Réseaux d'Opérateurs

M1 RISE

Jean-Jacques Pansiot pansiot@unistra.fr

Réseaux IP d'opérateurs

Objectifs

– « mieux utiliser leur réseau IP» :
  - Meilleure utilisation des ressources
  - Meilleur service au client
    – donc facturé plus cher
  - Gérer relations avec leurs partenaires/ concurrents
– Nécessite de nouveaux mécanismes/ protocoles dans le réseau

Réseau IP « basique »

– Protocole de routage (RIP, OSPF, IS-IS)
  - Chaque routeur calcule plus court chemin
    – Vers chaque destination
      Repérée par un **préfixe :**
      130.79.90.0/23 192.168.16.32/27
  - Routage saut par saut
      Décision prise dans chaque routeur
  - Tous les paquets traités de la même façon
      Dépend uniquement de la destination

Exemple de table de routage

```
C       130.79.48.128/25 is directly connected, GigabitEthernet0/1.48
C       130.79.48.96/28 is directly connected, GigabitEthernet0/1.486
C       130.79.48.80/28 is directly connected, GigabitEthernet0/1.485
C       130.79.48.72/29 is directly connected, GigabitEthernet0/1.484
S       130.79.48.64/29 [1/0] via 130.79.48.73
C       130.79.48.48/28 is directly connected, GigabitEthernet0/1.483
C       130.79.48.32/28 is directly connected, GigabitEthernet0/1.482
C       130.79.48.16/28 is directly connected, GigabitEthernet0/1.481
C       130.79.208.224/29 is directly connected, GigabitEthernet0/0
C       130.79.48.0/28 is directly connected, GigabitEthernet0/1.480
S*   0.0.0.0/0 [1/0] via 130.79.208.228
```

Prochain saut/interface pour 130.79.48.20, 130.79.48.66, 130.79.48.121 ?

Quelques problèmes

- Router des paquets
  - Par un chemin « spécial »
    - Avec réservation de débit
      - => ingénierie de trafic (MPLS)
- Trafic d'un (opérateur) concurrent
  - Doit-on le laisser passer ?
    - Politique de routage (BGP)
- Isoler/protéger trafic d'une entreprise cliente
  - Virtual Private Networks VPN
  - => VPN BGP/MPLS

Routage inter-domaine et BGP

Structure d'Internet

- Internet
  - Interconnexion de réseaux
    - Hétérogènes en matériel/logiciels
    - Administrés par des entités différentes
    - Ayant des intérêts différents
  - Connectivité globale
  - Très grande taille
  - Centaines de millions de machines

Structure Internet

- Notion de domaine ou AS
  - Autonomous System
  - Ensemble connexe de réseaux IP
  - Géré par une seule entité
    - Entreprise, Etablissement
    - Réseaux d'opérateurs FAI/ISP
  - Routage interne
    - Un (ou plusieurs) IGP (Interior Gateway Protocol) :
      - RIP, OSPF, IS-IS …
      - Plus court chemin suivant métrique « technique »
      - Taille « raisonnable »
      - A priori tout le trafic doit être acheminé

Structure Internet

- AS interconnectés
  - Trafic entre AS
  - Trafic de **transit**
    - Traverse AS A
    - Source et destination extérieures à A
    - => utilise les ressources de A
  - Un AS peut
    - Être connecté à plusieurs autres AS
    - Être connecté à un AS par plusieurs liaisons/routeurs

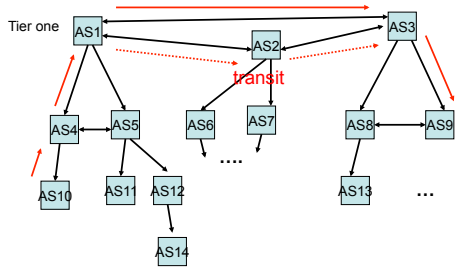
Structure Internet

- Relations entre AS :
  - en simplifiant deux types de relation
    - Relation **client fournisseur** A -> B
      - A offre un service à B
      - trafic de/vers B (ou de/vers ses clients)
        - » Peut transiter par A
      - Relation en général payante
      - Ex : entreprise cliente d'un FAI
    - Relation **pair à pair A<-> B**
      - Échange de trafic local
      - Transit mutuel
      - Relation non payante (échange)
      - Ex : entre deux FAI de même importance

Transit

- Pas toujours symétrique
  - A offre transit à B mais pas l'inverse
    - Ex B client de A
- Pas toujours transitif
  - A offre transit à B et B à C mais pas A à C
    - Les pairs sont aussi des concurrents
    - Le trafic de AS4(client de AS1) vers AS9 (client de AS3) ne peut pas transiter par AS2 (concurrent de AS1 et AS2)
    - => nécessité de créer une connexion directe AS1 <-> AS3 (ou de passer par un fournisseur commun

Graphe des AS



Graphe des AS

- Hiérarchie
  - Sommet (tier one) : client de personne
    - Opérateurs principaux (Sprint, Uunet, …)
  - Base : stub AS (ou feuille)
    - Connecté à un seul autre AS (fournisseur)
    - Protocole de Routage inter-domaine pas toujours nécessaire
  - Un AS peut être connecté à plusieurs autres sans admettre de transit
    - Client avec plusieurs fournisseurs
      - Redondance et/ou multihoming

Principaux tier 1

- AOL
- ATT
- Global Crossing
- Level3
- Verizon (UUNet)
- NTT (Verio)
- Qwest
- SAVVIS
- Sprint

Route et AS

- route de A à B traverse
  - AS_1, AS_2, AS_3, … AS_n
    - $A \in AS_1$, $B \in AS_n$
    - soit
      $AS_1 <= AS_2 <= \dots AS_i => AS_{i+1} => \dots AS_n$
    - soit
      $AS_1 <= AS_2 <= \dots AS_i = AS_{i+1} => \dots => AS_n$

« montée, plat éventuel, descente »

