RFC 8374 : BGPsec Design Choices and Summary of Supporting Discussions

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 1 mai 2018

Date de publication du RFC : Avril 2018

https://www.bortzmeyer.org/8374.html

Ce RFC est un peu spécial : il ne normalise pas un protocole, ni des procédures internes à l'IETF, et il n'est pas non plus la description d'un problème à résoudre, ou un cahier des charges d'une solution à développer. Non, ce RFC est la documentation a posteriori des choix qui ont été effectués lors du développement de BGPsec, une solution de sécurisation du routage Internet. C'est donc un document utile si vous lisez les RFC sur BGPsec, comme le RFC 8205 [1] et que vous vous demandez « mais pourquoi diable ont-ils choisi cette approche et pas cette autre, qui me semble bien meilleure ? »

Un petit rappel du contexte : le protocole de routage BGP fonctionne en échangeant des informations entre pairs, sur les routes que chaque pair sait joindre. Par défaut, un pair peut raconter n'importe quoi, dire qu'il a une route vers `2001:db8::/32` alors qu'il n'est pas le titulaire de ce préfixe et n'a pas non plus appris cette route d'un de ses pairs. Cela rend donc le routage Internet assez vulnérable. Pour le sécuriser, il existe plusieurs mécanismes qui font que, en pratique, ça ne marche pas trop mal. L'IETF a développé une solution technique, qui a deux couches : une infrastructure à clés publiques, la RPKI, normalisée dans les RFC 6480 et RFC 6481, et une seconde couche, les services qui utilisent la RPKI pour authentifier tel ou tel aspect du routage. Deux de ces services sont normalisés, les ROA ("*Route Origin Authorization*") des RFC 6482 et RFC 6811, qui servent à authentifier l'AS d'origine d'un préfixe, et BGPsec (RFC 8205), qui sert à authentifier le **chemin d'AS**, la liste des AS empruntés par une annonce de route (cf. section 2.1.1 de notre RFC). Sans BGPsec, les ROA, utilisés seuls, ne peuvent pas arrêter certaines attaques (cf. RFC 7132, qui explique quelles menaces traite BGPsec, et RFC 7353, cahier des charges de BGPsec). Par exemple, si l'AS 64641, malhonnête, veut tromper son pair, l'AS 64642, à propos du préfixe `2001:db8::/32`, et que ce préfixe a un ROA n'autorisant que 64643 à être à l'origine, le malhonnête peut fabriquer une annonce avec le chemin d'AS `64641 [éventuellement d'autres AS] 64643` (rappelez-vous que les chemins d'AS se lisent de droite à gauche) et l'envoyer à 64641. Si

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc8205.txt`

celui-ci vérifie le ROA, il aura l'impression que tout est normal, l'AS 64643 étant bien indiqué comme l'origine. Et cela marchera même si les annonces de l'AS 64643 ne sont jamais passées par ce chemin ! BGPsec répare ce problème en transportant un chemin d'AS signé, authentifiant toutes les étapes.

Lors du développement de BGPsec, un document avait été rédigé pour documenter tous les choix effectués, mais n'avait pas été publié. C'est désormais chose faite avec ce RFC, qui documente les choix, les justifie, et explique les différences entre ces choix initiaux et le protocole final, modifié après un long développement. Parmi les points importants du cahier des charges (RFC 7353) :

— Valider la totalité du chemin d'AS (et pas seulement détecter certains problèmes, comme le font les ROA),
— Déployable de manière incrémentale puisqu'il est évidemment impossible que tous les opérateurs adoptent BGPsec le même jour,
— Ne pas diffuser davantage d'informations que celles qui sont déjà diffusées. Par exemple, les opérateurs ne souhaitent pas forcément publier la liste de tous leurs accords d'appairage.

