RFC 8021 : Generation of IPv6 Atomic Fragments Considered Harmful

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 janvier 2017

Date de publication du RFC : Janvier 2017

https://www.bortzmeyer.org/8021.html

C'est quoi, un « fragment atomique » ? Décrits dans le RFC 6946 [1], ces charmants objets sont des datagrammes IPv6 qui sont des fragments... sans en être. Ils ont un en-tête de fragmentation sans être fragmentés du tout. Ce RFC estime qu'ils ne servent à rien, et sont dangereux, et devraient donc ne plus être générés.

Le mécanisme de fragmentation d'IPv6 (assez différent de celui d'IPv4) est décrit dans le RFC 2460, sections 4.5 et 5. Que se passe-t-il si un routeur génère un message ICMP "*Packet Too Big*" (RFC 4443, section 3.2) en indiquant une MTU **inférieure** à 1 280 octets, qui est normalement la MTU minimale d'IPv6 ? (Le plus beau, c'est que ce routeur n'est pas forcément en tort, cf. RFC 6145, qui décrivait leur utilisation pour être sûr d'avoir un identificateur de datagramme.) Eh bien, dans ce cas, l'émetteur du datagramme trop gros doit mettre un en-tête « Fragmentation » dans les datagrammes suivants, même s'il ne réduit pas sa MTU en dessous de 1 280 octets. Ce sont ces datagrammes portant un en-tête « Fragmentation » mais pas réellement fragmentés (leur bit M est à 0), qui sont les **fragments atomiques** du RFC 6946.

Malheureusement, ces fragments atomiques permettent des attaques contre les machines IPv6 (section 2 du RFC). Il existe des attaques liées à la fragmentation (RFC 6274 et RFC 7739). Certaines nécessitent que les datagrammes soient réellement fragmentés mais ce n'est pas le cas de toutes : il y en a qui marchent aussi bien avec des fragments atomiques. Un exemple d'une telle attaque exploite une énorme erreur de certaines "*middleboxes*", jeter les datagrammes IPv6 ayant un en-tête d'extension, quel qu'il soit (y compris, donc, l'en-tête Fragmentation). Ce comportement est stupide mais hélas répandu (cf. RFC 7872). Un attaquant peut exploiter cette violation de la neutralité du réseau pour faire une attaque par déni de service : il émet des faux messages ICMP "*Packet Too Big*" avec une MTU inférieur à 1 280 octets,

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc6946.txt

la source se met à générer des fragments atomiques, et ceux-ci sont jetés par ces imbéciles de "*middleboxes*".

Le RFC décrit aussi une variante de cette attaque, où deux pairs BGP jettent les fragments reçus (méthode qui évite certaines attaques contre le plan de contrôle du routeur) mais reçoivent les ICMP "*Packet Too Big*" et fabriquent alors des fragments atomiques. Il serait alors facile de couper la session entre ces deux pairs. (Personnellement, le cas me parait assez tiré par les cheveux...)

Ces attaques sont plus faciles à faire qu'on ne pourrait le croire car :

— Un paquet ICMP peut être légitimement émis par un routeur intermédiaire et l'attaquant n'a donc pas besoin d'usurper l'adresse IP de la destination (donc, BCP 38 <https://www.bortzmeyer.org/bcp38.html> ne sert à rien).
— Certes, l'attaquant doit usurper les adresses IP contenues dans le message ICMP lui-même mais c'est trivial : même si on peut en théorie envisager des contrôles du style BCP 38 de ce contenu, en pratique, personne ne le fait aujourd'hui.
— De toute façon, pas mal de mises en œuvres d'IP ne font aucune validation du contenu du message ICMP (malgré les recommandations du RFC 5927).
— Un seul message ICMP suffit, y compris pour plusieurs connexions TCP, car la MTU réduite est typiquement mémorisée dans le cache de l'émetteur.
— Comme la seule utilisation légitime connue des fragments atomiques était celle du RFC 6145 (qui a depuis été remplacé par le RFC 7915), on pourrait se dire qu'il suffit de limiter leur génération aux cas où on écrit à un traducteur utilisant le RFC 6145. Mais cela ne marche pas, car il n'y a pas de moyen fiable de détecter ces traducteurs.

Outre ces problèmes de sécurité, le RFC note (section 3) que les fragments atomiques ne sont de toute façon pas quelque chose sur lequel on puisse compter. Il faut que la machine émettrice les génère (elle devrait, mais la section 6 du RFC 6145 note que beaucoup ne le font pas), et, malheureusement, aussi bien les messages ICMP "*Packet Too Big*" que les fragments sont souvent jetés par des machines intermédiaires.

D'ailleurs, il n'est même pas certain que la méthode du RFC 6145 (faire générer des fragments atomiques afin d'obtenir un identificateur par datagramme) marche vraiment, l'API ne donnant pas toujours accès à cet identificateur de fragment. (Au passage, pour avoir une idée de la complexité de la mise en œuvre des fragments IP, voir cet excellent article sur le noyau Linux <https://github.com/NICMx/Jool/wiki/nf_defrag_ipv4-and-nf_defrag_ipv6>.)

En conclusion (section 4), notre RFC demande qu'on abandonne les fragments atomiques :

— Les traducteurs du RFC 7915 (la seule utilisation légitime connue) devraient arrêter d'en faire générer.
— Les machines IPv6 devraient désormais ignorer les messages ICMP "*Packet Too Big*" lorsqu'ils indiquent une MTU inférieure à 1 280 octets.