BGP et le désormais célèbre attribut 99

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 3 septembre 2010

https://www.bortzmeyer.org/bgp-attribut-99.html

Vendredi dernier, le 27 août, une expérience de mesure BGP menée par le RIPE-NCC et Duke University a révélé une bogue latente dans le code de certains modèles de routeurs Cisco, menant à la corruption des annonces BGP échangées, qui elle-même a entrainé la coupure de plusieurs sessions BGP, et finalement des perturbations plus ou moins fortes dans une partie de l'Internet. Cette panne a réactivé des débats sur la sécurité et la stabilité de l'Internet, ou bien sur la légitimité de procéder à de telles expériences.

Je ne vais pas détailler la panne elle-même car je n'étais pas là pour l'observer, je n'ai pas d'accès à toutes les données et, de toute façon, depuis une semaine, plusieurs très bons articles sont sortis sur le sujet. La meilleure synthèse est celle du RIPE-NCC <https://labs.ripe.net/Members/erik/ripe-ncc-and-duke-university-bgp-experiment/>. Bien plus technique, pour les amateurs de décodage de paquets, l'excellent article de Tassos <http://ccie-in-3-months.blogspot.com/2010/08/decoding-ripe-experiment.html> et évidemment le RFC 4271 [1], la norme de BGP (par exemple la section 4.3, sur les attributs, voir aussi le registre IANA des attributs <https://www.iana.org/assignments/bgp-parameters/bgp-parameters.xhtml#bgp-parameters-2>). Et naturellement l'avis de sécurité de Cisco <http://www.cisco.com/en/US/products/products_security_advisory09186a0080b4411f.shtml>, qui donne plein de détails. Notons toutefois un point peu mentionné, le fait que le seul « privilège » qu'avait le RIPE-NCC est d'être bien connecté, avec beaucoup de "*peerings*". Leurs annonces se propagent donc vite et loin. Mais, autrement, n'importe lequel des 40 ou 50 000 AS actifs pouvait faire pareil (si leur(s) opérateur(s) immédiats n'ont pas de Cisco bogué et transmettent l'annonce intacte). L'« expérience » aurait donc pu être faite par n'importe quel cyber-guerrier amateur.

Première question soulevée par cette panne : l'Internet est-il vraiment trop fragile ? Allons-nous tous mourir ? Al-Qaida va t-elle détruire la capacité de l'Occident à regarder YouTube toute la journée ? Faut-il refaire l'Internet en mieux ? Ou bien faire passer les routeurs sous le contrôle de l'"*US Army*" ? Ces questions resurgissent à chaque panne, surtout lorsqu'elle affecte des services dramatiquement essentiels <https://www.bortzmeyer.org/facebook-joue-bgp.html>. Des réunions sont organisés, des colloques causent, des rapports sont écrits. Mais, la plupart du temps, l'arbre cache la forêt et on oublie les points essentiels :

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc4271.txt

— Il est très facile de perturber l'Internet, mais on est vite repéré et le problème est vite guéri. La panne du 27 août n'a duré qu'une heure. Même si le RIPE-NCC n'avait pas retiré l'annonce anormale, les ingénieurs auraient vite reconfiguré le réseau, comme cela a été le cas dans les affaires similaires <https://www.bortzmeyer.org/securite-bgp-et-reaction-rapide.html>.
— L'Internet résiste aux soubresauts non pas parce que BGP est particulièrement bien conçu mais parce qu'il y a des gens compétents et dévoués derrière les routeurs. Rigidifier les procédures, augmenter le niveau de contrôle (dans le style UIT) aggraverait le problème au lieu de le résoudre, puisque cela empêcherait ces réactions intelligentes.
— Il est facile de se vanter (et beaucoup de gens le font <https://www.bortzmeyer.org/science-et-vie-
html>) qu'on peut construire un meilleur Internet. Techniquement, c'est faisable. Mais la vulnérabilité de l'Internet ne vient pas uniquement de ses protocoles mais aussi de leurs mises en œuvre. La faille du 27 août n'était pas dans BGP mais dans IOS XR. Les tenants d'une refonte de l'Internet ont-ils trouvé la méthode pour produire du logiciel sans bogue ? Si oui, il serait intéressant qu'ils la publient. De toute façon, la vulnérabilité fondamentale de l'Internet est qu'il connecte des organisations différentes et souvent ennemies. Tout réseau mondial aurait le même problème. Donc, soit on revient à un Minitel centralisé et national, soit on accepte le fait que la mondialisation a ses bons et ses mauvais côtés.
— Peu de commentateurs ont relevé que le problème venait encore d'une bogue de Cisco. Bien sûr, d'autres marques de routeurs ont connu des bogues liées au traitement de BGP. Mais Cisco a quand même le record. Seulement, si on peut taper facilement sur le RIPE-NCC, s'attaquer à une grosse entreprise états-unienne ayant beaucoup d'avocats est plus difficile. Donc, peu de commentateurs ont osé dire qu'il fallait peut-être songer à faire des choix techniques différents et, au minimum, à diversifier les logiciels des routeurs. Ceci dit, si on veut défendre Cisco, le meilleur argument serait que les clients de ce vendeur réclament tout le temps des nouvelles fonctions, des nouveaux services et que l'accroissement du taille du logiciel qui en résulte ne va pas dans le sens de la fiabilité. Il faut savoir si on veut, de la stabilité ou bien le dernier truc à la mode...
— Et enfin, un autre point doit être rappelé, lorsque je lis certaines indignations « Comment est-ce possible ? » ou « Que fait le gouvernement ? ». L'Internet n'est pas et ne doit pas être une infrastructure **vitale**. Bien sûr qu'il est important (par exemple, c'est lui qui justifie mon salaire) mais il n'y a pas de vies humaines en jeu. Si on voulait que des vies puissent être suspendues au bon fonctionnement de l'Internet (une très mauvaise idée), il faudrait en effet changer radicalement son architecture et en faire un réseau bien plus fermé, bien plus lent et bien moins innovant.

