Dans quels cas les résolveurs DNS auront-ils des problèmes le 5 mai ?

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 avril 2010

https://www.bortzmeyer.org/risques-reels-dns-limite.html

Le 5 mai de cette année <http://www.root-dnssec.org/>, le dernier serveur racine du DNS qui envoyait encore des réponses DNS non-signées va s'aligner sur les autres et enverra, lui aussi, les signatures. Qu'est-ce qui va casser et comment ?

Cet article est assez technique et rentre dans les détails du fonctionnement du DNS. Si vous êtes seulement intéressé par les mesures pratiques à prendre pour tester votre configuration DNS, voyez plutôt mon article d'introduction <https://www.bortzmeyer.org/dns-size.html>. Plusieurs personnes m'ont demandé « mais si mon système rate le test signalé dans votre article, est-ce que je vais **forcément** être victime du 5 mai ? ». La réponse est « ça dépend » et cet article explique de quoi ça dépend.

Donc, quelles sont les conditions exactes qu'il faudra pour qu'un résolveur, à partir du 5 mai, ne puisse plus joindre les serveurs de la racine et n'aie donc plus accès au DNS ? D'abord, un point de vocabulaire, lorsque je dis « résolveur », je parle du logiciel qui envoie les requêtes DNS aux serveurs de la racine et lorsque je dis « client DNS », je parle de l'entité vue par le serveur racine. La différence entre les deux ? Aucune dans le cas d'un chemin normal entre le résolveur et le serveur racine. Mais, si un intermédiaire (pare-feu, par exemple) modifie la requête, le serveur racine ne verra pas la même chose que ce qu'a envoyé le résolveur.

Donc, pour qu'il y aie un problème avec la signature de la racine, il faut que le client DNS envoie une demande de signatures mais que la réponse à celle-ci ne puisse pas être reçue en raison de sa trop grande taille. Commençons par le premier point, la demande de signatures.

Pour des raisons de compatibilité avec la base de logiciels existants, le RFC sur DNSSEC, le RFC 4035 [1] (section 3.1), dit que le serveur faisant autorité doit envoyer les informations DNSSEC (notamment

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc4035.txt

les signatures) lorsque le client DNS a utilisé EDNS0 (RFC 2671) **et** a mis le bit DO (RFC 3225) à 1. Sinon, le serveur ne doit pas noyer le client sous ces informations (RFC 3225, section 3).

Un très vieux résolveur, qui ne connait pas DNSSEC (a fortiori s'il en connait pas EDNS0) n'aura donc pas de problèmes. Il ne pourra pas faire de DNSSEC mais n'aura de la racine que les informations non signées, dont la taille reste bien inférieure aux 512 octets. (Avec dig, l'option `+dnssec` active ce bit DO.)

Mais le résolveur le plus utilisé, BIND, met le bit DO à 1 par défaut, **même s'il n'est pas configuré pour valider** (par exemple parce qu'il a l'option `dnssec-validation no`). Donc, tout résolveur BIND va envoyer le bit DO, la racine enverra les signatures (aujourd'hui, de 700 à 800 octets) et il faut que le chemin de retour soit propre, que la réponse de plus de 512 octets ne soit pas bloquée.

En fait, c'est encore un peu plus compliqué que cela, car cela dépendra d'un autre paramètre envoyé grâce à EDNS0, la taille maximale des réponses. Elle est de 4096 octets par défaut avec BIND mais peut être réduite par configuration. Si on la met à 512 octets, même avec le bit DO, les serveurs racine réagissent alors en diminuant le nombre de colles (adresses IP des serveurs) envoyées, repassant ainsi sous la limite tragique. Si le serveur a plus d'informations à faire passer et qu'elles ne tiennent pas dans 512 octets, même en laissant tomber ce qui n'est pas indispensable, alors il doit mettre le bit TC ("*TrunCation*") à 1 dans la réponse et le résolveur doit normalement réessayer avec TCP, qui n'a jamais eu la limite de 512. Si le résolveur ne peut pas faire de TCP (par exemple parce qu'un pare-feu a été configuré par un ignorant qui croit que le DNS utilise uniquement UDP), il sera alors fichu.