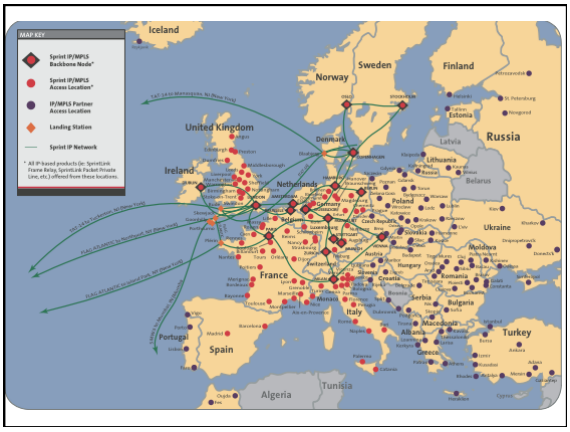
```
traceroute to 203.222.38.67 (203.222.38.67), 64 hops max, 40 byte packets
 1  api-rc1-ge-0-2-0-2 (130.79.91.253) [AS2200]  1 ms  1 ms  1 ms     en fait AS2259
 2  crc-rc1-ge-0-1-0-0 (130.79.20.13) [AS2200]  117 ms  1 ms  1 ms
 3  strasbourg-g3-3.cssi.renater.fr (193.51.184.42) [AS2200]  1 ms  1 ms  0 ms
 4  besancon-pos2-0.cssi.renater.fr (193.51.180.9) [AS2200] [MPLS: Label 194 Exp 0]  14 ms
 5  dijon-pos1-0.cssi.renater.fr (193.51.179.229) [AS2200] [MPLS: Label 37 Exp 0]  14 ms
 6  lyon-pos5-0.cssi.renater.fr (193.51.180.46) [AS2200]  13 ms  13 ms  13 ms
 7  ftld-lyon.cssi.renater.fr (193.51.185.29) [AS2200]  14 ms  14 ms  14 ms
 8  po6-0.auvbb1.aubervilliers.opentransit.net (193.251.241.94) [AS5511]  14 ms  14 ms  14 ms
 9  pos0-7-0-1.auvtr1.aubervilliers.opentransit.net (193.251.128.106) [AS5511]  15 ms * *
10  po0-0.ashcr2.ashburn.opentransit.net (193.251.243.170) [AS5511]  93 ms  102 ms  209 ms
11  so-1-0-0-0.atlcr1.atlanta.opentransit.net (193.251.240.154) [AS5511]  110 ms  106 ms  192
12  so-0-0-0-0.dalcr2.dallas.opentransit.net (193.251.243.114) [AS5511]  126 ms  128 ms  126
13  sprint.gw.opentransit.net (193.251.247.158) [AS5511]  126 ms  126 ms  126 ms
14  sl-bb20-fw-6-0.sprintlink.net (144.232.20.80) [AS1239]  127 ms  129 ms  127 ms
15  sl-bb21-fw-14-0.sprintlink.net (144.232.11.218) [AS1239]  127 ms  127 ms  132 ms
16  sl-bb22-ana-12-0.sprintlink.net (144.232.20.131) [AS1239]  164 ms  164 ms  164 ms
17  sl-bb20-ana-15-0.sprintlink.net (144.232.1.178) [AS1239]  168 ms  168 ms  168 ms
18  sl-bb24-sj-4-0.sprintlink.net (144.232.9.93) [AS1239]  168 ms  168 ms  169 ms
19  sl-bb20-hk-3-3.sprintlink.net (144.232.8.216) [AS1239]  342 ms  338 ms  338 ms
20  sl-gw1-hk-0-0-0.sprintlink.net (203.222.38.67) [AS1239/AS4657]  346 ms  343 ms *
```

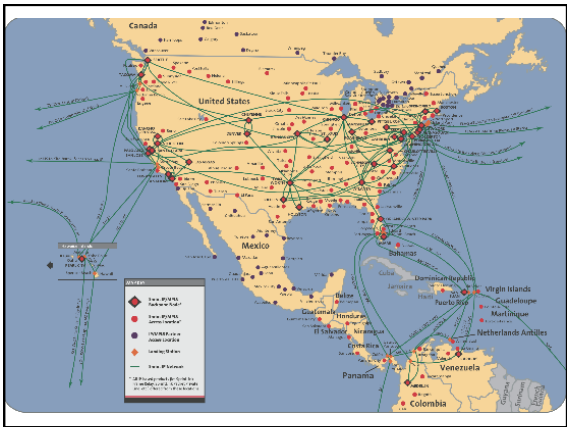


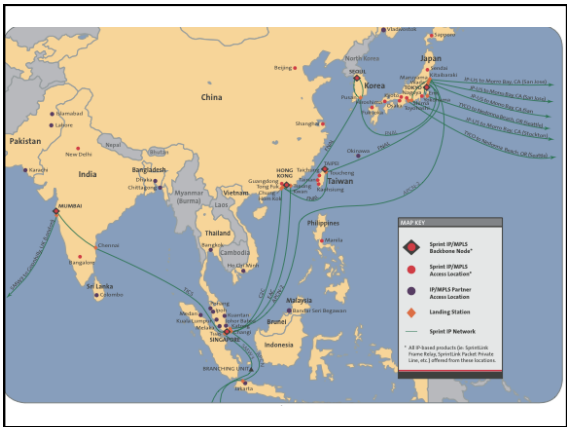
```
sudo ./lft -A 203.222.38.67
TTL LFT trace to sl-gw1-hk-.sprintlink.net (203.222.38.67):80/tcp
 1  [2259] api-rc1-ge-0-2-0-2.u-strasbg.fr (130.79.91.253) 0.9/0.8ms
 2  [2259] crc-rc1-ge-0-1-0-0.u-strasbg.fr (130.79.20.13) 1.0/2.3ms
**  [neglected] no reply packets received from TTL 3
 4  [2200] te0-2-0-0-besancon-rtr-011.noc.renater.fr (193.51.189.122) 11.7/13.6ms
 5  [2200] te0-1-0-0-dijon-rtr-011.noc.renater.fr (193.51.189.114) 14.4/12.0ms
 6  [2200] te0-3-2-0-lyon1-rtr-001.noc.renater.fr (193.51.189.141) 10.0/16.7ms
 7  [3356] xe-8-0-0.edge5.Paris1.Level3.net (212.73.207.173) 76.4/15.9ms
 8  [3356] ae-33-51.ebr1.Paris1.Level3.net (4.69.139.193) 16.5/16.2ms
 9  [3356] ae-2-2.ebr1.London1.Level3.net (4.69.142.106) 23.4/23.2ms
10  [3356] ae-11-51.car1.London1.Level3.net (4.69.139.66) 23.8/23.4ms
11  [1239] sl-bb21-lon-10-0-0.sprintlink.net (213.206.131.21) 23.8/23.1ms
12  [1239] sl-bb22-lon-3-0-0.sprintlink.net (213.206.129.153) 23.6/23.6ms
13  [1239] sl-bb20-bru-14-0-0.sprintlink.net (213.206.129.42) 32.6/32.7ms
14  [1239] sl-bb21-bru-15-0-0.sprintlink.net (80.66.128.42) 32.8/32.5ms
15  [1239] sl-bb20-ams-14-0-0.sprintlink.net (213.206.129.45) 32.9/32.2ms
16  [1239] sl-bb21-ham-6-0-0.sprintlink.net (213.206.129.145) 38.6/39.0ms
17  [1239] sl-bb21-cop-13-0-0.sprintlink.net (213.206.129.57) 43.8/43.7ms
18  [1239] sl-bb21-sto-14-0-0.sprintlink.net (213.206.129.34) 52.6/52.1ms
19  [1239] sl-bb20-hk-11-1-3.sprintlink.net (213.206.129.131) 283.9ms
**  [80/tcp no reply from target]. Try advanced options (use -VV to see packets).
```



```
traceroute to 203.222.38.67 (203.222.38.67), 64 hops max, 40 byte packets
 1  api-rc1-ge-0-2-0-2 (130.79.91.253) [AS2200]  1 ms  1 ms  1 ms
 2  crc-rc1-ge-0-1-0-0 (130.79.20.13) [AS2200]  1 ms  1 ms  1 ms
 3  * * *
 4  te0-2-0-0-besancon-rtr-011.noc.renater.fr (193.51.189.122) [AS2200] [MPLS: Label 16040 Exp
    0]  14 ms  13 ms  14 ms
 5  te0-1-0-0-dijon-rtr-011.noc.renater.fr (193.51.189.114) [AS2200] [MPLS: Label 16033 Exp 0]  14
    ms  13 ms  14 ms
 6  te0-3-2-0-lyon1-rtr-001.noc.renater.fr (193.51.189.141) [AS2200]  10 ms  10 ms  10 ms
 7  xe-8-0-0.edge5.Paris1.Level3.net (212.73.207.173) [AS9057/AS3356]  16 ms  16 ms  16 ms
 8  ae-33-51.ebr1.Paris1.Level3.net (4.69.139.193) [AS3356]  17 ms  17 ms  16 ms
 9  ae-2-2.ebr1.London1.Level3.net (4.69.142.106) [AS3356]  23 ms  23 ms  23 ms
10  ae-11-51.car1.London1.Level3.net (4.69.139.66) [AS3356]  24 ms  23 ms  23 ms
11  sl-bb21-lon-10-0-0.sprintlink.net (213.206.131.21) [<NONE>]  24 ms  24 ms  24 ms
12  sl-bb22-lon-3-0-0.sprintlink.net (213.206.129.153) [<NONE>]  24 ms  24 ms  27 ms
13  sl-bb20-bru-14-0-0.sprintlink.net (213.206.129.42) [<NONE>]  32 ms  34 ms  33 ms
14  sl-bb21-bru-15-0-0.sprintlink.net (80.66.128.42) [<NONE>]  33 ms  33 ms  33 ms
15  sl-bb20-ams-14-0-0.sprintlink.net (213.206.129.45) [<NONE>]  33 ms  33 ms  33 ms
16  sl-bb21-ham-6-0-0.sprintlink.net (213.206.129.145) [<NONE>]  39 ms  39 ms  39 ms
17  sl-bb21-cop-13-0-0.sprintlink.net (213.206.129.57) [<NONE>]  44 ms  44 ms  44 ms
18  sl-bb21-sto-14-0-0.sprintlink.net (213.206.129.34) [<NONE>]  53 ms  53 ms  52 ms
19  sl-bb20-hk-11-1-3.sprintlink.net (213.206.129.131) [<NONE>]  284 ms  284 ms  284 ms
20  sl-gw1-hk-.sprintlink.net (203.222.38.67) [AS4657]  281 ms  281 ms *
```


