

RFC 9229 : IPv4 Routes with an IPv6 Next Hop in the Babel Routing Protocol

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 11 mai 2022

Date de publication du RFC : Mai 2022

<https://www.bortzmeyer.org/9229.html>

Voici une extension sympathique qui réjouira tous les amateurs de routage IP. Elle permet à un routeur qui parle le protocole de routage Babel d'annoncer un préfixe IPv4 où le routeur suivant pour ce préfixe n'a pas d'adresse IPv4 sur cette interface.

Un petit rappel de routage : les annonces d'un routeur qui parle un protocole de routage dynamique comme Babel (RFC 8966¹) comprennent un préfixe IP (comme 2001:db8:543::/48) et l'adresse IP du routeur suivant ("*next hop*" dans la langue de Radia Perlman). Traditionnellement, le préfixe et l'adresse du routeur suivant étaient de la même version (famille) : tous les deux en IPv4 ou tous les deux en IPv6. Mais, si on veut router IPv6 et IPv4, cela consomme une adresse IP par interface sur chaque routeur, or les adresses IPv4 sont désormais une denrée très rare. D'où la proposition de ce RFC : permettre des annonces d'un préfixe IPv4 avec une adresse de routeur suivant en IPv6. (Notez que cela ne concerne pas que Babel, cela pourrait être utilisé pour d'autres protocoles de routage dynamique.)

La section 1 du RFC résume ce que fait un protocole de routage. Son but est de construire des tables de routage, où chaque entrée de la table est indexée par un préfixe d'adresses IP et a une valeur, l'adresse du routeur suivant. Par exemple :

destination	next hop
2001:db8:0:1::/64	fe80::1234:5678%eth0
203.0.113.0/24	192.0.2.1

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc8966.txt>

Lorsqu'un routeur doit transmettre un paquet, il regarde dans cette table et trouve l'adresse du routeur suivant. Il va alors utiliser un protocole de résolution d'adresses (ARP - RFC 826 - pour IPv4, ND - RFC 4861 - pour IPv6) pour trouver l'adresse MAC du routeur suivant. Rien dans cette procédure n'impose que le préfixe de destination et l'adresse du routeur suivant soient de la même version d'IP. Un routeur qui veut transmettre un paquet IPv4 vers un routeur dont il ne connaît que l'adresse IPv6 procède à la résolution en adresse MAC, puis met simplement le paquet IPv4 dans une trame portant l'adresse MAC en question (l'adresse IPv6 du routeur suivant n'apparaît pas dans le paquet).

En permettant des annonces de préfixes IPv4 avec un routeur suivant en IPv6, on économise des adresses IPv4. Un réseau peut fournir de la connectivité IPv4 sans avoir d'adresses IPv4, à part au bord. Et comme les adresses IPv6 des routeurs sont des adresses locales au lien allouées automatiquement (cf. RFC 7404), on peut avoir un réseau dont le cœur n'a aucune adresse statique, ce qui peut faciliter sa gestion.

Notre RFC documente donc une extension au protocole Babel (qui est normalisé dans le RFC 8966), nommée `v4-via-v6`. (Comme dit plus haut, le principe n'est pas spécifique à Babel, voir le RFC 5549 pour son équivalent pour BGP.)

Bon, le concret, maintenant. Les préfixes annoncés en Babel sont précédés de l'AE ("*Address Encoding*"), un entier qui indique la version (famille) du préfixe. Ce RFC crée un nouvel AE, portant le numéro 4 <<https://www.iana.org/assignments/babel/babel.xml#ae>>, qui a le même format que l'AE 1 qui servait à IPv4. Une annonce d'un préfixe ayant l'AE 4 devra donc contenir un préfixe IPv4 et un routeur suivant en IPv6. Un routeur qui sait router IPv4 mais n'a pas d'adresse IPv4 sur une de ses interfaces peut donc quand même y annoncer les préfixes connus, en les marquant avec l'AE 4 et en mettant comme adresse l'adresse IPv6 pour cette interface. (Cet AE est uniquement pour les préfixes, pas pour l'indication du routeur suivant.)

Les routeurs qui ne gèrent pas l'extension `v4-via-v6` ignoreront cette annonce, comme ils doivent ignorer toutes les annonces portant un AE inconnu (RFC 8966). Pour cette raison, si un routeur a une adresse IPv4 sur une interface, il faut qu'il utilise des annonces traditionnelles, avec l'AE 1, pour maximiser les chances que son annonce soit acceptée.

Arrivé ici, mes lectrices et mes lecteurs, qui sont très malins, se demandent depuis longtemps « mais un routeur doit parfois émettre des erreurs ICMP (RFC 792), par exemple s'il n'a pas d'entrée dans sa table pour cette destination ». Comment peut-il faire s'il n'a pas d'adresse sur l'interface où il a reçu le paquet coupable? Ne rien émettre n'est pas acceptable, certains protocoles en dépendent. C'est par exemple le cas de la découverte de la MTU du chemin, documentée dans le RFC 1191, qui a besoin des messages ICMP "*Packet too big*". Certes, il existe un algorithme de découverte de cette MTU du chemin qui est entièrement de bout en bout, et ne dépend donc pas des routeurs et de leurs messages (RFC 4821). Mais ICMP sert à d'autres choses, on ne peut pas y renoncer.

Si le routeur Babel a une adresse IPv4 quelque part, sur une de ses interfaces, il peut l'utiliser comme adresse IP source pour ses messages ICMP. S'il n'en a pas, il peut toujours utiliser `192.0.0.8` comme suggéré par la section 4.8 du RFC 7600. Bien sûr, si tout le monde le fait, des outils d'administration du réseau très pratiques comme `traceroute` seront moins utiles.

La nouvelle extension peut soulever des questions de sécurité (section 6 du RFC). Par exemple, si l'administrateur réseaux croyait que, parce que les routeurs n'avaient pas d'adresse IPv4 sur une interface, les paquets IPv4 ne seraient pas traités sur cette interface, cette supposition n'est plus vraie. Ainsi, une île de machines IPv4 qui se croyaient séparées de l'Internet IPv4 par un ensemble de routeurs qui n'avaient pas d'adresse IPv4 de leur côté peut se retrouver subitement connectée. Si ce n'est pas ce

qu'on veut, il faut penser à ne pas se contenter du protocole de routage pour filtrer, mais aussi à filtrer explicitement IPv4.

Question mise en œuvre, cette extension figure dans babeld <<https://www.irif.fr/~jch/software/babel/>> (voir ici un compte-rendu d'expérience <<https://alioth-lists.debian.net/pipermail/babel-users/2022-May/003944.html>>) à partir de la version 1.12, ainsi que dans BIRD.