RFC 7908 : Problem Definition and Classification of BGP Route Leaks

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 8 juillet 2016

Date de publication du RFC : Juin 2016

https://www.bortzmeyer.org/7908.html

Il est bien connu que le protocole de routage BGP connait fréquemment des « fuites de route » ("*route leaks*"). Les forums où les opérateurs discutent sont plein de discussions et d'alertes sur la dernière fuite en cours, il existe un compte Twitter dédié pour les alertes automatiques de fuites <https://twitter.com/bgpstream>, le rapport sur la résilience de l'Internet en France <https://www.afnic.fr/fr/l-afnic-en-bref/actualites/actualites-generales/9991/show/l-observatoire-de-l-html> en parle longuement. Bref, le sujet est bien connu. Curieusement, il n'y a pas de définition simple et consensuelle de ce qu'est une fuite. Il en existe en fait plusieurs sortes, et ce nouveau RFC tente de faire le point et de définir rigoureusement les différents types de fuite.

La fuite est en effet un sérieux problème : des routeurs vont recevoir des routes incorrectes et, s'ils les acceptent, le routage normal peut être perturbé, menant à des pannes, ou à du "*tromboning*" (paquets IP routés par un chemin bien plus long que l'optimum).

La liste des incidents de routage médiatiquement connus est longue : le détournement de YouTube par Pakistan Telecom <https://www.bortzmeyer.org/pakistan-pirate-youtube.html>, la fuite venue de Malaisie <https://www.bortzmeyer.org/bgp-malaisie.html>, et bien d'autres décrites dans l'abondante bibliographie de notre RFC (cf. section 1 du RFC). L'IETF travaille depuis longtemps à des solutions techniques à ce problème (notamment via son groupe de travail SIDR <https://tools.ietf.org/wg/sidr>) mais, pour résoudre un problème, il vaut mieux bien le comprendre. C'est le travail de ce nouveau RFC, qui tente une taxonomie des fuites.

Donc, d'abord, une (tentative de) définition (section 2 du RFC). **Une fuite est la propagation de l'annonce d'une route au-delà de la portée prévue.** Ce n'est donc pas la seule annonce qui est le problème, c'est sa propagation. Si le Pakistan veut annoncer les routes de YouTube dans le pays, à des fins de censure, ce n'est pas une fuite de route, mais un détournement délibéré (comme celui de Google Public DNS

en Turquie <https://www.bortzmeyer.org/dns-routing-hijack-turkey.html>). La fuite a commencé quand Pakistan Telecom a bêtement laissé se propager cette annonce au monde entier.

La définition s'appuie donc sur une **politique de routage** qui définit « ce qui est prévu ». La politique de routage est décidée par chaque acteur (l'Internet n'a pas de Chef, chacun est maître de sa politique) et mise en œuvre dans la configuration des routeurs, sous forme de règles disant ce qui est accepté et ce qui est refusé. Si les opérateurs ne commettaient jamais d'erreurs, on pourrait dire que ces règles suffiraient à décrire la politique de routage. Mais ils en commettent (par exemple, le transitaire de Pakistan Telecom aurait dû n'accepter de son client qu'un ensemble fini de routes, celles correspondant aux préfixes alloués à Pakistan Telecom et à ses clients). Une fuite de route est donc l'équivalent pour les opérateurs réseau de ce qu'est une bogue pour les programmeurs : « ce n'est pas ce que je voulais, mais c'est ce que j'ai dit ».

La définition de la politique de routage dépend des relations entre acteurs (on n'accepte pas la même chose d'un client, d'un transitaire, d'un pair...) Le RFC cite (section 2) plusieurs études expliquant ces relations compliquées.

Au fait, on a supposé que les « fuites de routes » étaient mauvaises et qu'il fallait les combattre. Mais pourquoi ? Parce qu'elles peuvent mener à des pannes (la route annoncée à tort ne fonctionne pas) ou à des chemins sous-optimaux (un long détour). Cela concerne les fuites accidentelles, de loin les plus nombreuses (il y a davantage de maladroits que de malveillants). Mais il y a aussi des fuites délibérées, provoquées pour faire une attaque par déni de service, ou bien pour forcer le trafic à passer en un point où il pourra être plus facilement surveillé.

C'est pour cela que les gens qui ne chiffrent pas leurs communications avec des arguments du genre « non, mais ça ne sort pas du pays, de toute façon », ont gravement tort. La vulnérabilité du routage fait que leur trafic peut soudainement partir bien plus loin.

Voyons maintenant la classification des fuites (section 3). Le RFC les classe en différents types, identifiés par un numéro. Le type 1 est « virage en épingle à cheveux avec un préfixe complet ». C'est ce qui se produit quand un AS de bordure apprend un préfixe d'un de ses transitaires et le ré-annonce à un autre transitaire (d'où l'épingle à cheveux). Le second transitaire, s'il ne filtre pas les préfixes que peut annoncer son client, risque fort d'accepter la route (préférence donnée aux routes des clients sur celles des fournisseurs) et le trafic sera donc envoyé à l'AS de bordure (qui pourra ou pas le gérer). Ce fut le cas de l'incident Dodo <http://labs.apnic.net/blabs/?p=139/>, ainsi que de la fuite en Malaisie citée plus haut.

Le type 2, lui, est « fuite latérale ». Un acteur reçoit une route d'un pair et la transmet à un autre pair. (Voir la conférence de Jared Mauch à NANOG <https://www.nanog.org/meetings/nanog41/presentations/mauch-lightning.pdf>, où il observe qu'il n'est pas facile de détecter automatiquement ces fuites, car les relations entre acteurs peuvent être compliquées.)

L'incident de type 3, lui, se produit lorsque un AS apprend d'un transitaire des routes qu'il annonce à son pair (normalement, on n'annonce à un pair que ses routes à soi). L'exposé de Mauch en cite également des exemples.

Le type 4 est l'inverse : un pair laisse fuir les préfixes d'un pair vers un transitaire. Ce fut le cas de l'incident Moratel contre Google <https://blog.cloudflare.com/why-google-went-offline-today-and-a>, ou d'un problème frappant Amazon <https://blog.thousandeyes.com/route-leak-causes-amazon-a

Le type 5 se nomme « ré-origination ». L'AS maladroit annonce les routes comme si elles venaient de lui, et c'est son numéro d'AS qui apparait comme origine (le début du chemin d'AS). Ce fut le cas lors de grande fuite chinoise de 2010 <http://research.dyn.com/2010/11/chinas-18-minute-mystery/>, ou pendant le curieux détournement islando-biélorusse <http://research.dyn.com/2013/11/mitm-internet-hijacking/> (un des rares cas où le "*shunt*" <https://www.bortzmeyer.org/bgp-shunt.html> semblait volontaire).

Le type 6 est la « fuite de préfixes plus spécifiques ». Un AS typique annonce dans son IGP des préfixes bien plus spécifiques que ce qu'il annonce publiquement, afin de contrôler plus finement le routage interne. Une erreur (redistribution de l'IGP dans l'EGP...) et paf, les préfixes spécifiques fuient. Étant plus spécifiques que le préfixe « normal », ils seront préférés lors de la transmission des paquets. Le cas le plus spectaculaire fut celui de l'AS 701 <http://www.bgpmon.net/what-caused-todays-internet-hiccup/>.

Le RFC ne discute pas des solutions possibles, ce n'est pas son but. Les curieux pourront regarder du côté des systèmes d'alerte <https://www.bortzmeyer.org/alarmes-as.html> ou de RPKI/ROA <https://www.bortzmeyer.org/securite-routage-bgp-rpki-roa.html>.