Finie, l'introduction, passons tout de suite à certains des choix qui ont été effectués. (Il est évidemment recommandé de lire le RFC 8205 avant, puis de lire notre RFC 8374 après, si on veut tous les choix documentés.) D'abord, à quoi ressemblent les signatures, et les nouveaux messages BGP (section 2 de notre RFC). Les signatures BGPsec signent le préfixe IP, l'AS signataire, et l'AS suivant à qui on va transmettre l'annonce. Le premier point important est qu'on signe l'AS suivant, de manière à créer une chaîne de signatures vérifiables. Autrement, un attaquant pourrait fabriquer un faux chemin à partir de signatures correctes.

Si on utilise le "*prepending*" d'AS, la première version de BGPsec prévoyait une signature à chaque fois, mais on peut finalement mettre une seule signature pour la répétition d'AS, avec une variable `pCount` qui indique leur nombre (RFC 8205, section 3.1), afin de diminuer le nombre de signatures (les signatures peuvent représenter la grande majorité des octets d'une annonce BGP).

Le second point important est que certains attributs ne sont pas **pas** signés, comme par exemple la préférence locale (RFC 4271, section 5.1.5) ou les communautés (cf. RFC 1997). Ces attributs pourront donc être modifiés par un attaquant à sa guise. Ce n'est pas si grave que ça en a l'air car la plupart de ces attributs n'ont de sens qu'entre deux AS (c'est le cas de la communauté `NO_EXPORT`) ou sont internes à un AS (la préférence locale). Sur ces courtes distances, on espère de toute façon que la session BGP sera protégée, par exemple par AO (RFC 5925).

La signature est transportée dans un attribut BGP optionnel et non-transitif (qui n'est pas transmis tel quel aux routeurs suivants). Optionnel car, sinon, il ne serait pas possible de déployer BGPsec progressivement. Et non-transitif car un routeur BGPsec n'envoie les signatures que si le routeur suivant lui a dit qu'il savait gérer BGPsec.

Dans le message BGP signé, le routeur qui signe est identifié par sa clé publique, pas par le certificat entier. Cela veut dire qu'on ne peut valider les signatures que si on a accès à une dépôt de la RPKI, avec les certificats.

La section 3 de notre RFC traite le cas des retraits de route : contrairement aux annonces, ils ne sont pas signés. Un AS est toujours libre de retirer une route qu'il a annoncée (BGP n'accepte pas un retrait venant d'un autre pair). Si on a accepté l'annonce, il est logique d'accepter le retrait (en supposant évidemment que la session entre les deux routeurs soit raisonnablement sécurisée).

La section 4, elle, parle des algorithmes de signature. Le RFC 8208 impose ECDSA avec la courbe P-256 (cf. RFC 6090). RSA avait été envisagé mais ECDSA avait l'avantage de signatures bien plus petites

(cf. l'étude du NIST <http://www.nist.gov/itl/antd/upload/BGPSEC_RIB_Estimation.pdf> sur les conséquences de ce choix).

Une autre décision importante dans cette section est la possibilité d'avoir une clé par routeur et pas par AS (cf. RFC 8207). Cela évite de révoquer un certificat global à l'AS si un seul routeur a été piraté. (Par contre, il me semble que c'est indiscret, permettant de savoir quel routeur de l'AS a relayé l'annonce, une information qu'on n'a pas forcément actuellement.)

Décision plus anecdotique, en revanche, celle comme quoi le nom dans le certificat ("*Subject*") sera la chaîne `router` suivie du numéro d'AS (cf. RFC 5396) puis de l'identité BGP du routeur. Les détails figurent dans le RFC 8209.

Voyons maintenant les problèmes de performance (section 5). Par exemple, BGPsec ne permet pas de regrouper plusieurs préfixes dans une annonce, comme on peut le faire traditionnellement avec BGP. C'est pour simplifier le protocole, dans des cas où un routeur recevrait une annonce avec plusieurs préfixes et n'en transmettrait que certains. Actuellement, il y a en moyenne quatre préfixes par annonce (cf. l'étude faite par l'auteur du RFC <http://www.nist.gov/itl/antd/upload/BGPSEC_RIB_Estimation.pdf>). Si tout le monde adoptait BGPsec, on aurait donc quatre fois plus d'annonces, et il faudra peut-être dans le futur optimiser ce cas.

