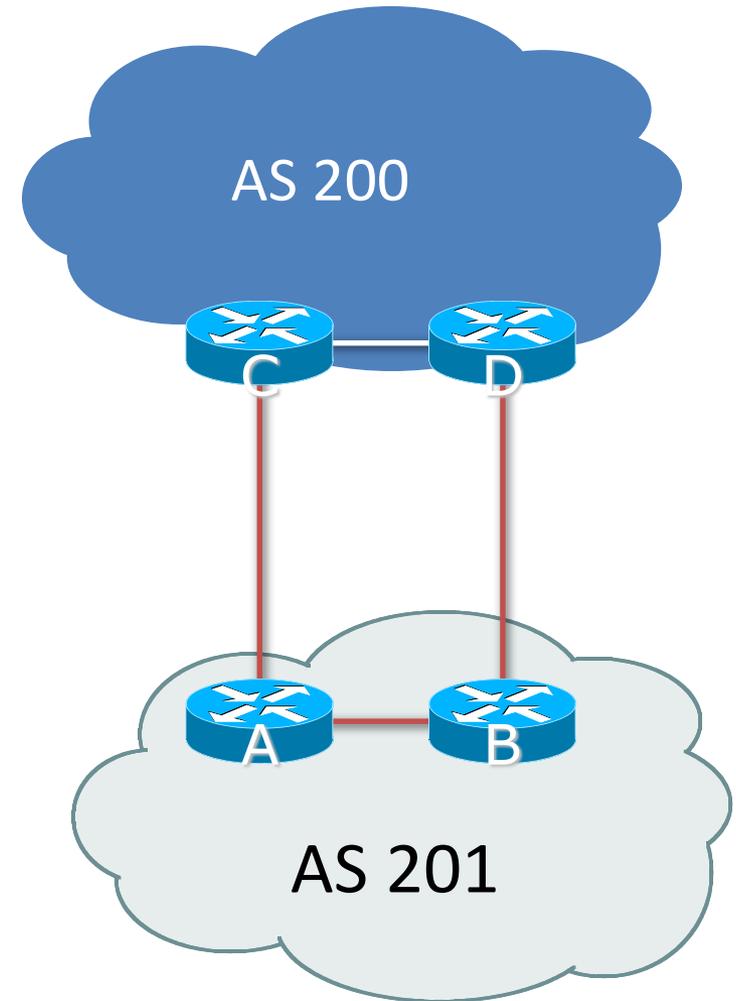
```
router bgp 100m
  neighbor 1.1.2.1 remote-as 200
  neighbor 1.1.2.5 remote-as 200
  neighbor 1.1.2.9 remote-as 200
  maximum-paths 3
```



# Plusieurs Sessions à un AS

## – bgp attributs & filtres

- Le plus simple schéma est d'utiliser les défauts
- Apprendre / Annoncez préfixes pour un meilleur contrôle
- Planification et certains travaux sont requis pour réaliser le partage de charge
  - Point par défaut vers un FAI
  - Apprendre les préfixes sélectionnés du second ISP
  - Modifier le nombre de préfixes appris à réaliser le partage de charge acceptable
- Pas de solution magique



# Principes de base du multi-homing

Apprenons à marcher avant que nous  
essayions de courir...

# Les Principes de base

- Annoncer l'espace d'adressage attire le trafic
  - (À moins que la politique des fournisseurs en amont interfère)
- En annonçant l'agrégat du FAI sur un lien se traduira par un trafic de cet agrégat à venir dans ce lien
- Annonçant un lien d'un sous-préfixe d'un agrégat signifie que tout le trafic pour ce sous-préfixe viendra dans ce lien, même si l'agrégat est annoncé ailleurs
  - **L'annonce la plus spécifique gagne!**

# Les Principes de base

- Pour séparer le trafic entre deux liens:
  - Annoncer l'agrégat sur les deux liens - assure la redondance
  - Annoncer une moitié de l'espace d'adressage sur chaque lien
  - (C'est la première étape, toutes choses étant égales)
- Résulte en:
  - Trafic pour la première moitié de l'espace d'adressage vient en premier lien
  - Trafic pour la deuxième moitié de l'espace d'adressage est disponible en deuxième lien
  - Si l'un ou l'autre de lien échoue, le fait que l'agrégat est annoncé s'assure qu'il y a un chemin de secours

# Les Principes de base

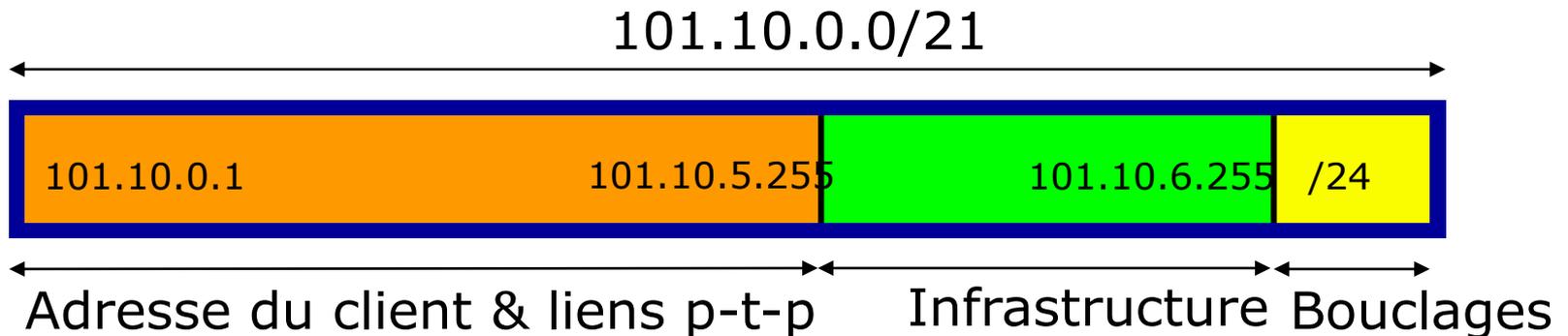
- Les clés de la réussite de configuration multihoming:
  - Garder les annonces de préfixe d'ingénierie du trafic indépendants du client iBGP
  - Comprendre comment annoncer les agrégats
  - Comprendre le but de l'annonce des sous-préfixes des agrégats
  - Comprendre comment manipuler les attributs BGP
  - Trop de chemins en amont / externes rendent plus difficile le multi-homing (2 ou 3 est assez!)