MAP KEY
Sprint IP/MPLS Backbone Node*
Sprint IP/MPLS Access Location*
IP/MPLS Partner Access Location
Landing Station
Sprint IP Network
* All IP-based products (ie. SprintLink Frame Relay, SprintLink Packet Private Line, etc.) offered from these locations.
Iceland
Norway
Sweden
Finland
Russia
Estonia
Latvia
Lithuania
Belarus
Ukraine
Moldova
Romania
Bulgaria
Turkey
United Kingdom
Ireland
Denmark
Netherlands
Belgium
Poland
Germany
Slovakia
Hungary
Slovenia
Croatia
Serbia
Macedonia
Albania
Greece
France
Switzerland
Austria
Italy
Monaco
Portugal
Spain
Algeria
Tunisia


Canada
United States
Mexico
Bahamas
Virgin Islands
Guadeloupe
Martinique
Netherlands Antilles
Venezuela
Colombia
Panama
Honduras


China
Japan
Korea
Taiwan
Pakistan
India
Bangladesh
Vietnam
Thailand
Cambodia
Philippines
Malaysia
Brunei
Indonesia
Sri Lanka
MUMBAI
SINGAPORE
MAP KEY
Sprint IP/MPLS Backbone Node*
Sprint IP/MPLS Access Location*
IP/MPLS Partner Access Location
Landing Station
Sprint IP Network
* All IP-based products (ie. SprintLink Frame Relay, SprintLink Packet Private Line, etc.) offered from these locations.

Structure d'un POP (Sprint)

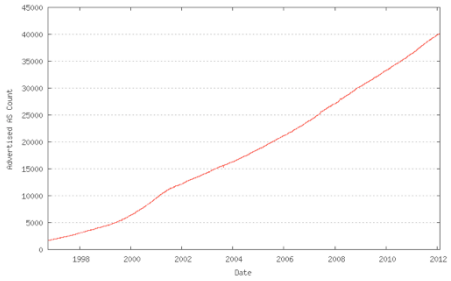
- Point of Presence
- plusieurs routeurs de backbone rbb
  - full mesh
  - au moins 2 liens (de 2 routeurs diff) vers autres POP
  - 2,4 ou 10 Gb/s
- plusieurs routeurs « gateway » gw
  connectés à au moins 2 rbb du POP
  connectés aux clients ou pairs


```
traceroute to 203.222.38.67 (203.222.38.67), 64 hops max, 40 byte packets
 1  api-rc1-ge-0-2-0-2 (130.79.91.253) [AS2200]  1 ms  1 ms  1 ms
 2  crc-rc1-ge-0-1-0-0 (130.79.20.13) [AS2200]  117 ms  1 ms  1 ms
 3  strasbourg-g3-3.cssi.renater.fr (193.51.184.42) [AS2200]  1 ms  1 ms  0 ms
 4  besancon-pos2-0.cssi.renater.fr (193.51.180.9) [AS2200] [MPLS: Label 194 Exp 0]  14 ms
 5  dijon-pos1-0.cssi.renater.fr (193.51.179.229) [AS2200] [MPLS: Label 37 Exp 0]  14 ms
 6  lyon-pos5-0.cssi.renater.fr (193.51.180.46) [AS2200]  13 ms  13 ms  13 ms
 7  ftld-lyon.cssi.renater.fr (193.51.185.29) [AS2200]  14 ms  14 ms  14 ms
 8  po6-0.auvbb1.aubervilliers.opentransit.net (193.251.241.94) [AS5511]  14 ms  14 ms  14 ms
 9  pos0-7-0-1.auvtr1.aubervilliers.opentransit.net (193.251.128.106) [AS5511]  15 ms * *
10  po0-0.ashcr2.ashburn.opentransit.net (193.251.243.170) [AS5511]  93 ms  102 ms  209 ms
11  so-1-0-0-0.atlcr1.atlanta.opentransit.net (193.251.240.154) [AS5511]  110 ms  106 ms  192
12  so-0-0-0-0.dalcr2.dallas.opentransit.net (193.251.243.114) [AS5511]  126 ms  128 ms  126
13  sprint.gw.opentransit.net (193.251.247.158) [AS5511]  126 ms  126 ms  126 ms
14  sl-bb20-fw-6-0.sprintlink.net (144.232.20.80) [AS1239]  127 ms  129 ms  127 ms
15  sl-bb21-fw-14-0.sprintlink.net (144.232.11.218) [AS1239]  127 ms  127 ms  132 ms
16  sl-bb22-ana-12-0.sprintlink.net (144.232.20.131) [AS1239]  164 ms  164 ms  164 ms
17  sl-bb20-ana-15-0.sprintlink.net (144.232.1.178) [AS1239]  168 ms  168 ms  168 ms
18  sl-bb24-sj-4-0.sprintlink.net (144.232.9.93) [AS1239]  168 ms  168 ms  169 ms
19  sl-bb20-hk-3-3.sprintlink.net (144.232.8.216) [AS1239]  342 ms  338 ms  338 ms
20  sl-gw1-hk-0-0-0.sprintlink.net (203.222.38.67) [AS1239/AS4657]  346 ms  343 ms *
```


