Curiosité BGP du mois : annonces ultra-larges

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 10 avril 2011

https://www.bortzmeyer.org/annonces-bgp-larges.html

L'Internet, on le sait, est une jungle pas civilisée, pleine de dangers et où les chinois, les pédonazis, les musulmans qui sont trop nombreux et les téléchargeurs de céline-dion.mp3 commettent d'innombrables crimes tous les jours. Notamment, il est bien connu que BGP est souvent l'innocente victime de leurs amusements. Ce protocole de routage est à la fois la base technique de l'Internet et l'un de ses points les plus fragiles. Curieusement, la plupart du temps, cette vulnérabilité n'a pas de conséquence : l'Internet est très robuste et la réparation est rapide. C'est seulement si on regarde bien qu'on s'aperçoit qu'il y a effectivement plein de choses bizarres dans BGP, masquées par cette robustesse. J'ai découvert très récemment le problème des annonces BGP trop larges, lorsqu'un opérateur annonce un préfixe bien plus général que ce qu'il devrait.

Le point de départ était le débogage d'un problème de routage. Utilisant des services qui affichent les annonces BGP reçues par divers routeurs, on s'aperçoit qu'il y a une annonce pour 64.0.0.0/3... Elle est faite par Global Crossing, avec le numéro d'AS 3549. Pourquoi est-ce un problème ? Parce qu'un /3, c'est un huitième de l'Internet ! Il n'est pas possible que Global Crossing ait autant de clients (l'examen des bases du RIR ARIN ne montre d'ailleurs aucune allocation correspondante : et ce /3 couvre des RIR différents). Cette annonce est donc clairement anormale.

Regardons-la plus en détail en utilisant l'excellent service RIS <http://www.ris.ripe.net> ("*Routing Information Service*") du RIPE-NCC : <http://www.ris.ripe.net/dashboard/64.0.0.0/3>. Que voit-on ?

— Que le problème a apparemment commencé vers le 23 mars (observé à l'époque mais RIS ne garde que les deux dernières semaines) et, qu'au jour de cet article, il continue.
— Que la visibilité de cette annonce est très faible : 1 %, dit le RIS, qui l'affiche en rouge. C'est très certainement parce que la plupart des pairs BGP sérieux jettent les annonces X.Y.Z.T/N où N ¡ 8 (il n'y a jamais eu d'allocation IPv4 plus générales qu'un /8). En regardant sur les routeurs BGP de mon employeur (situé assez loin du cœur de l'Internet, donc ne voyant que des annonces déjà très filtrées), on ne voit pas 64.0.0.0/3.
— Mais cette annonce est quand même propagée. Le service de surveillance Cyclops <http://cyclops.cs.ucla.edu/> l'a vue, relayée sur le chemin 3130 2914 3549 donc il y a quand même des gens qui transmettent un /3 sans hésiter.

— Comment se fait-il que Global Crossing n'ait pas réagi ? Ils ont été notifiés presque tout de suite mais on sait bien que ce genre de grosses sociétés ne lit pas son courrier <https://www.bortzmeyer.org/abuse-ne-repond-pas.html>.

Est-ce que ce genre de « détournement » (Global Crossing n'est pas gestionnaire de ce /3 et n'a aucun droit de l'annoncer) a des conséquences pratiques ? Probablement pas, et c'est sans doute ce qui explique que Global Crossing n'ait pas réagi, et que le problème continue. D'une part, cette annonce est largement filtrée et peu de routeurs BGP la verront. D'autre part, même si elle atteint un routeur, celui-ci aura presque à coup sûr une route plus spécifique vers les destinations incluses dans ce /3. Le routage IP privilégiant les routes plus spécifiques (règle dite du "*longest match*"), ce seront elles qui seront utilisées (à noter qu'on lit parfois que c'est BGP qui privilégie les routes plus spécifiques mais c'est tout à fait faux : BGP ne fait de sélection qu'entre deux préfixes de même longueur, sinon il transmet les deux préfixes). Les seules destinations atteignables via cette annonce sont celles qui correspondent à des trous dans `64.0.0.0/3`. Elles sont sans doute rares.

Mais, si cette annonce ne sert pas à grand'chose (filtrée, ou bien ignorée suite à une route plus spécifique), pourquoi Global Crossing le fait ? Si on est d'esprit paranoïaque, on peut imaginer plein de choses désagréables (et, si Global Crossing était une entreprise chinoise, il n'y a pas de doute que Fox News ferait un article sur ces choses) mais le plus probable est qu'il s'agit simpement d'une erreur. Certes, sans ce filtrage des annonces ¡ 8 et sans la règle IP de la route la plus spécifique, ces annonces trop larges poseraient un problème de sécurité : un méchant qui annoncerait un /1 pourrait capter la moitié du trafic de l'Internet. Mais, en pratique, il ne semble pas qu'il faille s'en inquiéter, le /3 de Global Crossing n'est pas allé très loin.

À noter qu'il peut s'agir d'une « demi-erreur ». Il est fréquent qu'un opérateur annonce à ses clients directs (et uniquement à ceux-ci) des routes très larges, notamment une route par défaut (`0.0.0.0/0`). Comme il s'agit d'un lien privé, ce n'est pas en soi un problème. Parfois, l'opérateur fait la même annonce aux moniteurs des services comme le RIS (parce que c'est plus simple de ne pas avoir de cas spécial) et l'information se retrouve alors publique.

Est-ce que ce genre d'erreurs (des annonces d'un /3, /4, /5...) arrive souvent ? Je n'ai pas trouvé de traces de discussion à ce sujet. Comme pour beaucoup de trucs bizarres de l'Internet, tant qu'on ne regarde pas de trop près, on ne voit pas le problème. Dès qu'on cherche, on voit des annonces rigolotes comme celle de Swisscom annonçant `80.0.0.0/5`, préfixe qui n'est pas enregistré tel quel dans la base du RIPE-NCC et qui est certainement une autre erreur. Le phénomène est donc sans doute fréquent mais ignoré.

Profitons au moins de ce problème pour résumer quelques méthodes pour explorer les annonces BGP récentes. Même si on peut se connecter sur les routeurs BGP de son entreprise et taper des `show route`, on ne verra (sauf à travailler chez un "*Tier-1*" et se connecter sur beaucoup de routeurs) qu'une partie de l'Internet. Il est donc préférable de compter sur des services extérieurs, en général accessibles via le Web.

Pour connaître ce qui existe en ce moment même, la meilleure source, ce sont les « "*looking glasses*" ». Une bonne liste figure en <http://www.traceroute.org/#Looking%20Glass>. (Attention, cette liste change tout le temps et un certain nombre de liens sont cassés.) L'un des « "*looking glasses*" » les plus connus et les plus stables est celui de Hurricane Electric, en <http://bgp.he.net>. On peut féliciter HE pour son ouverture. Malheureusement, il fournit peu de détails sur l'annonce BGP vue.

Pour connaître ce qui a existé dans le passé, les meilleures sources sont le RIS <http://www.ris.ripe.net> du RIPE-NCC et Cyclops <http://cyclops.cs.ucla.edu/>.

Il y a aussi des services qui stockent tout et produisent de très intéressantes analyses mais ils ne vous laissent pas interroger votre préfixe de votre choix. BGPmon <http://www.bgpmon.net> et Renesys <http://www.renesys.com> sont deux bons exemples.

Comme l'Internet lui-même, cet article a nécessité la mobilisation de compétences différences donc merci à Olivier Perret pour avoir détecté et signalé le problème, à Sylvain Busson pour l'analyse des routeurs BGP, à Michel Py et Stéphane Chevalier pour leurs bonnes remarques.