On l'a vu plus haut, il n'est pas envisageable de déployer BGPsec partout du jour au lendemain. Il faut donc envisager les problèmes de coexistence entre BGPsec et BGP pas sécurisé (section 6 de notre RFC). Dès que deux versions d'un protocole, une sécurisée et une qui ne l'est pas, coexistent, il y a le potentiel d'une attaque par repli, où l'attaquant va essayer de convaindre une des parties de choisir la solution la moins sécurisée. Actuellement, BGPsec ne dispose pas d'une protection contre cette attaque. Ainsi, il n'y a pas de moyen de savoir si un routeur qui envoie des annonces non-signées le fait parce qu'il ne connait pas BGPsec, ou simplement parce qu'un attaquant a modifié le trafic.

La possibilité de faire du BGPsec est négociée à l'établissement de la session BGP. Notez qu'elle est asymétrique. Par exemple, il est raisonnable qu'un routeur de bordure signe ses propres annonces mais accepte tout de la part de ses transitaires, et ne lui demande donc pas d'envoyer les signatures. Pendant un certain temps (probablement plusieurs années), nous aurons donc des ilots BGPsec au milieu d'un océan de routeurs qui font du BGP traditionnel. On peut espérer qu'au fur et à mesure du déploiement, ces ilots se rejoindront et formeront des iles, puis des continents.

La question de permettre des chemins d'AS partiellement signés avait été discutée mais cela avait été rejeté : il faut signer tout le chemin, ou pas du tout. Des signatures partielles auraient aidé au déploiement progressif mais auraient été dangereuses : elle aurait permis aux attaquants de fabriquer des chemins valides en collant des bouts de chemins signés - et donc authentiques - avec des bouts de chemins non-signés et mensongers.

La section 7 de notre RFC est consacrée aux interactions entre BGPsec et les fonctions habituelles de BGP, qui peuvent ne pas bien s'entendre avec la nouvelle sécurité. Par exemple, les très utiles communautés BGP (RFC 1997 et RFC 8092). On a vu plus haut qu'elles n'étaient pas signées du tout et donc pas protégées. La raison est que les auteurs de BGPsec considèrent les communautés comme mal fichues, question sécurité. Certaines sont utilisées pour des décisions effectives par les routeurs, d'autres sont juste pour le débogage, d'autres encore purement pour information. Certaines sont transitives, d'autres utilisées seulement entre pairs se parlant directement. Et elles sont routinement modifiées en route. Le RFC conseille, pour celles qui ne sont utilisées qu'entre pairs directs, de plutôt sécuriser la session BGP.

Pour les communautés qu'on veut voir transmises transitivement, il avait été envisagé d'utiliser un bit libre pour indiquer que la communauté était transitive et donc devait être incluse dans la signature. Mais la solution n'a pas été retenue. Conseil pratique, dans la situation actuelle : attention à ne pas utiliser des communautés transmises transitivement pour des décisions de routage.

Autre cas pratique d'interaction avec un service très utile, les serveurs de route. Un point d'échange peut fonctionner de trois façons :

— Appairages directs entre deux membres et, dans ce cas, BGPsec n'a rien de particulier à faire.
— Utilisation d'un serveur de routes qui ajoute son propre AS dans le chemin (mais, évidemment, ne change pas le `NEXT_HOP`, c'est un serveur de routes, pas un routeur). Cette méthode est de loin la plus rare des trois. Là aussi, BGPsec n'a rien de particulier à faire.
— Utilisation d'un serveur de routes qui n'ajoute pas son propre AS dans le chemin. Sur un gros point d'échange, cette méthode permet d'éviter de gérer des dizaines, voire des centaines d'appairages. Pour BGPsec, c'est le cas le plus délicat. Il faut que le serveur de routes mette son AS dans le chemin, pour qu'il puisse être validé, mais en positionnant `pCount` à 0 (sa valeur normale est 1, ou davantage si on utilise le "*prepending*") pour indiquer qu'il ne faut pas en tenir compte pour les décisions de routage (fondées sur la longueur du chemin), seulement pour la validaton.