Ex : routeur peering

144.232.20.81 (sl-st21-dal-13-0.sprintlink.net) [version 12.0]:
144.228.241.244 -> 0.0.0.0
144.232.20.17 -> 144.232.20.16 (sl-bb20-fw-13-0.sprintlink.net)
144.228.250.73 -> 0.0.0.0
144.228.250.77 -> 0.0.0.0
144.232.20.81 -> 144.232.20.80 (sl-bb20-fw-6-0.sprintlink.net)
208.173.178.134 -> 208.173.178.133 (dpr1-so-0-1-0.dallasequinix.savvis.net)
193.251.247.158 -> 193.251.247.157 (so-1-1-0-0.dalcr2.Dallas.opentransit.net)
144.232.8.137 -> 0.0.0.0 (local)
208.51.134.34 -> 208.51.134.33 (so1-1-0-2488M.ar1.DAL2.gblx.net)
144.232.9.95 -> 144.232.9.94 (sl-st20-dal-3-0.sprintlink.net)

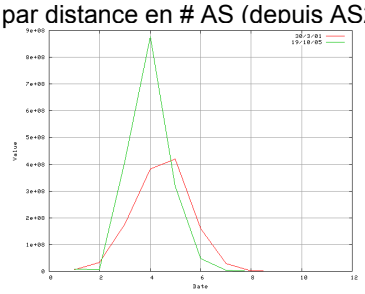
Nombre AS vus 97-2012
source www.potaroo.net



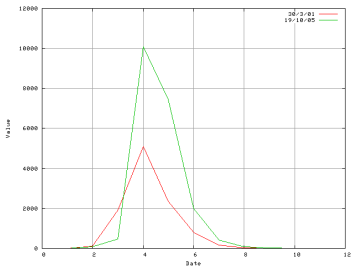
Graphe des AS

- Près de 40 000 AS effectifs (plus réservés)
  - AS codés sur 16 bits => nouveau codage 32 bits
- Plus de 14 500 feuilles
  - Origine d' annonces mais pas utilisé par transit
- Pratiquement tous les autres AS
  origine ET transit
- Longueur moyenne depuis le « centre » 3,6
  - Degré moyen 4 (en 2000)
  - Degré max 1704 (en 2000) uunet
  - Diamètre 10

Répartition des adresses
par distance en # AS (depuis AS2221)



Répartition des AS
par distance en # AS depuis AS2221



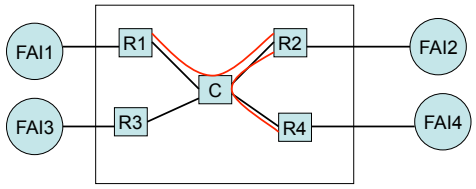
Politique de routage

- Chaque AS souhaite
  - Sélectionner le trafic autorisé en transit
    - Sur quel critère (destination …)
    - Ne pas filtrer le transit, le « détourner »
  - Choisir par quel(s) voisins/sortie
    - Envoyer le trafic vers une destination
    - Et coordonner avec le routage intra domaine
  - Choisir/influencer par quel voisin
    - Recevoir certains flux (plus difficile)
  - Le tout
    - Extensible à 40K AS, 400K destination,
    - avec un réseau qui change en permanence

Interconnexion d'AS

- Deux (routeurs d')AS différents
  - Connectés à distance par liaison point-à-point
    - Pas de matériel commun
    - Limite de responsabilité = liaison
  - Coûteux si nombreux AS voisins
  - Points d'échanges (GIX) (ou NAP)
    - http://www.eurogix.org/fr/accueil.php

GIX /IXP



R routeur, C commutateur

peering

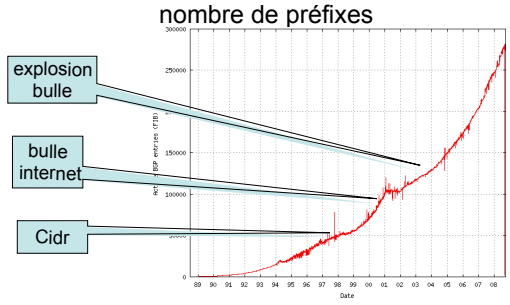
Adressage

- Destination
  - Jusqu'en 1995
    - Réseaux répartis en classes (A, B, C)
    - Table contient
      - Réseaux distants
      - Sous-réseaux du réseau local
      - Adresse par défaut (0.0.0.0)
    - => épuisement des classes A et B
    - => utilisation massive des classes C
    - => explosion des tables de routage

CIDR

- Classless InterDomain Routing RFC 1518 et 1519
  - Allocation de blocs d'adresses
    - Représenté par un préfixe ex : 192.168.36.0 /23
    - « super réseau » de taille 2^n, n quelconque
  - Allocation hiérarchique
    - Préfixe client inclus dans bloc fournisseur
    - Autorise agrégation dans le réseau (adresses PA)
    - Problème pour multihoming et
    - chgt fournisseur => renumérotation
  - Routage par « longest match »
    - Dans la FIB prendre plus long préfixe correspondant à l'adresse destination

Taille table de commutation (FIB) nombre de préfixes



Protocoles de routage inter-domaine

- Protocoles IGP non adaptés
  - Extensibilité
  - Expression de la politique de routage
- EGP Exterior Gateway Protocol
  - Apparu en 1984 (ne gère pas les cycles)
- BGP Border Gateway Protocol
  - 1ère version 1989
  - BGP 4 en 1995 RFC 1771 support CIDR
  - M-BGP (BGP4+) extensions multi protocoles
    - RFC 2283 => 2858

BGP

- Objectifs
  - Diffuser l'existence et l'accessibilité de destinations (préfixe CIDR)
  - Contrôler que
    - le transit respecte la politique de l'AS
    - Que les routes sont valides (pas de boucles)
    - Sont les « meilleures » (selon critères locaux et globaux)
  - Dynamiquement en cas de pannes/changements

Sessions BGP

- Un routeur BGP ouvre une session avec ses voisins (peer) BGP
  - Sessions fiables : connexion TCP port 179
    - Voisins déclarés (sécurité)
    - Possibilité d'authentification
  - Deux types de peer
    - Voisins dans l'AS iBGP (interior BGP)
    - Voisins extérieur eBGP (exterior BGP)

Annonce

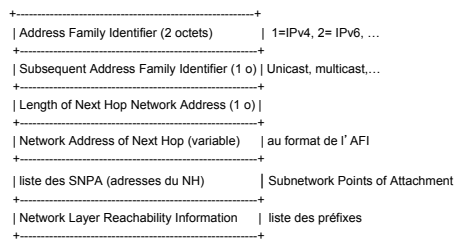
- Destination
  - AFI : address family (ipv4, ipv6, …)
  - SAFI (sub address family: unicast, VPN, …)

  Ex : préfixe CIDR en IPv4
- Attributs permettant de
  - filtrer les annonces indésirables
  - Sélectionner la meilleure annonce (à utiliser)
  - Influencer les routeurs en aval
  - Ex : AS-Path, Origin, Next-Hop, Local Pref
- Annonce peut être retirée explicitement
  - Withdraw

Update

Unfeasible Routes Length (2 octets)	« withdraw »
Withdrawn Routes (variable)	liste préfixes retirés 0 à n
Total Path Attribute Length (2 octets)	« annonce 1 route »
Path Attributes (variable)	liste des attributs de la route
Network Layer Reachability Information	liste préfixes access. via cette route

Attribut Multi-proto NLRI (rfc 2858)