Et la deuxième question, celle de la légitimité à faire des tests sur le vrai Internet ? Plusieurs commentateurs (comme Pierre Col <http://www.zdnet.fr/blogs/infra-net/encore-2-incidents-majeurs-su
htm> ou Earl Zmijewski <http://www.renesys.com/blog/2010/08/house-of-cards.shtml>
- dans un article qui était autrement très intéressant) ont mis en cause le RIPE-NCC pour avoir fait des essais avec le vrai Internet. Des noms d'oiseaux ont circulé avec des arguments du genre « on n'est plus à l'époque du réseau universitaire pour jouer, il faut maintenant être très prudent avec le réseau ».

Cet argument semble de bon sens mais il revient en fait à taper sur le messager parce qu'on n'aime pas le message. La panne est gênante et c'est justement pour cela qu'il faut féliciter le RIPE-NCC pour avoir testé et permis qu'on la découvre avant les méchants, qui l'auraient utilisé de manière bien plus agressive ! Pour la sécurité et la stabilité de l'Internet, il **faut** continuer à tester. Si on engueule le messager parce que le message nous déplait, plus personne n'osera faire de tests et on ne découvrira les problèmes que le jour d'une vraie attaque !

Au moins, aurait-il été possible de tester plus prudemment, comme semble le promettre le premier communiqué public du RIPE-NCC <http://www.ripe.net/news/ris-outage.html> ? Par exemple, ne pouvait-on pas tester dans le laboratoire avant d'essayer sur le vrai Internet ? La lecture de l'article du RIPE-NCC <https://labs.ripe.net/Members/erik/ripe-ncc-and-duke-university-bgp-
montre bien que l'expérience avait été soigneusement étudiée (comme le montre le fait d'interposer un routeur BGP standard pour être sûr de n'émettre que du BGP légal, je n'y aurai pas pensé). Aurait-il fallu tester avec divers modèles de routeurs pour s'assurer qu'on ne les cassait pas ? Peut-être mais il n'y a de toute façon aucun chance pour qu'un laboratoire ait à sa disposition toutes les combinaisons

de routeur et de logiciel (d'autant plus que la bogue « attribut 99 » n'affectait que le routeur **suivant** le routeur bogué). Il n'aurait donc pas été possible de tout tester. À un moment ou à l'autre, il faut passer aux essais en vrai grandeur. Comme l'avait noté un participant sur la liste Nanog, « *"I'm planning on announcing x.y.z.0/20 later in the week – x, y and z are all prime and the sum of all 3 is also a prime. There is a non-zero chance that something somewhere will go flooie, shall I send mail now or later ?"* ».

À défaut de pouvoir tout tester à l'avance, le RIPE-NCC aurait-il pu mieux communiquer ? Toute mesure active (ce qui était le cas de l'« expérience attribut 99 ») peut potentiellement pertuber le système qu'elle mesure. Le RIPE-NCC n'a, semble t-il, pas prévenu à l'avance, sa première annonce était envoyée sur une liste fermée (elle a ensuite été relayée sur une liste publique <http://mailman.nanog.org/pipermail/nanog/2010-August/024837.html>). Donc, oui, sur ce point et seulement sur celui-là, le RIPE-NCC aurait pu faire mieux. Ne nous faisons cependant pas trop d'illusions : les messages d'avertissement (« Nous allons annoncer pendant soixante minutes un attribut BGP non enregistré, depuis deux POP et pour le préfixe 203.0.113.0/24 ») sont en général complètement ignorés par les équipes opérationnelles déjà débordées...

Ah au fait, cette expérience, à quoi elle était destinée ? À la sécurité ! Il s'agissait de voir si certaines propositions techniques de sécurisation de BGP (pour empêcher des embrouilles à la Pakistan Telecom <https://www.bortzmeyer.org/pakistan-pirate-youtube.html>) étaient réalistes, car elles reposaient sur des attributs BGP nouveaux. Si on ne teste pas la possibilité de déployer ces extensions à BGP dans le vrai Internet, on peut arrêter tout de suite tout le travail sur le BGP sécurisé.

Un article résumant la panne et ses conséquences sur BGP, telles que vues par Cyclops <http://cyclops.cs.ucla.edu/>, est "*Cisco bug crashed Internet*" <http://cyclops.cs.ucla.edu/blog/?p=96>. Si vous aimez les graphiques, une jolie représentation de la baisse de trafic liées à la panne : <http://inl.info.ucl.ac.be/system/files/16all.png>. Pour une analyse des conséquences de l'« expérience 99 » sur le DNS, voir le communiqué de l'AFNIC <http://operations.afnic.fr/fr/2010/09/02/perturbation-du-ripe.html>.