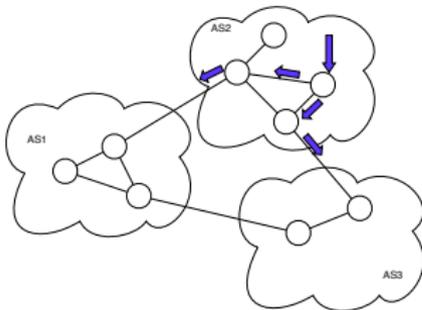


BGP — *Border Gateway Protocol*

Martin Heusse

Routage interne / externe

1. Pour les algorithmes de routage externe,
un sommet (du graphe) = un AS! (réseaux de réseaux...)
2. Problématique : les routes externes aux systèmes autonomes doivent être connues à l'intérieur

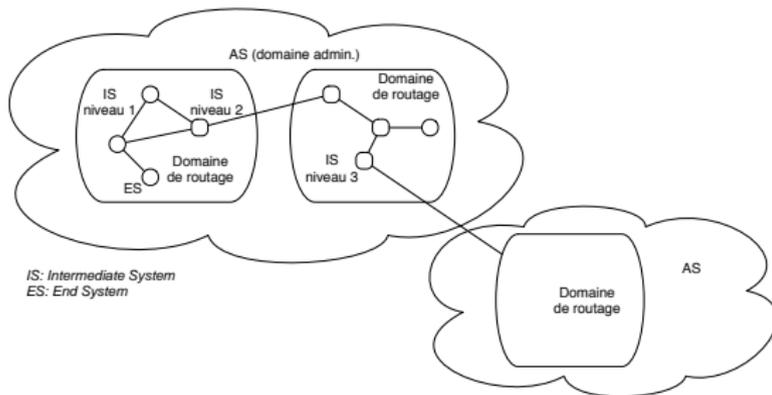


Le point 1 est traité par le protocole BGP

Le point 2 est traité par la redistribution par l'IGP (*Internal Gateway Protocol*)
des routes de BGP

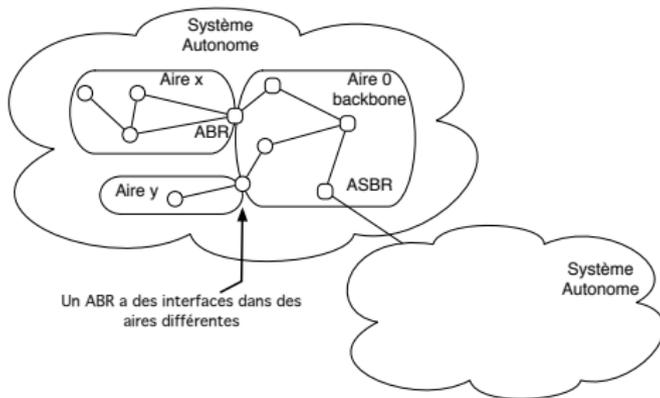
Routage interne / externe (suite)

Classification des routeurs selon l'ISO



Routage interne / externe (suite)

Classification des routeurs pour OSPF



Tout ABR doit faire partie du *backbone*

Sinon, il faut créer un lien virtuel

(évite les problèmes de convergence d'un vecteur de distance :
le réseau est en étoile)

La hiérarchie des domaines de routage

1 AS = réseau géré par une seule organisation

1. Réseau local : pas de protocole de routage sauf *rdisc* (ou parfois un IGP)
rdisc : RFC 1256 — messages ICMP type 9 (*advert.*) type 10 (*req.*)
2. Au sein de l'AS
Protocole IGP, éventuellement prise en compte de la charge (EIGRP)
Attention !
3. Routage entre AS : Protocole EGP (*External Gateway Protocol*)
Mises à jour aussi rares que possible ex. : BGP 4

BGP 4

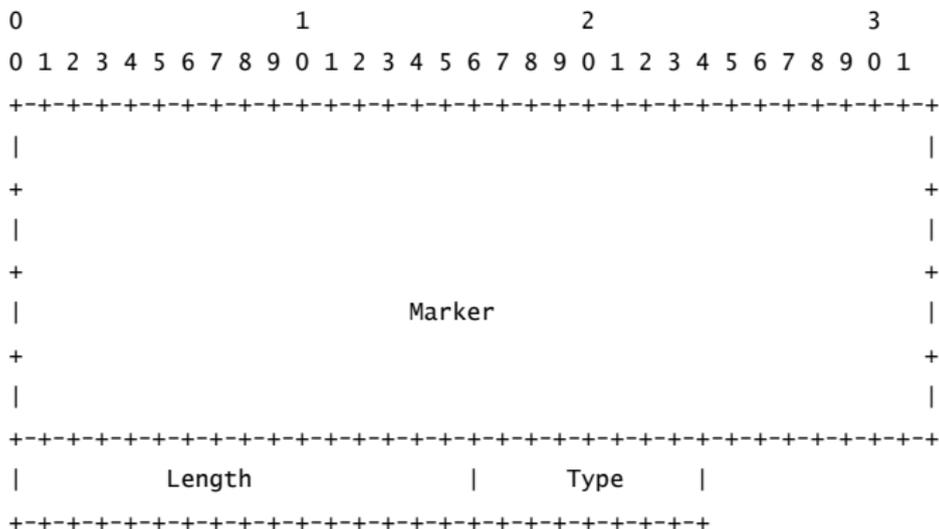
- Le protocole EGP
- Pas de notion de distance, C'est le nombre d'AS traversés qui compte
- Éventuellement, plusieurs chemins vers une destination
- La structure de données de routage est un vecteur de chemins
→ pas de cycle
- Possibilité de spécifier un filtrage des routes en fonction des contrats passés, ou des préférence (IBM veut atteindre Motorola sans passer pas un réseau opéré par intel)
- Toute publication est volontaire (commande network)