Un point de transition intéressant est celui des numéros d'AS de quatre octets, dans le RFC 4893. La technique pour que les AS ayant un tel numéro puisse communiquer avec les vieux routeurs qui ne comprennent pas ces AS est un bricolage utilisant un AS spécial (23456), bricolage incompatible avec BGPsec, qui, d'ailleurs, exige que les AS de quatre octets soient acceptés. En pratique, on peut espérer que les derniers routeurs ne gérant pas les AS de quatre octets auront disparu bien avant que BGPsec soit massivement déployé.

La section 8 du RFC discute de la validation des chemins d'AS signés. Par exemple, le RFC 8205 demande qu'un routeur transmette les annonces ayant des signatures invalides. Pourquoi ? Parce que la RPKI n'est que modérement synchrone : il est parfaitement possible qu'un nouveau certificat ne soit arrivé que sur certains routeurs et que, donc, certains acceptent la signature et d'autres pas. Il ne faut donc pas préjuger de ce que pourront valider les copains.

Une question qui revient souvent avec les techniques de sécurité (pas seulement BGPsec mais aussi des choses comme DNSSEC) est « et si la validation échoue, que doit-on faire des données invalides ? » Vous ne trouverez pas de réponse dans le RFC : c'est une décision locale. Pour BGPsec, chaque routeur, ou plus exactement son administrateur, va décider de ce qu'il faut faire avec les annonces dont le chemin d'AS signé pose un problème. Contrairement à DNSSEC, où la validation peut donner trois résultats (oui, en fait, quatre, mais je simplifie, cf. RFC 4035), « sûr », « non sûr », et « invalide », BGPsec n'a que deux résultats possibles, « valide » et « invalide ». L'état « invalide » regroupe aussi bien les cas où le chemin d'AS n'est pas signé (par exemple parce qu'un routeur non-BGPsec se trouvait sur le trajet) que le cas où une signature ne correspond pas aux données (les deux états « non sûr » et « invalide » de DNSSEC se réduisent donc à un seul ici). Il avait été discuté de faire une division plus fine entre les différents cas d'invalidité mais il avait semblé trop complexe de rendre en compte tous les cas possibles. Notez que « invalide » couvre même le cas où un ROA valide l'origine (un cas particulier des chemins partiellement signés, déjà traités).

Donc, si une annonce est invalide, que doit faire le routeur ? D'abord, la décision d'accepter ou pas une route dépend de plusieurs facteurs, la validation BGPsec n'en étant qu'un seul. Ensuite, il n'est pas évident de traiter tous les cas. D'où la décision de laisser le problème à l'administrateur réseaux.

Ah, et si vous utilisez iBGP, notez que la validation BGPsec ne se fait qu'en bord d'AS. Vous pouvez transporter l'information comme quoi une annonce était valide ou invalide comme vous voulez (une communauté à vous ?), il n'existe pas de mécanisme standard dans iBGP pour cela.

Enfin, la section 9 de notre RFC traite de quelques problèmes d'ordre opérationnel. Mais pensez à lire le RFC 8207 avant. Par exemple, elle estime que BGPsec, contrairement à la validation avec les ROA seuls, nécessitera sans doute du nouveau matériel dans les routeurs, comme un coprocesseur cryptographique, et davantage de RAM. C'est une des raisons pour lesquelles on ne verra certainement pas de déploiement significatif de BGPsec avant des années. Ceci dit, au début, les routeurs BGPsec auront peu de travail supplémentaire, précisément puisqu'il y aura peu d'annonces signées, donc pourront retarder les mises à jour matérielles. D'ici que BGPsec devienne banal, des optimisations comme celles décrites dans cet exposé <http://csrc.nist.gov/groups/ST/ecc-workshop-2015/papers/session6-adalier-Mehmet.pdf> ou celui-ci <https://www.nanog.org/meetings/abstract?id=3043>, ou encore dans l'article « "*Design and analysis of optimization algorithms to minimize cryptographic processing in BGP security protocols*" <https://www.sciencedirect.com/science/article/pii/S0140366417303365> » aideront peut-être.