# Adressage IP & Multi-homing

Comment les plans d'IP adresse  
assistent le multi-homing

# Adressage IP & Multi-homing

- Planification de l'adresse IP est une partie importante du Multi-homing
- Précédemment nous avons discuté de la séparation de:
  - Espace d'adressage du client
  - Espace d'adressage du client p-t-p
  - Espace d'adressage de lien d'infrastructure p-t-p
  - Espace d'adressage de bouclage

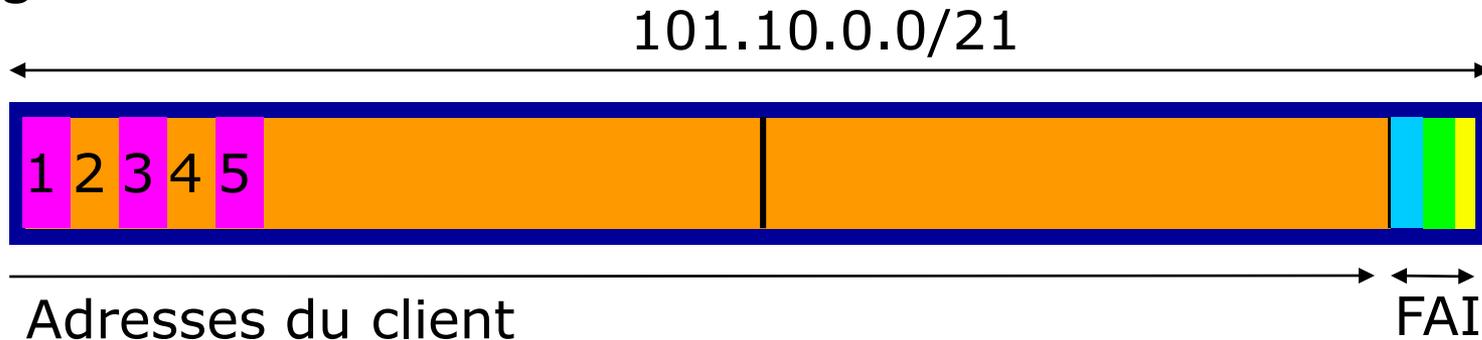


# Adressage IP & Multi-homing

- Les loopback du routeur du FAI et les liens point à point du backbone ne constituent qu'une petite partie de l'espace d'adressage totale
  - Et ils n'attirent pas le trafic, contrairement à l'espace d'adresse du client
- Liens du bord de l'agrégation du FAI au routeur du client ont besoin d'un /30
  - Faibles exigences par rapport à l'espace d'adressage totale
  - Certains FAI utilisent de nombreuses IP
- Planification des affectations du client est une partie très importante de multi-homing
  - L'ingénierie du trafic implique de subdiviser l'agrégat en morceaux jusqu'à ce que l'équilibrage de charge fonctionne

# L'adressage IP non planifié

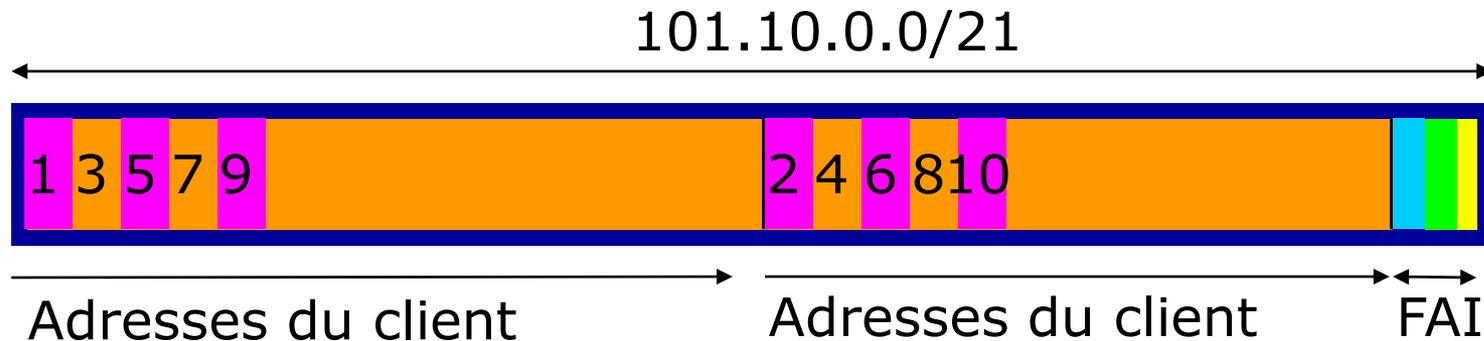
- FAI remplit les IP des clients à partir d'une extrémité de la gamme:



- Les clients génèrent un trafic
  - Diviser la plage en deux morceaux se traduira par un /22 avec tous les clients, et un /22 avec seulement les adresses de l'infrastructure du FAI
  - Aucun équilibrage de charge : tout le trafic ne viendra que du premier /22
  - subdivision supplémentaire du premier /22 = plus de travail

# L'adressage IP planifié

- Si FAI assigne les clients à partir des deux extrémités de la gamme:



- Le schéma est alors:
  - Premier client du premier /22, le deuxième client du deuxième /22, le troisième à partir du premier /22, etc
- Cela fonctionne également pour du résidentiel par rapport aux clients commerciaux:
  - Résidentiel du premier /22
  - Commercial du deuxième /22

# L'adressage IP planifié

- Cela fonctionne bien pour le multi-homing entre deux liens en amont (fournisseurs identiques ou différents)
- Peut également subdiviser l'espace d'adressage pour adapter à plus de deux opérateurs en amont
  - Suivre un schéma similaire pour le remplissage de chaque portion de l'espace d'adressage
- Ne pas oublier de toujours annoncer un agrégat de chaque lien

# Multi-homing de base

Essayons quelques exemples  
simples...

# Multi-homing de base

- Pas de superflu multi-homing
- Regardons deux cas:
  - Multi-homing avec le même FAI
  - Multi-homing avec plusieurs FAI
- Gardons les exemples faciles
  - Comprendre les concepts faciles rendra les scénarios plus complexes plus facile à comprendre
  - Supposer que le site multi-homing a un bloc d'adresse de /19

# Multi-homing de base

- Ce type est le lieu le plus commun à l'avant garde de l'Internet
  - Ici les réseaux ici sont généralement intéressés par les flux de trafic entrants
  - Les flux de trafic sortant étant “le plus proche de la sortie” suffit généralement
- Peut s'adresser au FAI final ainsi que les réseaux d'entreprise