Annonces (suite)

- Routeur maintient
  - Base d'annonces (acceptées) Adj-RIB
  - Possibilité plusieurs annonces même dest.
  - Seule meilleure annonce
    - Utilisée par routeur (FIB)
    - Propagée (éventuellement) aux voisins
  - Base d'annonce« **Hard State** »
    - Pas de timeout sur les annonces
    - extensibilité

Session / messages

- Au démarrage routeur/BGP/liaison
  - **Open**
    - **négociation de « capabilities » : IPv6, multicast, …**
- Cycliquement
  - **Keep Alive** (vérification session)
- Si changement RIB
  - **Update** : envoi annonces/retraits
- Si anomalie
  - **Notification** (fin session), suppression annonces de la session

BGP et routage intra

- iBGP assure même vue de tous les routeurs
  - Full mesh
    - connexion avec tous les routeurs BGP de l'AS
  - Annonce reçue par iBGP non propagées par iBGP
    - Évite les boucles
  - Coûteux => Possibilité de Route Reflector
- Connexion iBGP à travers routeurs intra possible (et problème du NextHop)
- Nécessité de redistribution de routes
  - de BGP vers IGP et réciproquement
  - Problème des métriques incomparables

Algorithme de BGP

- Principe :
  - Envoyer annonce préfixe P vers routeur R
  - => accepter le trafic venant de R pour P
- Algorithme non complètement spécifié
  - Ordre de traitement des attributs indéfini
- « tous les coups sont permis »
  - Filtrage
  - Modification des attributs

Algorithme de BGP

- Réception d'une annonce
  - Filtrer annonces indésirables (et MaJ attributs)
    - politique d'**importation**, détermine trafic **sortant**
  - Insérer annonce dans Adj-RIB (si non filtrée)
    - Calculer meilleure route pour préfixe
      - Dépend ordre des attributs
  - Si meilleure route change (ou nouveau préfixe)
    - Maj FIB et redistribution IGP (si meilleur que IGP)
    - Pour chaque voisin BGP V
      - Maj attributs (AS Path notamment) et/ou filtrage
      - Si non-filtrage envoyer Update à V
      - Politique d'**exportation (et donc transit si V externe)**
      - Détermine le trafic **entrant** et le **transit**

Quelques attributs (1)

- **AS-Path**
  - Indique suite d'AS traversés par annonce
  - Chaque AS ajoute son ASN (a priori)
    - aux annonces vers l'extérieur
    - BGP : protocole à **vecteur de chemins**
    - Détection de boucle (refus des annonces avec son ASN)
    - Longueur AS Path critère principal par défaut de choix
    - Possibilité de faire du « padding »

Exemple

Network	Next Hop	Metric	LP	W	Path
*130.79.0.0	204.42.253.253			0	267 1225 2914 5511 2200 2259 i
*	195.66.225.254			0	5459 5511 2200 2259 i
*	158.43.206.96			0	1849 702 4000 5511 2200 2259 i
*	205.158.2.126			0	2828 5511 2200 2259 i
*	193.0.0.56			0	3333 1103 8933 2200 2259 i
*	12.127.0.249			0	7018 1740 1239 5511 2200 2259 i
*	4.0.0.2		120	0	1 1239 5511 2200 2259 i
*	134.55.24.6			0	293 1800 5511 2200 2259 i
*>	129.250.0.1	16		0	2914 5511 2200 2259 i
*	204.212.44.129			0	234 2914 5511 2200 2259 i
*	204.70.4.89			0	3561 1239 5511 2200 2259 i
*	202.232.1.8			0	2497 1239 5511 2200 2259 i
*	192.121.154.25			0	1755 2603 8933 2200 2259 i
*	129.250.0.3	48		0	2914 5511 2200 2259 i

Quelques attributs (2)

- **Origin**
  - Indique comment premier AS connaît P
    - i = IGP (connu par networks),
    - e = EGP (le protocole EGP),
    - ? = inconnu (agrégation, redistribution IGP)

Exemple

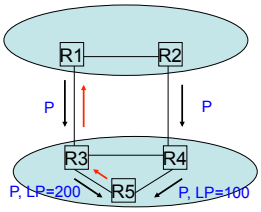
Network	Next Hop	Metric	LP	W Path
*130.79.0.0	204.42.253.253			0 267 1225 2914 5511 2200 2259 i
*	195.66.225.254			0 5459 5511 2200 2259 i
*	158.43.206.96			0 1849 702 4000 5511 2200 2259 i
*	205.158.2.126			0 2828 5511 2200 2259 i
*	193.0.0.56			0 3333 1103 8933 2200 2259 i
*	12.127.0.249			0 7018 1740 1239 5511 2200 2259 i
*	4.0.0.2		120	0 1 1239 5511 2200 2259 i
*	134.55.24.6			0 293 1800 5511 2200 2259 i
*>	129.250.0.1	16		0 2914 5511 2200 2259 i
*	204.212.44.129			0 234 2914 5511 2200 2259 i
*	204.70.4.89			0 3561 1239 5511 2200 2259 i
*	202.232.1.8			0 2497 1239 5511 2200 2259 i
*	192.121.154.25			0 1755 2603 8933 2200 2259 i
*	129.250.0.3	48		0 2914 5511 2200 2259 i

Quelques attributs (3)

- **Local Preference**
  - Signification locale (iBGP)
  - Sélection meilleure sortie
    - entre routeurs **même domaine**
      - influencer sortie
    - a priori LP pas envoyée en eBGP
    - et ignorée d'un voisin eBGP
  - Nombre + grand = meilleur, défaut =100

Ex local preference

Configuration de R3 et R4 : meilleure sortie R3



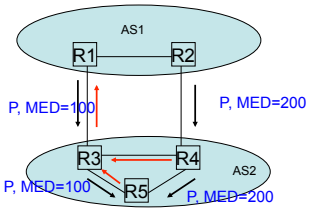
Remarque
Si R3 envoie
son annonce avant R4,
Alors R4 n'envoie pas
la sienne

Quelques attributs (4)

- **Multi Exit Discriminator** (MED, metric)
  - Nombre, + petit meilleur, défaut 0
  - Envoyé voisins externes
  - Ex :
    - AS1 envoie à AS2 annonce pour P
      - avec MED = 0 via routeur V1
      - avec MED = 100 via routeur V2
    - Influer sur porte d'entrée
      - V1 et V2 échangent annonces => meilleure via V1

Ex MED

Configuration R1, R2 : meilleure entrée R1
AS1 influence trafic sortant de AS2



Exemple Metric

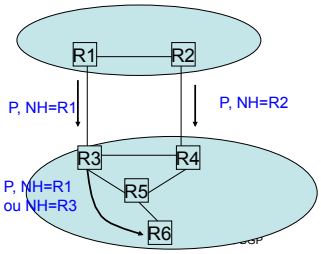
Network	Next Hop	Metric	LP	W Path
*130.79.0.0	204.42.253.253			0 267 1225 2914 5511 2200 2259 i
*	195.66.225.254			0 5459 5511 2200 2259 i
*	158.43.206.96			0 1849 702 4000 5511 2200 2259 i
*	205.158.2.126			0 2828 5511 2200 2259 i
*	193.0.0.56			0 3333 1103 8933 2200 2259 i
*	12.127.0.249			0 7018 1740 1239 5511 2200 2259 i
*	4.0.0.2	120		0 1 1239 5511 2200 2259 i
*	134.55.24.6			0 293 1800 5511 2200 2259 i
*>	129.250.0.1	16		0 2914 5511 2200 2259 i
*	204.212.44.129			0 234 2914 5511 2200 2259 i
*	204.70.4.89			0 3561 1239 5511 2200 2259 i
*	202.232.1.8			0 2497 1239 5511 2200 2259 i
*	192.121.154.25			0 1755 2603 8933 2200 2259 i
*	129.250.0.3	48		0 2914 5511 2200 2259 i

Quelques attributs (3)

- **NextHop**
  - Normalement adresse routeur source annonce (source TCP)
    - Voisin si eBGP
    - Dans iBGP : par défaut NextHop reçu par eBGP
      - (≠ prochain saut FIB)
  - Le NextHop doit être accessible
    - Annoncé par IGP si pas directement connecté
      - Prochain_saut_FIB(P) = prochain_saut_IGP(NextHop_BGP(P))
    - Possibilité de le ré-écrire (adresse loopback p.e.)

