RFC 9234 : Route Leak Prevention and Detection Using Roles in UPDATE and OPEN Messages

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 25 mai 2022

Date de publication du RFC : Mai 2022

https://www.bortzmeyer.org/9234.html

Vous savez sans doute que l'Internet repose sur le protocole de routage BGP et que ce protocole a été cité dans des accidents (voire des attaques) spectaculaires, par exemple lorsqu'un routeur BGP annonçait ou réannonçait des routes qu'il n'aurait pas dû annoncer (ce qu'on nomme une fuite de routes). Si BGP lui-même est pair-à-pair, en revanche, les relations entre les pairs ne sont pas forcément symétriques et des règles sont largement admises ; par exemple, on n'est pas censé annoncer à un transitaire des routes apprises d'un autre transitaire. Ce RFC étend BGP pour indiquer à l'ouverture de la session le type de relations qu'on a avec son pair BGP, ce qui permet de détecter plus facilement certaines fuites de routes.

La solution proposée s'appuie sur un modèle des relations entre opérateurs, le modèle « sans vallée » ("*valley-free*" <https://blog.ipspace.net/2018/09/valley-free-routing.html>, même si ce terme n'est pas utilisé dans le RFC). On le nomme aussi modèle Gao-Rexford <https://ieeexplore.ieee.org/document/974523>. Ce modèle structure ces relations de telle façon à ce que le trafic aille forcément vers un opérateur aussi ou plus important (la « montagne »), avant de « redescendre » vers la destination ; le trafic ne descend pas dans une « vallée », sauf si elle est la destination finale. Le but est de permettre un routage optimum et d'éviter boucles et goulets d'étranglement, mais ce modèle a aussi une conséquence politico-économique, maintenir la domination des gros acteurs (typiquement les "*Tier-1*"). Les relations entre participants à une session BGP sont de type Client-Fournisseur ou Pair-Pair, et le trafic va toujours dans le sens du client vers le fournisseur, sauf pour les destinations finales (on n'utilise donc pas un pair pour du transit, c'est-à-dire pour joindre d'autres réseaux que ceux du pair).

En raison de ce modèle, un routeur BGP n'est pas censé annoncer à un transitaire des routes apprises d'un autre transitaire, ou annoncer à un pair des routes apprises d'un transitaire, ou bien apprises d'un autre pair (RFC 7908 [1]). Une fuite de routes est une annonce qui viole cette politique (qu'elle soit explicite ou implicite).

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc7908.txt

Traditionnellement, on empêche ces fuites par des règles dans la configuration du routeur. Par exemple, le client d'un transitaire va bloquer l'exportation par défaut, puis lister explicitement les routes qu'il veut annoncer. Et le transitaire va bloquer l'importation par défaut, et lister ensuite les routes qu'il acceptera de son client. Ce nouveau RFC ajoute un autre mécanisme, où la relation entre les deux partenaires BGP est explicitement marquée comme Client-Fournisseur ou Pair-Pair, facilitant la détection des fuites. Un concept qui était auparavant purement "*business*", le type de relation entre acteurs BGP, passe ainsi dans le protocole.

Le RFC utilise donc ces termes, pour désigner la place d'un acteur BGP dans le processus de routage :
— Fournisseur ("*Provider*") : peut envoyer n'importe quelle route (un transitaire est un Fournisseur).
— Client ("*Customer*") : ne peut envoyer que ses propres routes, ou bien celles apprises d'un de ses clients (l'opérateur qui est Client pour une relation BGP peut être Fournisseur pour une autre).
— Serveur de routes ("*Route Server*") : peut envoyer toutes les routes qu'il connait à ses clients.
— Client d'un serveur de routes ("*Route Server Client*") : ne peut envoyer au serveur de routes que ses propres routes, ou bien celles apprises d'un de ses clients.
— Pair ("*Peer*") : ne peut envoyer à un de ses pairs que ses propres routes, ou bien celles apprises d'un de ses clients.

Une violation de ces règles est une fuite de routes (RFC 7908). Le RFC précise qu'il s'agit de relations « techniques », qu'elles n'impliquent pas forcément une relation "*business*" mais, en pratique, l'échange entre pairs est souvent gratuit et la relation Client-Fournisseur payante. Le RFC précise aussi qu'il peut exister des cas plus complexes, comme une relation Client-Fournisseur pour certains préfixes IP et Pair-Pair pour d'autres.

Bon, assez de politique et de "*business*", place à la technique et au protocole. Le RFC définit une nouvelle capacité (RFC 5492) BGP, "*Role*" <https://www.iana.org/assignments/capability-codes/capability-codes.xml#capability-codes-2>, code 9 et dont la valeur, stockée sur un octet, indique un des rôles listés ci-dessus <https://www.iana.org/assignments/capability-codes/capability-codes.xml#bgp-role-value> (0 pour Fournisseur, 3 pour Client, 4 pour Pair, etc). D'autres rôles pourront être ajoutés dans le futur (politique « Examen par l'IETF », cf. RFC 8126). Ce rôle indique la place de l'AS et est configuré avec la session BGP (rappelez-vous qu'un AS peut être Client d'un côté et Fournisseur de l'autre). À l'ouverture de la connexion, les routeurs vérifient la compatibilité des rôles (si tous les deux annoncent qu'ils sont Fournisseur, c'est un problème, d'où un nouveau code d'erreur BGP, "*Role mismatch*" <https://www.iana.org/assignments/bgp-parameters/bgp-parameters.xml#bgp-parameters-6>).

Une fois que les deux routeurs sont d'accord sur les rôles respectifs de leurs AS, les routes annoncées peuvent être marquées avec un nouvel attribut, OTC ("*Only To Customer*" <https://www.iana.org/assignments/bgp-parameters/bgp-parameters.xml#bgp-parameters-2>, code 35), dont la valeur est un numéro d'AS. Une route ainsi marquée ne sera acceptée que si elle vient d'un Fournisseur, ou d'un pair dont le numéro d'AS coïncide. Autrement, elle sera considérée comme résultant d'une fuite.

Ah, et pour les cas compliqués dont on a parlé plus haut, comme les rôles différents pour chaque préfixe ? Dans ce cas-là, le mécanisme des rôles de ce RFC n'est pas adapté et ne doit pas être utilisé.

Ce système des rôles est apparemment mis en œuvre dans des patches (non encore intégrés ?) pour FRRouting et BIRD. Par exemple, pour FRRouting, on configurerait :

```
neighbor 2001:db8:999::1 local-role customer
```

Pour BIRD, cela serait :

```
configuration
    ...
    protocol bgp {
        ...
        local_role customer;
    }
```

Et la gestion des rôles devrait être bientôt dans BGPkit <https://twitter.com/heymingwei/status/1529223382822055936>. Et une version antérieure à la sortie du RFC tourne sur les routeurs MikroTik (voir leur documentation <https://help.mikrotik.com/docs/display/ROS/BGP>, `local.role ebgp-customer`...). Par contre, je ne sais pas pour Cisco ou Juniper. Si quelqu'un a des informations...