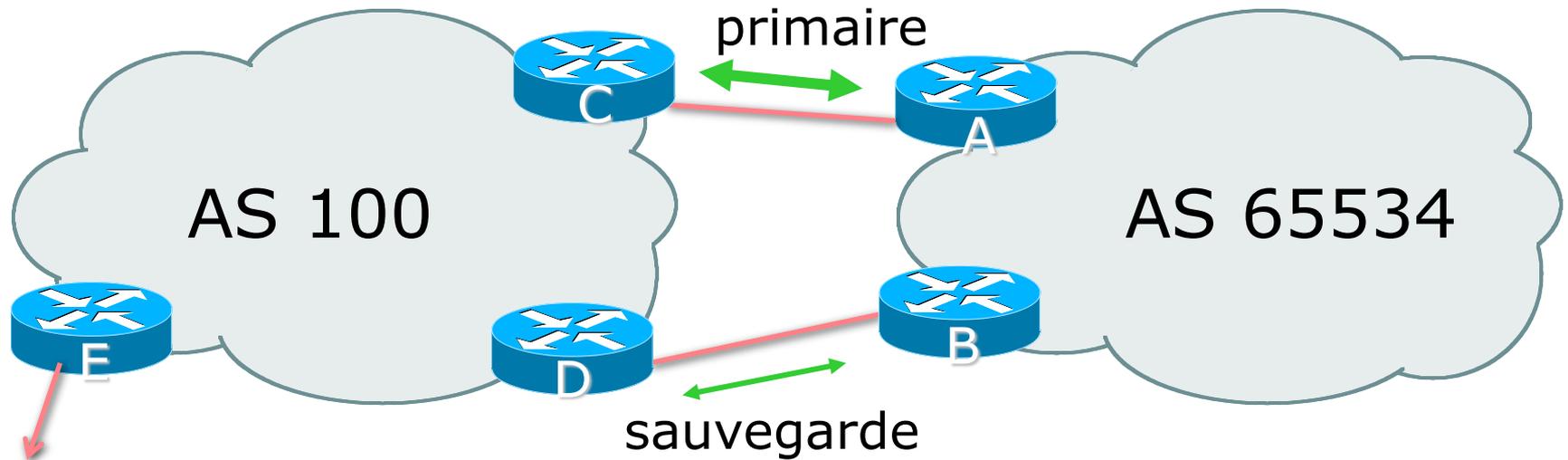
# deux liens vers le même FAI

Un lien primaire, l'autre lien de sauvegarde uniquement

# Deux liens vers le même FAI (dont un de secours seulement)

- S'applique lorsque le site final a acheté un grand nombre de liens RGD primaires à leur opérateur en amont et un petit nombre de liens RGD secondaires comme sauvegarde
  - Par exemple, chemin principal pourrait être un E1, la sauvegarde pourrait être 64kbps

# Deux liens vers le même FAI (dont un de secours seulement)



- AS100 supprime AS privé et l'annonce des sous-préfixes des clients de d'Internet

# Deux liens vers le même FAI (dont un de secours seulement)

- Annoncer l'agrégat / 19 sur chaque lien
  - lien primaire:
    - Sortants – annonce /19 inchangée
    - Entrant – reçoit route par défaut
  - lien de sauvegarde:
    - Sortants – annonce /19 dont la métrique est accrue
    - Entrants – reçu par défaut et réduit la préférence locale
- Quand un lien tombe, l'annonce de l'agrégat / 19 via l'autre lien assure une connectivité continue

# Deux liens vers le même FAI (dont un de secours seulement)

- Routeur A Configuration

```
router bgp 65534
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.2 remote-as 100
```

```
neighbor 122.102.10.2 description RouterC
```

```
neighbor 122.102.10.2 prefix-list aggregate out
```

```
neighbor 122.102.10.2 prefix-list default in
```

```
!
```

```
ip prefix-list aggregate 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 liens vers le même FAI (dont un de secours seulement)

- Routeur B Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.6 remote-as 100
  neighbor 122.102.10.6 description RouterD
  neighbor 122.102.10.6 prefix-list aggregate out
  neighbor 122.102.10.6 route-map routerD-out out
  neighbor 122.102.10.6 prefix-list default in
  neighbor 122.102.10.6 route-map routerD-in in
```

!

..diapositive suivante

# Deux liens vers le même FAI (dont un de secours seulement)

```
ip prefix-list aggregate 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
!
route-map routerD-out permit 10
  set metric 10
!
route-map routerD-in permit 10
  set local-preference 90
!
```

# Deux liens vers le même FAI (dont un de secours seulement)

- Routeur C Configuration (lien principal)

```
router bgp 100m
```

```
neighbor 122.102.10.1 remote-as 65534
```

```
neighbor 122.102.10.1 default-originate
```

```
neighbor 122.102.10.1 prefix-list Customer in
```

```
neighbor 122.102.10.1 prefix-list default out
```

```
!
```

```
ip prefix-list Customer permit 121.10.0.0/19
```

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

# Deux liens vers le même FAI (dont un de secours seulement)

- Routeur D Configuration (lien de secours)

```
router bgp 100m
  neighbor 122.102.10.5 remote-as 65534
  neighbor 122.102.10.5 default-originate
  neighbor 122.102.10.5 prefix-list Customer in
  neighbor 122.102.10.5 prefix-list default out
!
ip prefix-list Customer permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

# Deux liens vers le même FAI (dont un de secours seulement)

- Routeur E Configuration

```
router bgp 100m
```

```
neighbor 122.102.10.17 remote-as 110
```

```
neighbor 122.102.10.17 remove-private-AS
```

```
neighbor 122.102.10.17 prefix-list Customer out
```

```
!
```

```
ip prefix-list Customer permit 121.10.0.0/19
```

- Routeur E supprime l'AS privé et les sous-préfixes du client des annonces externes
- AS privé toujours visible à l'intérieur de AS100

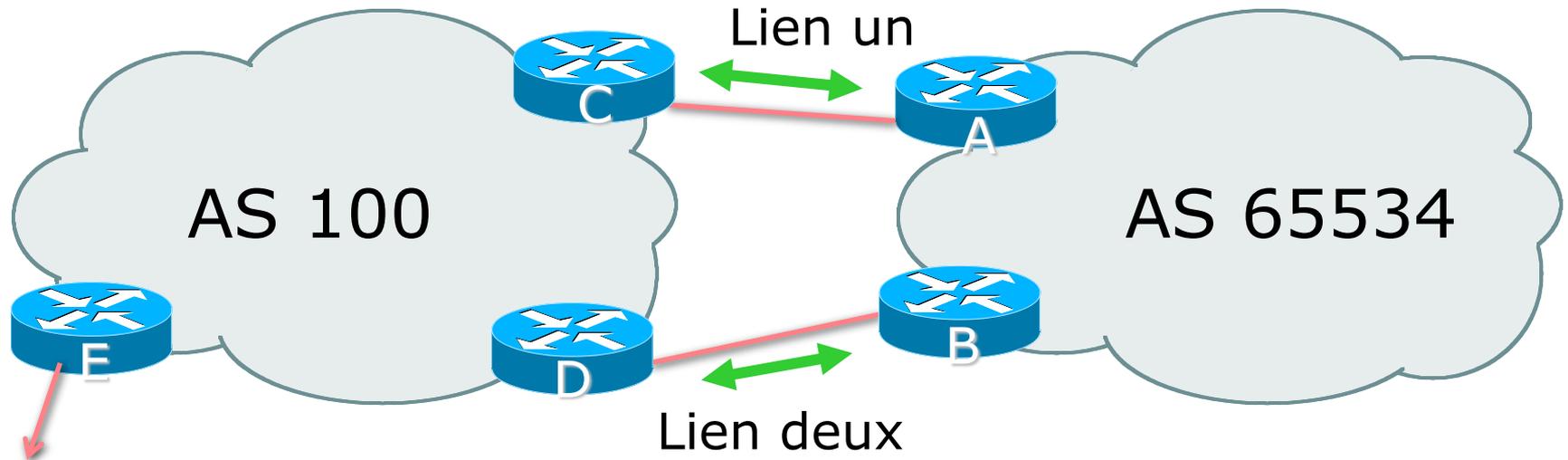
deux liens vers le même FAI

Avec le partage de charge

# Le partage de charge avec le même FAI

- Cas les plus courants
- Sites finaux tendent à acheter des circuits et de les laisser inactif, seulement utilisé pour la sauvegarde, comme dans l'exemple précédent
- Cet exemple suppose des circuits à capacité égale
  - Circuits à capacité inégales nécessite plus de raffinement - voir plus loin

# Le partage de charge avec le même FAI



- Le routeur E de frontière dans AS100 enlève l'AS privé et les sous-préfixes de tous les clients de l'annonce d'Internet

# Le partage de charge avec le même FAI (avec redondance)

- Annoncer / 19 agrégat sur chaque lien
- Diviser /19 et annoncer comme deux / 20s, un sur chaque lien
  - partage de charge de base entrant
  - suppose la capacité de circuit égale et même la propagation du trafic à travers le bloc d'adresses
- Varier le fractionnement jusqu'à " parfait " partage de la charge atteint
- Acceptez la valeur par défaut de l'opérateur en amont
  - base de partage de charge sortant par la sortie la plus proche
  - Ok dans le premier approximatif comme la plupart des FAI et site final trafic est entrant

# Le partage de charge avec le même FAI (avec redondance)

- Routeur A Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.0.0 mask 255.255.240.0
  neighbor 122.102.10.2 remote-as 100
  neighbor 122.102.10.2 prefix-list routerC out
  neighbor 122.102.10.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerC permit 121.10.0.0/20
ip prefix-list routerC permit 121.10.0.0/19
!
ip route 121.10.0.0 255.255.240.0 null0
ip route 121.10.0.0 255.255.224.0 null0
```