Ex NextHop

R3 envoie à R6 (iBGP) un NH :
R1 (recopie annonce) ou R3 (self)



R6 devra envoyer au prochain saut interne vers R1 ou R3
ex : R5

Exemple

Network	Next Hop	Metric	LP	W Path
*130.79.0.0	204.42.253.253			0 267 1225 2914 5511 2200 2259 i
*	195.66.225.254			0 5459 5511 2200 2259 i
*	158.43.206.96			0 1849 702 4000 5511 2200 2259 i
*	205.158.2.126			0 2828 5511 2200 2259 i
*	193.0.0.56			0 3333 1103 8933 2200 2259 i
*	12.127.0.249			0 7018 1740 1239 5511 2200 2259 i
*	4.0.0.2	120		0 1 1239 5511 2200 2259 i
*	134.55.24.6			0 293 1800 5511 2200 2259 i
*>	129.250.0.1	16		0 2914 5511 2200 2259 i
*	204.212.44.129			0 234 2914 5511 2200 2259 i
*	204.70.4.89			0 3561 1239 5511 2200 2259 i
*	202.232.1.8			0 2497 1239 5511 2200 2259 i
*	192.121.154.25			0 1755 2603 8933 2200 2259 i
*	129.250.0.3	48		0 2914 5511 2200 2259 i

Quelques attributs (4)

- Attribut **Community**
  - Liste de communautés
    - 32 bits : 16 bits ASN (AS qui a attribué) 16 bits N°
  - Communautés prédéfinies
    - 0xFFFFFF01 no export (ne pas annoncer via eBGP)
    - 0xFFFFFF02 no advertise (ni eBGP ni iBGP)

Quelques attributs (4)

- Attribut **Community**
  - Classification globale des annonces
    - Ex : vient de pair, de client, de fournisseur
      - Marquage à l'entrée
      - Filtrage en sortie ex :
        - » Vers client envoyer tout
        - » Vers fournisseur envoyer clients (et pairs ?)
        - » Vers pairs envoyer clients (et fournisseurs ?)
  - Simplifie l'expression des politiques

Ordre des attributs

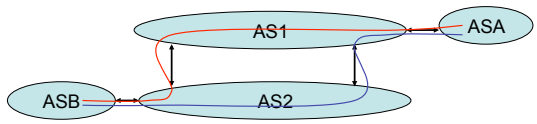
- Classement lexicographique (cisco, simplifié)

1 + grande local pref
2 plus court AS-Path
3 Origin i puis e puis ?
4 Plus petit MED (metric) du même AS
5 Plus proche NextHop (métrique IGP)
Hot potatoe routing
6 Plus ancienne route (évite oscillations)
7 Si ex æquo : identif. routeur BGP

http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a0080094431.shtml

Hot potatoe routing

dû à la règle 5



A -> B
B -> A

cause routage disymétrique

Exemple choix = >

	Network	Next Hop	Metric	LP	W Path
*130.79.0.0	204.42.253.253			0	267 1225 2914 5511 2200 2259 i
*	195.66.225.254			0	5459 5511 2200 2259 i
*	158.43.206.96			0	1849 702 4000 5511 2200 2259 i
*	205.158.2.126			0	2828 5511 2200 2259 i
*	193.0.0.56			0	3333 1103 8933 2200 2259 i
*	12.127.0.249			0	7018 1740 1239 5511 2200 2259 i
*	4.0.0.2		120	0	1 1239 5511 2200 2259 i
*	134.55.24.6			0	293 1800 5511 2200 2259 i
*>	129.250.0.1		16	0	2914 5511 2200 2259 i
*	204.212.44.129			0	234 2914 5511 2200 2259 i
*	204.70.4.89			0	3561 1239 5511 2200 2259 i
*	202.232.1.8			0	2497 1239 5511 2200 2259 i
*	192.121.154.25			0	1755 2603 8933 2200 2259 i
*	129.250.0.3		48	0	2914 5511 2200 2259 i

Exemple table (suite)

Bourrage d' AS-Path et découpage classe A

*	4.17.106.0/24	204.42.253.253	0 267 1225 701 13832 13832 13832 i
*		158.43.206.96	0 1849 702 701 13832 13832 13832 i

…

Origine inconnue, deux AS-Path identiques

*	9.20.0.0/17	204.42.253.253	0 267 1225 2914 2685 2686 ?
*		195.211.222.6	0 5409 2686 ?
*		195.66.225.254	0 5459 2686 ?

Origine externe

*	130.63.0.0	204.42.253.253	0 267 1225 701 549 802 e
*		158.43.206.96	0 1849 702 701 549 802 e

exemple avec history et dampening

h	24.48.44.0/22	204.42.253.253	0 267 1225 701 7843 i
h		204.70.4.89	0 3561 701 7843 i
*d		134.55.24.6	0 293 701 7843 i

Description des politiques

- Nécessité d'avoir une trace des politiques
- bases de données auprès des « registrars » régionaux
  - (RIPE, ARIN, …)
- langage de description de politique

qui annonce quoi à qui ?

Exemple base Ripe

- consultable via whois

```
whois -h whois.ripe.net 130.79.0.0
% This is the RIPE Whois query server #3.
% The objects are in RPSL format.
% Information related to '130.79.0.0 - 130.79.255.255'
…
inetnum:      130.79.0.0 - 130.79.255.255
netname:      FR-OSIRIS
descr:        Centre Reseau et Communication, Universite Louis Pasteur
descr:        7 rue Rene Descartes, 67084 Strasbourg CEDEX, France
country:      FR
admin-c:      PDA1-RIPE
tech-c:       PG25-RIPE
```


AS2259 (Osiris)

Extrait de la base RIPE whois.ripe.net

```
aut-num: AS2259
as-name: FR-U-STRASBOURG
descr: FR
import: from AS2200 action pref=100; accept ANY
export: to AS2200 announce AS2259
default: to AS2200 action pref=10; networks ANY
admin-c: GR1378-RIPE
tech-c: GR1378-RIPE
mnt-by: RENATER-MNT
```

