Un exemple de problème dans BGP

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 24 juillet 2015

https://www.bortzmeyer.org/un-exemple-de-probleme-bgp.html

« On » me demande parfois quels outils utiliser pour analyser un problème BGP quand on n'a pas accès à un routeur de la DFZ (d'ailleurs, certaines techniques ici peuvent servir même dans ce cas). Voici un exemple avec la panne de plus de deux heures de Swift <http://www.swiftco.net/> hier 22 juillet.

Point de départ, la machine 204.13.164.192 n'est plus joignable. Bête panne d'un serveur, comme cela arrive tout le temps sur l'Internet ? Non, cette fois, c'est plus rigolo, le préfixe IP 204.13.164.0/24 qui l'englobe a disparu de la table de routage mondiale. De même que d'autres préfixes du même opérateur, comme 204.8.32.0/24 (qui héberge notamment les serveurs DNS de swiftco.net, rendant ce nom de domaine inutilisable ; on ne le répétera jamais assez, il faut mettre ses serveurs DNS dans plusieurs réseaux différents).

Comment voir ce que contient cette table de routage ? Si on a accès à un routeur qui a une table complète, on peut le faire soi-même mais, si ce n'est pas le cas, on peut utiliser un des innombrables "*looking glasses*" qui vous donnent un accès indirect à ces routeurs. Par exemple, voici ce que donne celui de Hurricane Electric <http://lg.he.net/> une fois la panne réparée (on voit les routes pour le 204.8.32.0/24, pendant la panne, on avait « "*None of the BGP4 routes match the display condition*" ») :

Qu'on utilise ce "*looking glass*" ou bien qu'on passe par un de ses routeurs à soi, on n'a qu'une vision immédiate. Il serait intéressant de pouvoir regarder le passé, notamment si on a été prévenu trop tard et qu'on veut investiguer a posteriori. C'est ce que permet RIPEstat <https://stat.ripe.net> qui fournit tout un tas d'outils d'analyse (qui ne sont pas toujours d'un abord facile). L'un des plus simples est le "*BGP Update activity*" <https://stat.ripe.net/widget/bgp-update-activity#w.starttime=2015-07-09T15%3A00%3A00&w.endtime=2015-07-23T15%3A00%3A00&w.resource=204.8.32.0%2F24>. Voici ce qu'il affichait juste après la réparation :

On y voit une grosse activité BGP vers 0810 UTC au moment où le préfixe, pour une raison inconnue, était retiré. Cette activité comporte des retraits ("*withdraw*") mais aussi des annonces ("*announce*"). C'est normal, les routeurs BGP réagissent au retrait en annonçant des routes alternatives pendant une minute ou deux, le temps que tous réalisent que le préfixe est bien retiré, qu'il n'y a pas d'alternative. Puis on voit une autre période d'activité vers 1045 UTC au moment où ça repart. Celle-ci ne comporte que des annonces.