# Le partage de charge avec le même FAI (avec redondance)

- Routeur B Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.16.0 mask 255.255.240.0
  neighbor 122.102.10.6 remote-as 100
  neighbor 122.102.10.6 prefix-list routerD out
  neighbor 122.102.10.6 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerD permit 121.10.16.0/20
ip prefix-list routerD permit 121.10.0.0/19
!
ip route 121.10.16.0 255.255.240.0 null0
ip route 121.10.0.0 255.255.224.0 null0
```

# Le partage de charge avec le même FAI (avec redondance)

- Routeur C Configuration

```
router bgp 100m
```

```
neighbor 122.102.10.1 remote-as 65534
```

```
neighbor 122.102.10.1 default-originate
```

```
neighbor 122.102.10.1 prefix-list Customer in
```

```
neighbor 122.102.10.1 prefix-list default out
```

```
!
```

```
ip prefix-list Customer permit 121.10.0.0/19 le 20
```

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

- Le routeur C permet seulement en /19 et /20 préfixes du bloc client
- La configuration du routeur D est identique

# Le partage de charge avec le même FAI (avec redondance)

- Routeur E Configuration

```
router bgp 100m
```

```
neighbor 122.102.10.17 remote-as 110
```

```
neighbor 122.102.10.17 remove-private-AS
```

```
neighbor 122.102.10.17 prefix-list Customer out
```

```
!
```

```
ip prefix-list Customer permit 121.10.0.0/19
```

- AS privé toujours visible à l'intérieur de AS100

# Le partage de charge avec le même FAI (avec redondance)

- Route par défaut pour le trafic sortant?
  - Utiliser les informations par défaut qui proviennent de l'IGP et qui dépendent sur des métriques IGP pour la sortie la plus proche
  - e.g. sur routeur A:

```
router ospf 65534
```

```
default-information originate metric 2 metric-type 1
```

# Le partage de charge avec le même FAI (avec redondance)

- Configuration de partage de charge n'est que sur le routeur du client
- FAI En amont doit
  - retirer les sous-préfixes des clients des annonces externes
  - retirer AS privé des annonces externes
- Pourrait aussi utiliser les communautés BGP

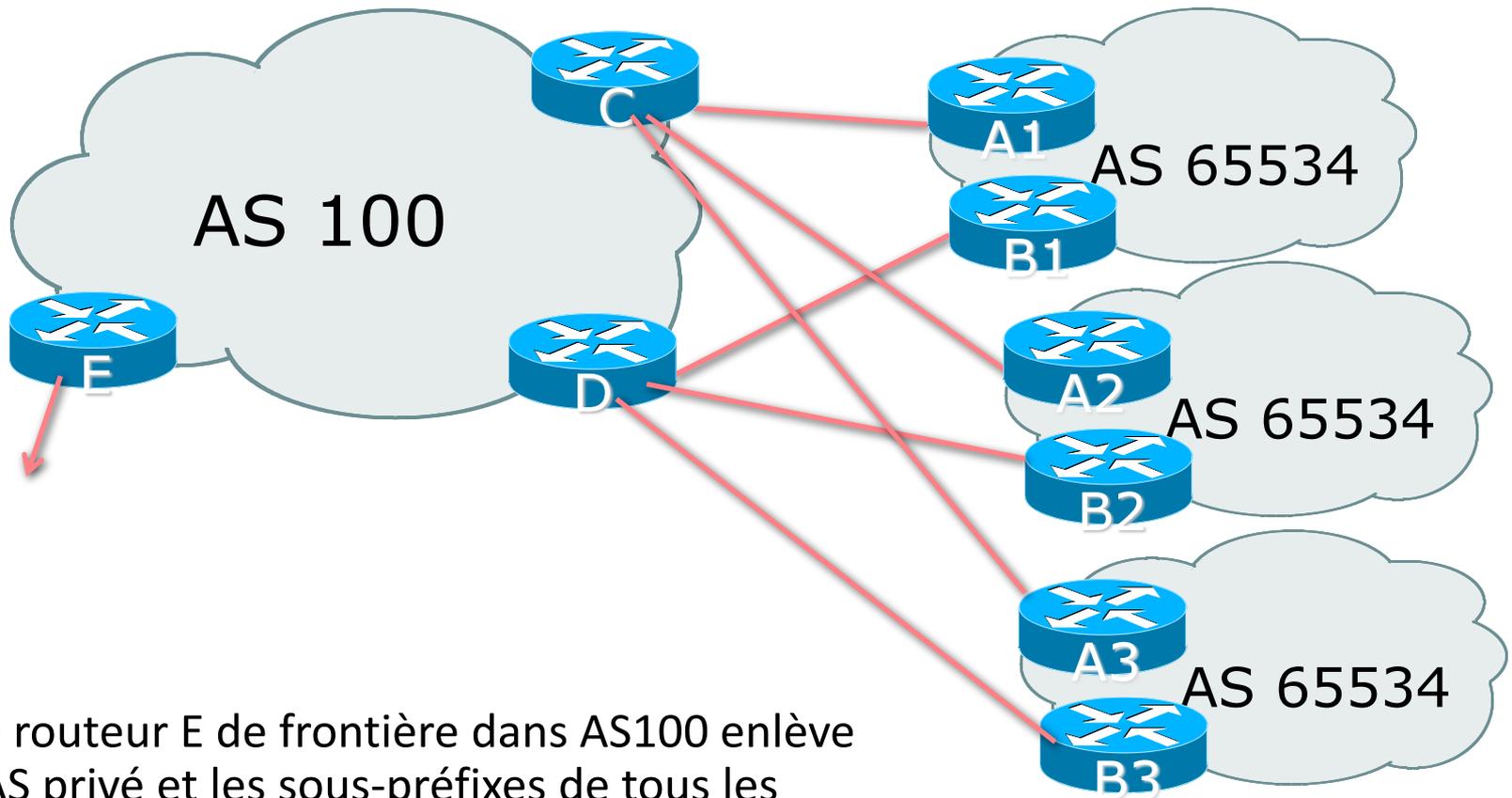
# Deux liens vers le même FAI

Plusieurs clients multi-homés  
(RFC2270)

# Plusieurs clients dual-homed (RFC2270)

- Inhabituel pour un FAI juste pour avoir un client multi-homé
  - Offre de service valide / valable pour un FAI avec plusieurs points de présence
  - Mieux pour les FAI que d'avoir multi-home client chez un autre fournisseur!
- Regarder à la réduction de la configuration
  - $\Rightarrow$  Simplification de la configuration
  - Utilisation des modèles, des pairs-groupes, etc
  - Chaque client possède la même configuration (essentiellement)

# Plusieurs clients dual-homeds (RFC2270)



- Le routeur E de frontière dans AS100 enlève l'AS privé et les sous-préfixes de tous les clients de l'annonce d'Internet

# Plusieurs clients dual-homeds (RFC2270)

- Annonces de client selon l'exemple précédent
- Utiliser le même AS privé pour chaque client
  - documenté dans RFC2270
  - l'espace d'adressage ne se chevauchent pas
  - chaque client entend le défaut uniquement
- Routeurs An et Bn même configuration que les routeurs A et B précédemment

# Plusieurs clients dual-homed (RFC2270)

- Routeur A1 Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.0.0 mask 255.255.240.0
  neighbor 122.102.10.2 remote-as 100
  neighbor 122.102.10.2 prefix-list routerC out
  neighbor 122.102.10.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerC permit 121.10.0.0/20
ip prefix-list routerC permit 121.10.0.0/19
!
ip route 121.10.0.0 255.255.240.0 null0
ip route 121.10.0.0 255.255.224.0 null0
```