Fait partie de

```
as-set: AS-RENATER
```

AS553 (Belwue)

```
aut-num: AS553
as-name: BELWUE
descr: Landeshochschulnetz Baden-Wuerttemberg (BelWue)
descr: Academic Network of the
descr: Federal State of Baden-Wuerttemberg
import: from AS174 action pref=100; accept AS-COGENT
import: from AS286 action pref=100; accept AS-KQ
…
import: from AS2259 action pref=100; accept AS2259
…
export: to AS2259 announce AS-BELWUE
```


AS2200 (Renater)

```
aut-num: AS2200
as-name: FR-RENATER
descr: Reseau National de telecommunications pour la Technologie
descr: l'Enseignement et la Recherche
descr: FR
import: from AS-SFINX-PEERS action pref=190; accept <^AS-SFINX-PEERS .$>
import: from AS5511 action pref=300; accept ANY   /* Opentransit*/
import: from AS7500 action pref=190; accept AS7500  /* M root server Wide Japon */
import: from AS9270 action pref=250; accept AS-KOREN  /* réseau recherche Coréen */
import: from AS20965 action pref=300; accept AS-GEANTNRN AS-AUCS
import: from AS2470 action pref=100; accept AS2470  /* La Réunion */
import: from AS-RENATER accept <^AS-RENATER+$>
export: to AS-SFINX-PEERS announce AS-RENATER
export: to AS5511 announce AS-RENATER AS7500
export: to AS9270 announce AS-RENATER AS-GEANTNRN
export: to AS12654 announce AS-RENATER
export: to AS21357 announce AS-RENATER
export: to AS20965 announce AS-RENATER AS-KOREN AS7500
export: to AS2470 announce AS-RENATER
export: to AS-RENATER announce ANY
```


AS2200 (Renater) (suite)

```
as-set:     AS-RENATER
descr:      RENATER
members:    AS261, AS513, AS775, AS776, AS777, AS779, AS781
members:    AS782, AS789
members:    AS1300, AS1301, AS1303, AS1304, AS1307, AS1309
members:    AS1707, AS1708, AS1712, AS1715, AS1716, AS1717, AS1719
members:    AS1721, AS1723, AS1724, AS1725, AS1726
members:    AS1935, AS1936, AS1937, AS1938, AS1939, AS1940, AS1941
members:    AS1942, AS1943, AS1945, AS1951
members:    AS2060, AS2061, AS2071, AS2072, AS2085, AS2089
members:    AS2187, AS2188, AS2194, AS2198, AS2199
members:    AS2200, AS2202, AS2222, AS2223, AS2231, AS2236, AS2239
members:    AS2258, AS2259, AS2264, AS2269
members:    AS2418, AS2422, AS2426, AS2439, AS2445, AS2457, AS2462
members:    AS2470, AS2485
members:    AS13019
…
```

AS5511 (Opentransit)

```
aut-num:      AS5511
as-name:      OPENTRANSIT
descr:        France Telecom
descr:        Worldwide IP Backbone
…
import:       from AS2200 action pref=10 accept AS-RENATER
…
export:       to AS2200 announce ANY
…
import:       from AS2914 action pref=25 accept AS-VERIO
…
export:       to AS2914 announce AS-OPENTRANSIT

AS-OPENTRANSIT : backbone and customers
```


AS2914 (Verio- NTT)

- Dépend de ARIN
  American Registry for Internet Numbers

BGP mise en oeuvre

- Langage de commande
  - Définition des sessions (simple)
  - Définition de la politique (compliqué)
    - Filtrage/manipulation conditionnels d'attributs
- CLI (Command Line Interface) cisco
  ~ CLI de quagga (ex zebra)

Démarrage BGP (1)

- Configurer interfaces, routage IP
- Lancer processus BGP
  - **Router BGP** <ASN_local>
- Déclarer les sessions iBGP et eBGP
  - **Neighbor** <adresse-IP-voisin> **remote-as** <ASN>
    - Si ASN = ASN_local alors iBGP sinon eBGP

Démarrage BGP (2)

- Spécifier préfixes redistribués, ex :
  - **Redistribute connected**
  - **Redistribute RIP (ou ospf, …)**
  - Et/ou indiquer que des préfixes sont internes :
    - **Network** <adresse-réseau> [mask *masque*] [route-map *nom*]
      - Annoncés que s'ils sont accessibles
  - ces préfixes auront cet AS comme origine
    - attribut **origin ?** ou **i**

Démarrage BGP (3)

- Synchronisation
  - [**no**] **synchronization**
  - Attend (ou non) annonce IGP avant d'annoncer par BGP
- Redistribuer BGP dans IGP, par exemple
  - **router RIP**
    - **redistribute BGP**
  - Attention à ne pas redistribuer en rond
    - BGP => IGP => BGP (filtres, commande **network**)
- Normalement à ce stade
  - Routage BGP fonctionne
  - Aucune politique, critère essentiel : AS-Path

Politique (1)

- Access list (ACL) spécifiques aux attributs/ annonces
  - **Distribute-list in|out**
    - Filtrage sans modification
    - Note : risque d'inaccessibilité
- Route-map : expressions conditionnelles
  - Conditions (**match**)
  - Modifications attributs (**set**)

Exemple filtrage (1)

Router BGP 1

Neighbor 192.168.10.1 **remote-as** 2
Neighbor 192.168.10.1 **distribute-list** 1 **out**
…
Access-list 1 **deny** 192.168.20.0 0.0.1.255
Access-list 1 **permit** 0.0.0.0 255.255.255.255

- Autorise annonce tous préfixes ≠ 192.168.20.0/23
- **deny all** implicite à la fin d'une ACL

Exemple filtrage (2)

- Filter-list : filtre sur divers attributs
- Ex : Filtrage sur AS-Path (filtrer AS200)

neighbor 2.2.2.2 **filter-list** 1 **out**

ip as-path access-list 1 **deny** ^200$
ip as-path access-list 1 **permit** .*

Noter la syntaxe analogue aux expressions unix
^200$ = chaîne commençant et finissant par 200

- Applicable en **in** ou **out**

filtrage

- filtrer une annonce
  - en entrée => interdire trafic sortant
  - en sortie => interdire trafic entrant
    - et donc transit
- s'il n'y a pas d'autre annonce pour P
  - => peut rendre P inaccessible
  - => à manier avec prudence
- cisco : possibilité d'annonces conditionnelles
  neighbor A.B.C.D **advertise-map** ADVERTISE **non-exist-map** NON-EXIST

route-map (1)

- Format général

Route-map nom **permit**|**deny** num_1
Match condition_1
Set action_1
Route-map nom **permit**|**deny** num_2
Match condition_2
Set action_2
…
Route-map nom **permit**|**deny** num_n
[**match** condition_n]
Set action_n
/* deny implicite en fin */

route-map (2)

- Liste interprétée ordre des num_n
  - Jusqu'à un **match** satisfait (ou fin liste)
  - Si **match** condition_i et **permit**
    - => appliquer action_i (**set**) éventuelle et sortir (garder annonce)
  - Si **match** condition_i et **deny**
    - => éliminer annonce et sortir
      - donc un **set** est inutile avec un **deny**
  - Sortie fin de liste sans match
    - => deny implicite

Exemples de match

Match as-path liste-as
ip as-path access-list liste-as {**permit**|**deny**} AS_ER
…
Match community liste-communauté
Ip community-list liste_communauté {**permit**|**deny**} C1
IP community-list liste_communauté {**permit**|**deny**} C2
…
Match ip address liste_adresses
Match metric nombre
Match ip next-hop liste_adresses