Bien, mettons-nous maintenant dans le cas où le client DNS demande les signatures, la taille qu'il indique est de 4096 octets, la réponse fait plus de 512 octets. Que va t-il se passer ? Normalement, rien, le DNS n'a plus de limite à 512 octets depuis plus de dix ans, et la réponse va arriver intacte.

Dans la réalité, il existe beaucoup de réseaux mal configurés avec les pieds. Il y a deux cas à considérer : la réponse DNS fait plus de 512 octets mais moins de 1500 (ou plutôt 1492, pour avoir une marge, par exemple pour PPPoE). Dans ce cas, il n'y aura pas de problème de MTU <`https://www.bortzmeyer.org/mtu-et-mss-sont-dans-un-reseau.html`> et le seul risque est celui d'un pare-feu ou autre "*middlebox*" mal programmé et/ou mal configuré qui bloque ce paquet (à noter que ladite "*middlebox*" peut être du côté du résolveur ou bien du côté du serveur faisant autorité), croyant bêtement que le DNS, c'est moins de 512 octets. Dans ce cas, le résolveur ne verra jamais venir la réponse. (Notons que ce problème ne date pas de DNSSEC, car, aujourd'hui, les réponses à des requêtes comme `NS .` ou `NS com.` font déjà plus de 512 octets.)

À noter que certains résolveurs, comme BIND, s'ajustent automatiquement dans ce cas, et que, ne voyant pas venir la réponse, ils réessaient sans EDNS. Cela prend du temps (il faut attendre l'expiration d'un délai), cela reviendra à couper DNSSEC mais, au moins, la résolution DNS fonctionnera toujours.

Deuxième cas, la réponse fait plus de 512 octets et même plus que la MTU du lien (typiquement 1500 octets, la MTU d'Ethernet mais elle peut être plus petite dans certains cas). Ce n'est pas le cas avec la racine aujourd'hui, mais cela pourrait changer. Normalement, la solution est la fragmentation des paquets. De nombreuses études ont montré que ce n'était pas un processus fiable (voir par exemple le RFC 2923), notamment en raison des équipements intermédiaires bogués qui refusent les fragments (surtout en UDP). La réponse n'arrivera alors pas et le comportement du résolveur semblera aléatoire car les réponses plus petites que la MTU, elles, passeront toujours. BIND, là encore, finira sans doute par essayer sans EDNS, ce qui empêchera d'utiliser DNSSEC. Certaines réponse seront tronquées, forçant le passage en TCP et, là encore, il faudra donc que TCP soit possible, sur le résolveur, sur le serveur faisant

autorité (voilà pourquoi Zonecheck <http://www.zonecheck.fr/> teste cela) et sur le chemin entre eux (pare-feux...)

Cela semble compliqué, expliqué en langage naturel ? Alors, essayez le même texte mais en pseudo-code, en syntaxe vaguement Ada : (en ligne sur https://www.bortzmeyer.org/files/dns-size-pseudocode.txt).

Alors, quelle conclusion en tirer ? On l'a vu, pour un blocage complet de la résolution DNS, il faudrait beaucoup de conditions simultanées, et qu'elles sont rarement réalisées pour un site qui abrite un résolveur (le client ordinaire d'un FAI avec lien ADSL n'a en général pas de résolveur chez lui). Il est donc probable que, le 5 mai, très peu de sites seront coupés complètement. Et ceux-ci cumulent vraiment les erreurs de configuration... Le plus probable est que beaucoup de sites verront une dégradation des performances mais que peu seront complètement hors-service.

Merci à Bert Hubert, Nicholas Weaver, Michael Sinatra et alnitak <http://serverfault.com/questions/106207/what-are-the-effects-of-the-l-root-server-now-publishing-durz/106240#106240> pour leurs premières explorations du problème.

Quelques références utiles :
— RFC 5625,
— Le rapport SSAC 35 <https://www.icann.org/committees/security/sac035.pdf>.