# Plusieurs clients dual-homed (RFC2270)

- Routeur B1 Configuration

```
router bgp 65534
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.16.0 mask 255.255.240.0
  neighbor 122.102.10.6 remote-as 100
  neighbor 122.102.10.6 prefix-list routerD out
  neighbor 122.102.10.6 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list routerD permit 121.10.16.0/20
ip prefix-list routerD permit 121.10.0.0/19
!
ip route 121.10.0.0 255.255.224.0 null0
ip route 121.10.16.0 255.255.240.0 null0
```

# Plusieurs clients dual-homeds (RFC2270)

- Routeur C Configuration

```
router bgp 100m
```

```
neighbor bgp-customers peer-group
```

```
neighbor bgp-customers remote-as 65534
```

```
neighbor bgp-customers default-originate
```

```
neighbor bgp-customers prefix-list default out
```

```
neighbor 122.102.10.1 peer-group bgp-customers
```

```
neighbor 122.102.10.1 description Customer One
```

```
neighbor 122.102.10.1 prefix-list Customer1 in
```

```
neighbor 122.102.10.9 peer-group bgp-customers
```

```
neighbor 122.102.10.9 description Customer Two
```

```
neighbor 122.102.10.9 prefix-list Customer2 in
```

# Plusieurs clients dual-homed (RFC2270)

```
neighbor 122.102.10.17 peer-group bgp-customers
neighbor 122.102.10.17 description Customer Three
neighbor 122.102.10.17 prefix-list Customer3 in
!
ip prefix-list Customer1 permit 121.10.0.0/19 le 20
ip prefix-list Customer2 permit 121.16.64.0/19 le 20
ip prefix-list Customer3 permit 121.14.192.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

- Le routeur C permet seulement en /19 et /20 préfixes du bloc client

# Plusieurs clients dual-homed (RFC2270)

- Routeur D Configuration

```
router bgp 100m
```

```
neighbor bgp-customers peer-group
```

```
neighbor bgp-customers remote-as 65534
```

```
neighbor bgp-customers default-originate
```

```
neighbor bgp-customers prefix-list default out
```

```
neighbor 122.102.10.5 peer-group bgp-customers
```

```
neighbor 122.102.10.5 description Customer One
```

```
neighbor 122.102.10.5 prefix-list Customer1 in
```

```
neighbor 122.102.10.13 peer-group bgp-customers
```

```
neighbor 122.102.10.13 description Customer Two
```

```
neighbor 122.102.10.13 prefix-list Customer2 in
```

# Plusieurs clients dual-homed (RFC2270)

```
neighbor 122.102.10.21 peer-group bgp-customers
neighbor 122.102.10.21 description Customer Three
neighbor 122.102.10.21 prefix-list Customer3 in
!
ip prefix-list Customer1 permit 121.10.0.0/19 le 20
ip prefix-list Customer2 permit 121.16.64.0/19 le 20
ip prefix-list Customer3 permit 121.14.192.0/19 le 20
ip prefix-list default permit 0.0.0.0/0
```

- Le routeur D permet seulement en /19 et /20 préfixes du bloc client

# Plusieurs clients dual-homed (RFC2270)

- Routeur E Configuration

- suppose que l'espace d'adressage du client ne fait pas parti du bloc d'adresse de l'opérateur en amont

```
router bgp 100m
```

```
neighbor 122.102.10.17 remote-as 110
```

```
neighbor 122.102.10.17 remove-private-AS
```

```
neighbor 122.102.10.17 prefix-list Customers out
```

```
!
```

```
ip prefix-list Customers permit 121.10.0.0/19
```

```
ip prefix-list Customers permit 121.16.64.0/19
```

```
ip prefix-list Customers permit 121.14.192.0/19
```

- AS privé toujours visible à l'intérieur de AS100

# Plusieurs clients dual-homed (RFC2270)

- Si les préfixes des clients proviennent de bloc d'adresse du FAI
  - Ne **PAS** les annoncer sur Internet
  - annoncer l'agrégat du FAI seulement
- Routeur E configuration:

```
router bgp 100m
  neighbor 122.102.10.17 remote-as 110
  neighbor 122.102.10.17 prefix-list my-aggregate out
!
ip prefix-list my-aggregate permit 121.8.0.0/13
```

# Résumé multi-homing

- Utiliser AS privé pour multi-homing à l'opérateur en amont même
- Divulguer les sous-préfixes vers l'opérateur en amont que pour faciliter du partage de charge
- La configuration du routeur E en amont est identique dans toutes les situations

# Multi-homing de base

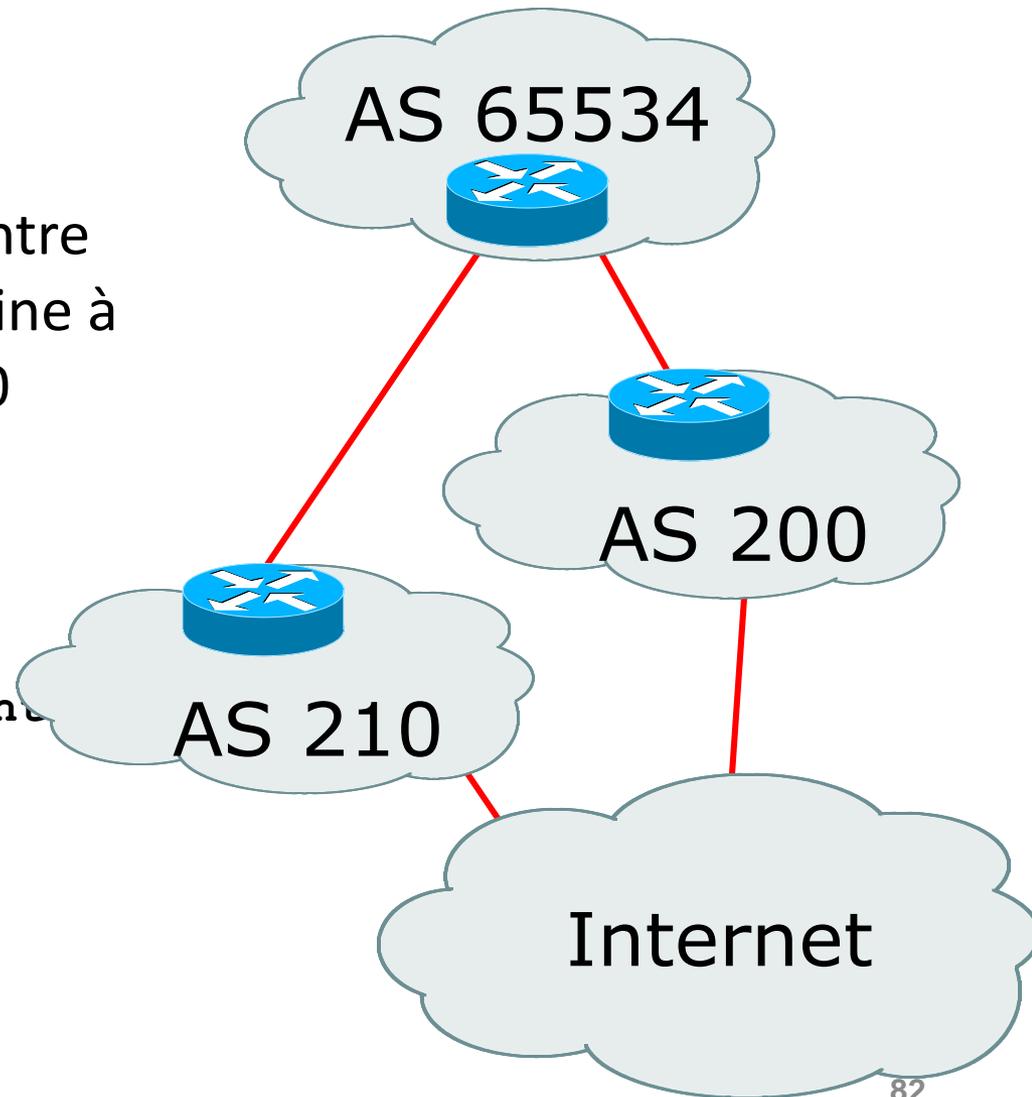
Multi-homing aux différents FAI

# deux liens vers différents FAI

- Utiliser un AS public
  - Ou utiliser un AS privé en cas d'accord avec d'autres FAI
  - Mais certaines personnes n'aiment pas le «incompatible-AS», qui résulte de l'utilisation d'un AS privé
- L'espace d'adressage vient de
  - Tous deux en amont ou
  - Registre Internet Régional
- Concepts de configuration très similaires

# Incompatibilité AS?

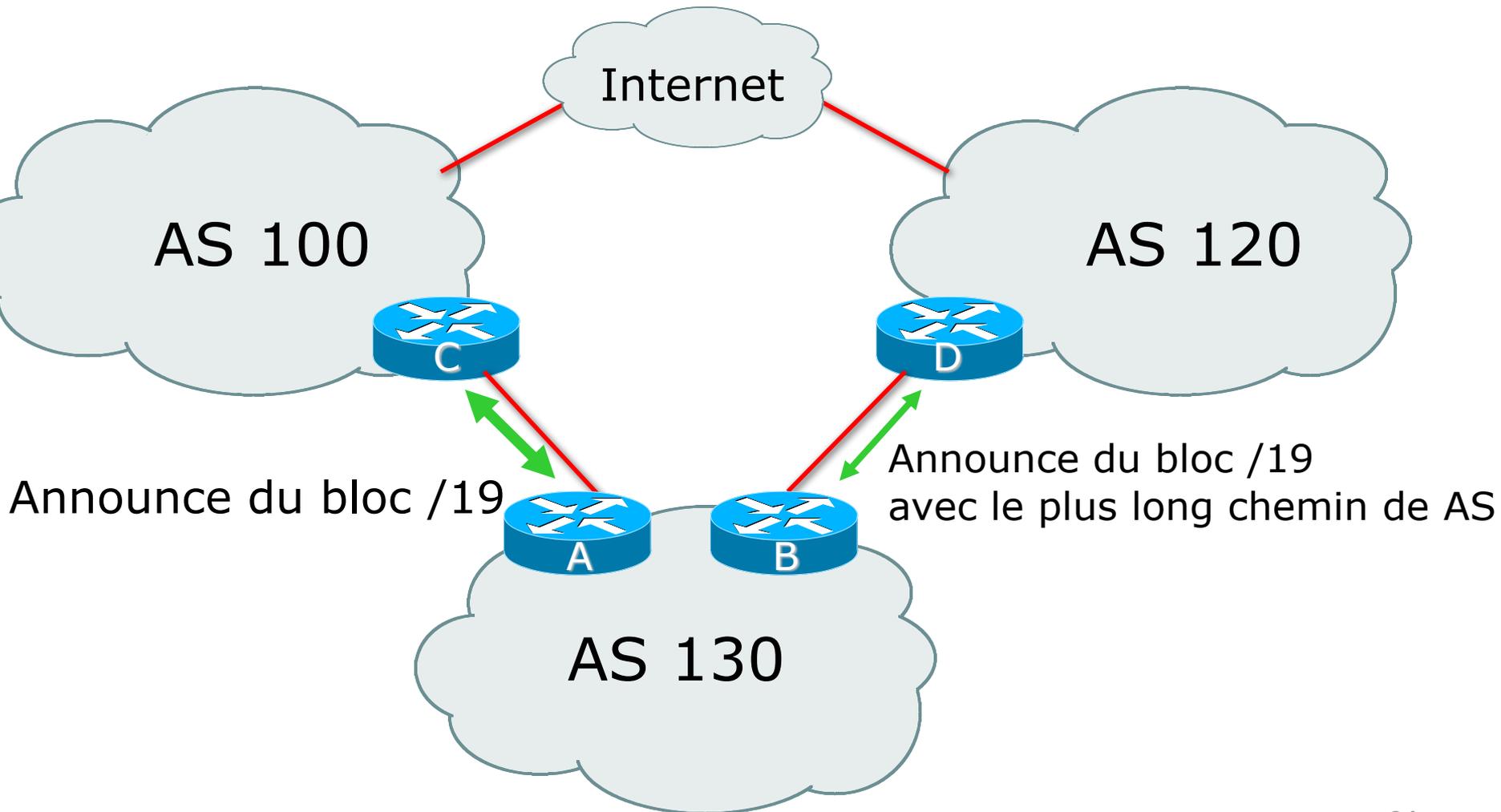
- Voir les préfixes émis par AS65534 sur Internet montre qu'ils semblent être l'origine à la fois par AS210 et AS200
  - Ce n'est pas mauvais
  - Ni illégal
- IOS command is `show ip bgp inconsistent`



# deux liens vers différents FAI

Un lien primaire, l'autre lien de sauvegarde uniquement

# deux liens vers différents FAI (dont un de secours seulement)



# deux liens vers différents FAI (dont un de secours seulement)

- Annoncer /19 agrégat sur chaque lien
  - lien principal rend l'annonce standard
  - lien de secours allonge l'AS PATH en utilisant AS PATH prepend
- Quand un lien échoue, l'annonce de l'agrégat /19 via l'autre lien assure une connectivité continue

# deux liens vers différents FAI (dont un de secours seulement)

- Routeur A Configuration