Exemples de set

Set as-path prepend ASN1 ASN2…
ajoute ASN1 ASN2 au début de l'as path
Set metric nombre
positionne le MED
Set origin {**igp**| **egp** asn | **incomplete**}
Set local-preference nombre

Utilisation route-map

Comme filtre en entrée ou sortie vers pair BGP
Neighbor adresse **remote-as** Asn
Neighbor adresse **route-map** nom_map **in**| **out**
les annonces reçues/envoyées au voisin sont filtrées/modifiés par nom_map

Next-hop

- Normalement Next-Hop
  - routeur émetteur en eBGP
  - recopié de eBGP en iBGP
  - => pourrait ne pas être accessible
    neighbor *adresse_voisin* next-hop-self
  - En iBGP utilisation adresse loopback
    - moins vulnérable panne

Route Reflector

- Permet de ne pas avoir de full mesh en iBGP
  - => limite le nombre de sessions
- un routeur est client d'un RR :
  router BGP ASNi
  neighbor add1 **remote-as** ASNi
  neighbor add1 **route-reflector-client**
  - idem clien ts add2, add3, …
  - Le routeur add1 ouvre une seule session iBGP
  - le RR redistribue entre ses clients,
    - et avec autres voisins BGP
    - possibilité plusieurs RR en full mesh

Confédération d'AS

- Permet d'étendre des fonctions iBGP
  - à travers plusieurs AS
  - local pref, next hop, MED
    - transmis par eBGP « spécial »
  - politique concertée de plusieurs AS
    - ou un seul découpé ?

Autres commandes

- **clear ip bgp** [* | adresse] [**soft** [**in** | **out**]]
  - reset de la (les) session
    - en entrée ou en sortie
    - permet de prendre en compte les changements de configuration de BGP
      - BGP : hard state
      - nécessite de retransmettre **toutes** les annonces
    - **soft** : permet de ne pas tout remettre à zéro
      - coûteux en mémoire pour **in**
      - => mémoriser toutes les annonces reçues
        - » car annonce filtrée peut devenir non filtrée
        - » amélioré si « soft route refresh capability » :
        - » routeur redemande annonces au voisin

sho …

- **sho ip bgp** *paramètres*
  - affiche les tables et configuration BGP
  - routage *adresse*
  - voisins *neighbor*
- sho ip route
  - affiche routes, y compris apprises par BGP

debug

- debug ip bgp paramètre
  - ex :

    debug ip bgp updates
  - attention au volume de debug …

Quelques problèmes classiques

- Pour communiquer
  - Routage doit être correct dans les deux sens
  - Redistribution inter vers intra et réciproquement
  - Accessibilité des réseaux d'interconnexion

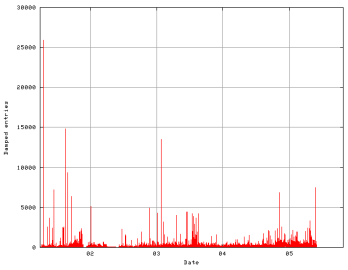
Problème de BGP

- Black hole
  - P $\in AS_i$ (qui l'annonce)
  - P annoncé à tort par AS_j
  - $\forall AS_k$ plus proche de AS_j que de AS_i
  - => trafic vers P passant par AS_k« perdu »
  - Filtrer explicitement préfixes annoncés par clients

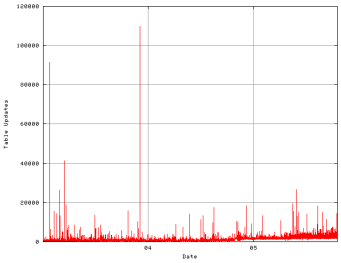
Problèmes BGP (2)

- Oscillation de routes (« route flapping »)
  - Lien physique vers P intermittent
    - Annonce / retrait de P propagé vers Internet
    - Coût en messages et calcul
  - Solution
    - Détecter route-flapping (maintenir historique par P)
    - Si flapping « punir » (« dampening »)
      - Ne pas propager annonce (P) pendant délai
      - Continuer de surveiller (annonce/withdraw)
    - + le dampening est proche du problème + efficace

Nb entrées avec dampening

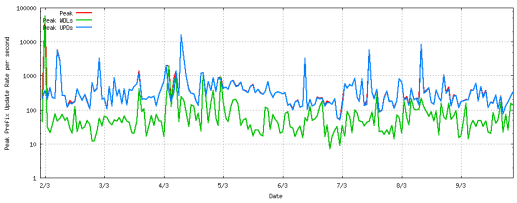


Udpdate par heure



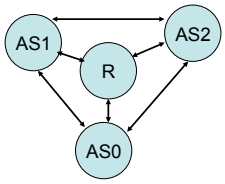
Moyenne : 1676 new, 596 changt, 2287 withdraw

Hourly Peak of Per-Second Updated and Withdrawn Prefix Rate



Problème de convergence

- Exemple :



Etat initial 0(*R, 1R, 2R) 1(0R,*R,2R) 2(0R, 1R, *R)

Convergence (1)

- Notation
  - Asi -> Asj { AS-Path | W } W = withdraw
- On considère un préfixe P accessible seulement par R
- R perd ce préfixe
  - R -> 0 W
  - R -> 1 W
  - R -> 2 W

Convergence (2)

- Traitement R -> 0 W R -> 1 W R -> 2 W
  0(∞, *1R, 2R) 1(*0R, ∞, 2R) 2(*0R,1R, ∞)
- Emission
  0 -> 1 01R, 0-> 2 01R, 1 -> 0, 10R 1-> 2 10R, 2 -> 0 20R, 2 -> 1 20R
- Traitement 0 -> 1 01R, 0-> 2 01R
  1(∞, ∞, *2R) 2(01R, *1R, ∞)
- Emission
  1 -> 0 12R, 1 -> 2 12R, 2 -> 0 21R, 2 -> 1 21R
- Traitement 1 -> 0, 10R 1-> 2 10R
  0(∞, ∞, *2R) 2(*01R, 10R, ∞) (ex æquo : ordre as)
- Emission 0 -> 1 02R, 0 -> 2 02R, 2 -> 0 201R, 2 -> 1 201R
- …. 48ème étape
  0(∞, ∞, ∞), 1 (∞, ∞, ∞), 2 (∞, ∞, ∞)
- Updates inutiles émis vers extérieur 47 étapes

Biblio

• [BGP] RFC1771, - A Border Gateway Protocol 4 (BGP-4), mars 1995.
• [CIDR] Fuller, V., Li, T., Yu, J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): An Address Assignment and Aggregation Strategy", RFC1519, September 1993.
• [RFC3513] IP Version 6 Addressing Architecture
• [MBGP] RFC 2858 - Multiprotocol Extensions for BGP-4, juin 2000.

et le tutoriel de Sam Halabi

• Internet Routing Architectures, Second Edition by Sam Halabi, Danny McPherson, Cisco Press
• la doc cisco en ligne, configuration BGP
http://www.cisco.com/en/US/docs/ios/12_0/np1/configuration/guide/1cbgp.html
• Tutoriel dynagen http://www.dynagen.org/tutorial.htm

[BGP convergence] Labovitz et al, Delayed Internet Routing Convergence, IEEE TON, Vol 9, Juin 2001.