BGP, protocole de routage à vecteur de chemins

- Pour chaque destination, on connaît l'ensemble du chemin constitué de numéros d'AS (gérés par RIPE en Europe, ARIN aux USA)
- Possibilité d'avoir plusieurs chemins
- Pas de cycle, même transitoire
- Tous les échanges se font sur TCP
 - ✓ Champs « *marker* » : 16 octets à l !
 - Redondant avec le champs *length*, détecte la perte de synchronisation

BGP, protocole de routage à vecteur de chemins (suite)

Entête « paquet » BGP :



Attributs de routage BGP

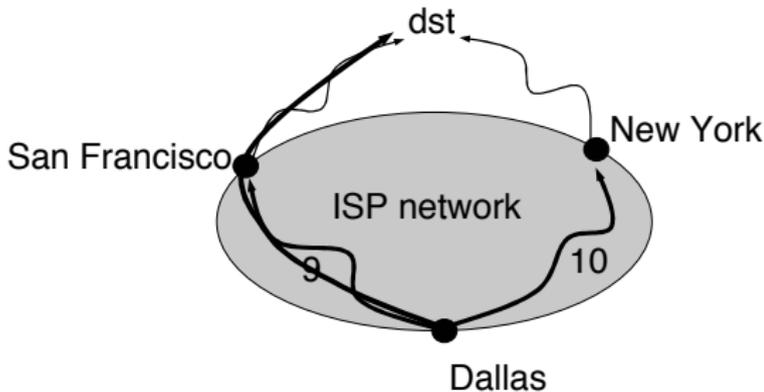
- Attributs obligatoires pour une route annoncée par BGP
 - ✓ La liste des AS traversés
 - ✓ l'adresse du prochain routeur
 - ✓ le mode d'acquisition de la route : apprise d'un IGP, d'un EGP ou autre
- Optionnels :
 - ✓ Préférence locale
 - ✓ Indication d'origine d'agrégation
 - ✓ Métrique (*multi exit discriminator*)(comparée seulement pour deux routes provenant du même AS)
 - ✓ Communauté (marquage commun à un ensemble d'adresses, utilisé pour le filtrage...)

Choix d'une route parmi plusieurs : par ordre de priorité

1. Le «prochain routeur» doit être accessible
2. Si la route vient de l'AS local, utiliser la route IGP
(si inconnue de l'IGP, on *n'utilise pas* une éventuelle route BGP)
3. Routeur de plus grand poids (parmi tous les voisins : attribut local au routeur)
4. Préférence locale (attribut propagé par iBGP)
choix de la direction pour sortir de l'AS
5. Chemin le plus court (nombre d'AS)
✓ Annonces multiples d'un AS \Rightarrow Allonge la liste d'AS...
6. Code d'origine le plus bas (IGP < EGP < INCOMPLETE)
7. Plus petite métrique (*multi exit discriminator*)(Utilisé seulement si l'AS suivant est le même pour toutes les routes)
8. Route externe plutôt que route interne

Choix d'une route parmi plusieurs : par ordre de priorité (suite)

9. Route pour laquelle le «prochain routeur» est le plus proche (métrique IGP)(Routage *hot potato*)

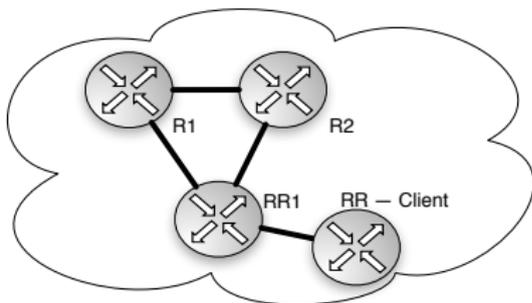


BGP : important

- Les routeurs voisins communiquent par TCP. Les voisins sont explicitement déclarés à la configuration du routeur. Les échanges commencent après l'établissement de la connexion.
- Voisinage interne / externe :
iBGP : le protocole BGP entre routeurs d'un même AS.
Ils forment un graphe complet !
- Les connexions internes sont généralement établies avec des interfaces de *loopback*
⇒ connexions indépendantes de l'état des liens «intérieurs» du routeur