```
router bgp 130
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
neighbor 122.102.10.1 remote-as 100
```

```
neighbor 122.102.10.1 prefix-list aggregate out
```

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

```
!
```

```
ip prefix-list aggregate 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 liens vers différents FAI (dont un de secours seulement)

- Routeur B Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  neighbor 120.1.5.1 remote-as 120
  neighbor 120.1.5.1 prefix-list aggregate out
  neighbor 120.1.5.1 route-map routerD-out out
  neighbor 120.1.5.1 prefix-list default in
  neighbor 120.1.5.1 route-map routerD-in in
!
ip prefix-list aggregate permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
route-map routerD-out permit 10
  set as-path prepend 130 130 130
!
route-map routerD-in permit 10
  set local-preference 80
```

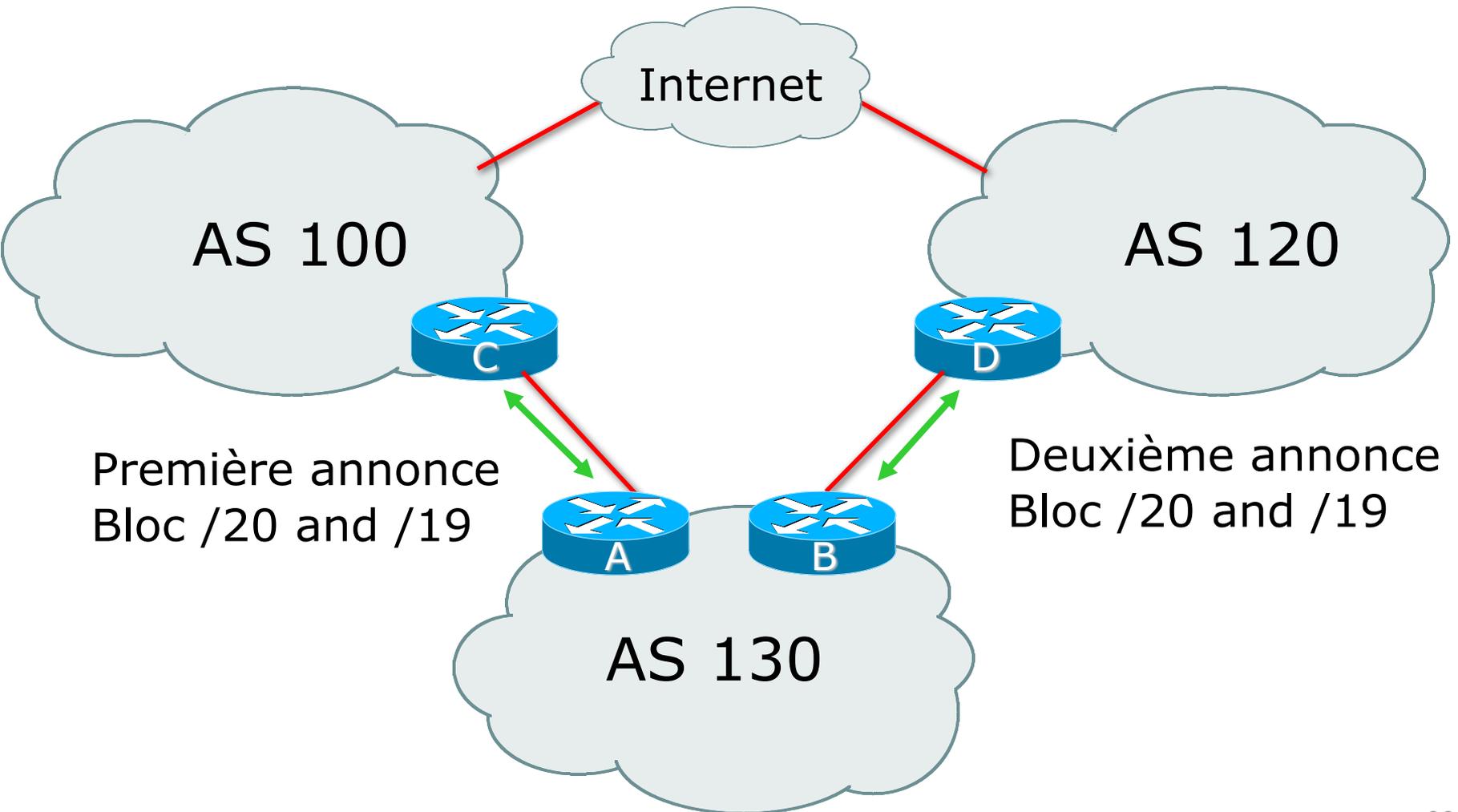
# deux liens vers différents FAI (dont un de secours seulement)

- Pas une situation courante comme la plupart des sites ont tendance à préférer utiliser toutes les capacités dont ils ont
  - (Utile lorsque deux FAI concurrents s'engagent à fournir une sauvegarde mutuelle les uns aux autres)
- Mais il présente les concepts de base de l'utilisation de local-prefs et AS-path preprends pour ingénierie du trafic dans la direction choisie

# Deux liens vers différents FAI

Avec le partage de charge

# deux liens vers différents FAI (Avec le partage de charge)



# deux liens vers différents FAI (Avec le partage de charge)

- Annoncer / 19 agrégat sur chaque lien
- Diviser /19 et annoncer comme deux /20s, un sur chaque lien
  - partage de charge de base entrant
- Quand un lien échoue, l'annonce de l'agrégat /19 via l'autre FAI assure une connectivité continue

# deux liens vers différents FAI (Avec le partage de charge)

- Routeur A Configuration

```
router bgp 130
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
network 121.10.0.0 mask 255.255.240.0
```

```
neighbor 122.102.10.1 remote-as 100
```

```
neighbor 122.102.10.1 prefix-list firstblock out
```

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

```
!
```

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

```
!
```

```
ip prefix-list firstblock permit 121.10.0.0/20
```

```
ip prefix-list firstblock permit 121.10.0.0/19
```

# deux liens vers différents FAI (Avec le partage de charge)

- Routeur B Configuration

```
router bgp 130
```

```
network 121.10.0.0 mask 255.255.224.0
```

```
network 121.10.16.0 mask 255.255.240.0
```

```
neighbor 120.1.5.1 remote-as 120
```

```
neighbor 120.1.5.1 prefix-list secondblock out
```

```
neighbor 120.1.5.1 prefix-list default in
```

```
!
```

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

```
!
```

```
ip prefix-list secondblock permit 121.10.16.0/20
```

```
ip prefix-list secondblock permit 121.10.0.0/19
```

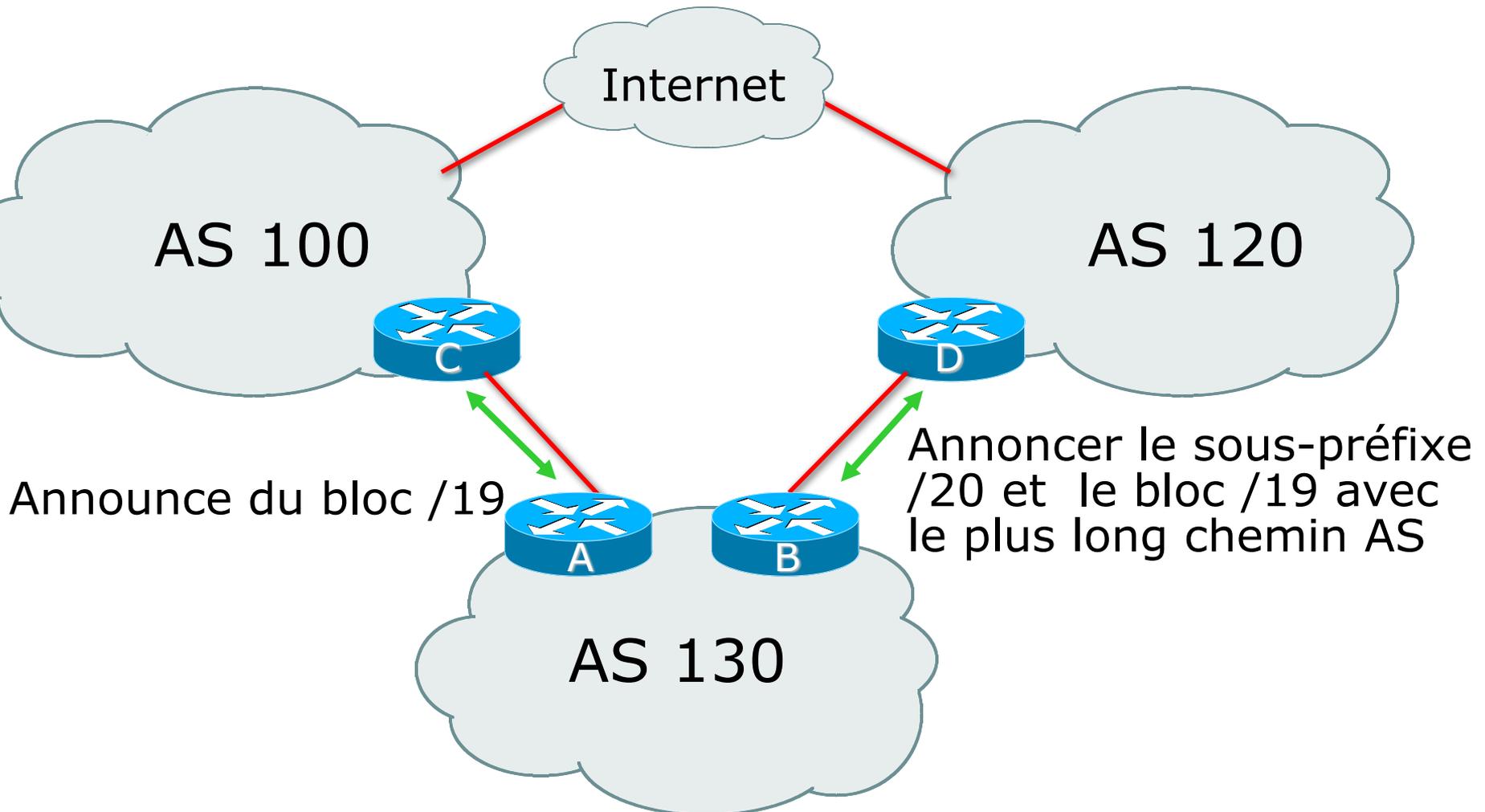
# deux liens vers différents FAI (Avec le partage de charge)

- Partage de charge dans ce cas est très basique
- Mais montre les premières étapes dans la conception d'une solution de répartition de charge
  - Commencer avec un concept simple
  - Et de le développer...!

# Deux liens vers différents FAI

Plus de partage de charge contrôlé

# Partage de charge avec différents FAI



# Partage de charge avec différents FAI

- Annoncer /19 agrégat sur chaque lien
  - Sur le premier lien, annoncer /19 comme d'habitude
  - Sur le deuxième lien, annoncer /19 avec le plus long AS PATH, et annoncer un sous-préfixe /20
    - Contrôles du partage de charge entre les opérateurs en amont et l'Internet
- Varier la taille du sous-préfixe et la longueur AS PATH jusqu'à ce que le partage de charge soit "parfait"
- Encore besoin de la redondance!

# Partage de charge avec différents FAI

- Routeur A Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 100
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list aggregate out
!
ip prefix-list aggregate 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
```

# Partage de charge avec différents FAI

- Routeur B Configuration

```
router bgp 130
  network 121.10.0.0 mask 255.255.224.0
  network 121.10.16.0 mask 255.255.240.0
  neighbor 120.1.5.1 remote-as 120
  neighbor 120.1.5.1 prefix-list default in
  neighbor 120.1.5.1 prefix-list subblocks out
  neighbor 120.1.5.1 route-map routerD out
!
route-map routerD permit 10
  match ip address prefix-list aggregate
  set as-path prepend 130 130
route-map routerD permit 20
!
ip prefix-list subblocks permit 121.10.0.0/19 le 20
ip prefix-list aggregate permit 121.10.0.0/19
```

# Partage de charge avec différents FAI

- Cet exemple est plus courant
- Montre comment les FAI et les sites finals subdivisent avec parcimonie l'espace d'adressage, ainsi que l'utilisation du concept AS-PATH prepend afin d'optimiser la répartition de charge entre les différents FAI
- Notez que le bloc agrégat /19 est TOUJOURS annoncé

# Résumé

# Résumé

- Les exemples précédents ont traité des cas simples
- L'équilibrage de charge des flux de trafic entrant
  - Obtenue en modifiant les annonces de routage sortants
  - Agrégat est toujours annoncé
- Nous n'avons pas regardé les flux de trafic sortant
  - Pour l'instant, cela est laissé comme "la plus proche sortie"

# 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 simple

Fin