BGP administration

- *Peer-group* : ensemble de routeurs partageant la même politique (une seule configuration pour le groupe)
- Réflecteurs (*route reflectors*) : diminue la taille de la clique de pairs iBGP
Les routeurs non associés à un réflecteur doivent former un graphe complet



Un exemple :

```
> show ip rout
```

```
Codes: C- connected, S- static, I- IGRP, R- RIP, M- mobile,  
B- BGP, D- EIGRP, EX- EIGRP external, O- OSPF,  
IA- OSPF inter area,  
N1- OSPF NSSA ext type 1, N2- OSPF NSSA ext type 2  
E1- OSPF external type 1, E2- OSPF ext type 2,  
i- IS-IS, L1- IS-IS level-1, L2- IS-IS level-2,
```

Gateway of last resort is 193.44.182.111 to network 0.0.0.0

```
O 195.220.30.0/24 [110/20] via 193.44.182.122, 04:54:05, Vlan5  
O IA 193.54.242.0/24 [110/21] via 193.44.182.124, 02:14:46, Vlan5  
O E2 194.199.31.0/24 [110/1] via 193.44.182.120, 02:14:46, Vlan5  
152.77.0.0/16 is variably subnetted, 8 subnets, 4 masks  
O E1 152.77.208.0/24 [110/71] via 193.44.182.124, 02:14:46, Vlan5  
O E2 152.77.5.128/25 [110/20] via 193.44.182.124, 02:14:46, Vlan5  
O E2 152.77.212.83/32 [110/2] via 193.44.182.124, 01:12:40, Vlan5  
O IA 152.77.0.0/16 [110/21] via 193.44.182.124, 02:14:46, Vlan5  
[etc.]
```

129.88.0.0/16 is variably subnetted, 36 subnets, 4 masks
O 129.88.1.250/32 [110/9] via 129.88.4.1, 05:55:09, Vlan54
C 129.88.253.0/24 is directly connected, Vlan30
C 129.88.1.254/32 is directly connected, Loopback0
C 129.88.69.0/24 is directly connected, Vlan15
O IA 129.88.103.0/24 [110/31] via 193.44.182.124, 00:44:38, Vlan5
R 129.88.25.0/24 [120/3] via 129.88.6.1, 00:00:29, Vlan59
R 129.88.24.0/24 [120/3] via 129.88.6.1, 00:00:29, Vlan59
C 129.88.27.0/24 is directly connected, Vlan7
R 129.88.26.0/24 [120/1] via 129.88.6.1, 00:00:29, Vlan59
C 129.88.29.0/24 is directly connected, Vlan14
C 129.88.28.0/24 is directly connected, Vlan8
C 129.88.31.0/24 is directly connected, Vlan11

Table de routage BGP

show ip bgp 129.88.0.0 sur e513-e-rci76-1.cern.ch

BGP routing table entry for 129.88.0.0/16, version 68976

Paths: (6 available, best #5, table Default-IP-Routing-Table)

Multipath: eBGP

Advertised to update-groups:

9 10 11 12

11537 11537 20965 2200 1942

198.32.11.85 from 198.32.11.85 (64.57.28.242)

Origin IGP, metric 100, localpref 100, weight 100, valid, external

Community: 2200:1001 2200:2002 2200:2200 11537:2501 20965:155

11537 20965 2200 1942, (received-only)

198.32.11.85 from 198.32.11.85 (64.57.28.242)

Origin IGP, metric 0, localpref 100, valid, external

Community: 2200:1001 2200:2002 2200:2200 11537:2501 20965:155

3356 3356 5511 2200 1942

213.242.73.145 from 213.242.73.145 (4.68.0.93)

Origin incomplete, metric 1000, localpref 10, weight 100, valid, external

Community: 513:92

3356 5511 2200 1942, (received-only)

Table de routage BGP

show ip bgp 129.88.0.0 sur e513-e-rci76-1.cern.ch

(suite)

213.242.73.145 from 213.242.73.145 (4.68.0.93)

Origin IGP, metric 0, localpref 100, valid, external

Community: 3356:2 3356:86 3356:501 3356:666 3356:2065

559 20965 2200 1942

192.65.184.221 from 192.65.184.221 (130.59.32.42)

Origin IGP, metric 10, localpref 100, weight 100, valid, external, best

Community: 2200:1001 2200:2002 2200:2200 20965:155

559 20965 2200 1942, (received-only)

192.65.184.221 from 192.65.184.221 (130.59.32.42)

Origin IGP, localpref 100, valid, external

Community: 2200:1001 2200:2002 2200:2200 20965:155

Paquets BGP

- Ouverture
- Mise à jour
- Erreur
- Sonde (émis périodiquement)
- Tout en TCP