Le DNS va t-il utiliser de plus en plus souvent TCP ?

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 20 mai 2013

https://www.bortzmeyer.org/dns-over-tcp.html

Le DNS utilise traditionnellement surtout UDP comme protocole de transport. TCP est parfaitement légal mais, en pratique, il a été cantonné aux transferts de zone et à quelques requêtes où la réponse était trop grosse pour passer en UDP. La montée des attaques utilisant le DNS avec **réflexion** et **amplification** a changé les choses et de plus en plus de gens se demandent si le DNS ne va pas utiliser TCP plus fréquemment.

Écartons d'abord un mythe encore propagé par certains ignorants : non, le DNS n'utilise pas que UDP. Outre les transferts de zone (cf. RFC 5936 [1]), le DNS utilise TCP dès que la réponse est de taille trop importante pour être transmise en UDP. C'est combien d'octets, « trop importante » ? Cela dépend. Autrefois, il y avait une limite en dur à 512 octets. Elle a été remplacée depuis longtemps par l'extension EDNS (aujourd'hui décrite par le RFC 6891) qui permet d'indiquer la taille des réponses qu'on peut recevoir. Le serveur répondeur ayant également sa propre limite, la taille maximale pratique est le minimum de la taille annoncée que le demandeur et de la taille configurée dans le serveur répondeur. Pour la plupart des logiciels DNS, ces deux tailles valent par défaut 4 096 octets, mais peuvent être modifiées. Vous verrez ainsi que les serveurs de noms de `.com` ont une limite configurée à 1 460 octets. Même si le demandeur propose d'avantage (8 192 octets dans l'exemple suivant), le serveur enverra une réponse tronquée (bit TC mis à un) et le demandeur réessaiera alors en TCP. Voyons avec dig :

```
% dig +bufsize=8192 @a.gtld-servers.net ANY com.
;; Truncated, retrying in TCP mode.
...
;; flags: qr aa rd; QUERY: 1, ANSWER: 21, AUTHORITY: 0, ADDITIONAL: 16
...
;; SERVER: 192.5.6.30#53(192.5.6.30)
;; WHEN: Mon May 20 17:21:15 2013
;; MSG SIZE  rcvd: 1792
```

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc5936.txt`

dig a automatiquement réessayé en TCP. Si on lui dit de ne pas le faire :

```
% dig +bufsize=8192 +noignore @a.gtld-servers.net ANY com.
;; flags: qr aa tc rd; QUERY: 1, ANSWER: 20, AUTHORITY: 0, ADDITIONAL: 1
...
;; SERVER: 192.5.6.30#53(192.5.6.30)
;; WHEN: Mon May 20 17:23:24 2013
;; MSG SIZE  rcvd: 1365
```

Le `tc` dans la réponse indique qu'elle a été tronquée (regardez le compteur `ANSWER` et la taille de la réponse).

Cette possibilité de se rabattre en TCP est cruciale si la réponse est de trop grande taille (ce qui est plus fréquent aujourd'hui, avec IPv6, les IDN et surtout DNSSEC). C'est pour cela qu'il est essentiel de s'assurer que la configuration du réseau permette les requêtes et les réponses TCP, comme exigé par le RFC 7766. C'est aussi pour cela que l'outil Zonecheck <http://www.zonecheck.fr/> a, par défaut, une politique de tests qui impose que le serveur réponde en TCP. (Un point qui a toujours fait l'objet d'un consensus chez les experts <http://www.circleid.com/posts/afnic_dns_server_redelegation/> à chaque discussion.)

Au fait, pourquoi le serveur a t-il une limite de taille en UDP et pas en TCP ? Car, en UDP, on n'a aucun moyen de garantir la véracité de l'adresse IP source utilisée. Cela permet des attaques par réflexion + amplification, qui ne sont pas possibles en TCP. Limiter la taille des réponses, comme le fait `.com`, limite donc les dégâts.

OK, bon, on a le droit d'utiliser TCP si on veut. Si un employé ou un consultant en sécurité dit qu'il faut débrayer / bloquer TCP, on sait qu'on peut muter l'employé et virer le consultant, ils ne connaissent pas leur métier. Mais le fait qu'on **puisse** utiliser TCP veut-il dire qu'il le **faut** ? Notons qu'un client DNS peut toujours utiliser TCP dès le début, sans attendre une réponse tronquée :

```
% dig +tcp @a.gtld-servers.net ANY com.
...
;; flags: qr aa rd; QUERY: 1, ANSWER: 21, AUTHORITY: 0, ADDITIONAL: 16
...
;; SERVER: 192.5.6.30#53(192.5.6.30)
;; WHEN: Mon May 20 17:51:26 2013
;; MSG SIZE  rcvd: 1792
```

Mais quels sont les problèmes si tous les clients DNS faisaient ainsi ? Il y en a deux : bien des réseaux bloquent stupidement les requêtes DNS TCP (cf. le cas du consultant en sécurité incompétent, ainsi que les excellents tests TCP de Zonecheck, cités plus haut). Ensuite, UDP est bien plus léger pour le serveur. Il peut être mis en œuvre sans état (notez que ce n'est pas toujours fait ainsi sur les systèmes modernes) et la même machine peut donc servir bien plus de requêtes par seconde en UDP.

Au dernier atelier OARC <https://www.dns-oarc.net> à Dublin, le 12 mai <https://indico.dns-oarc.net/indico/conferenceDisplay.py?ovw=True&confId=0>, deux excellents exposés étaient revenus sur cette question. Celui de Francis Dupont <https://indico.dns-oarc.net/indico/materialDisplay.py?contribId=21&materialId=slides&confId=0> présentait le problème

des performances TCP, les réglages souhaitables sur les serveurs **et** les clients (ceux par défaut conviennent rarement) et les résultats de ses mesures (après réglages, 30 kr/s en TCP sur une machine qui en fait 130 kr/s en UDP). Cela indique que TCP reste plus lent (ce qui est logique) mais que l'écart est peut-être supportable (le DNS va de toute façon nécessiter des investissements, entre autre en raison des attaques par réflexion). Et des optimisations sont possibles, comme de laisser les connexions TCP ouvertes (c'est l'établissement de la connexion TCP qui est coûteux, et on peut faire passer plusieurs requêtes DNS sur une seule connexion).

Autre exposé très intéressant, celui d'Ed Lewis <https://indico.dns-oarc.net/indico/materialDisplay.py?contribId=18&materialId=slides&confId=0>, exposé très provoquant sur la sécurité du DNS, le genre d'exposé qui remet tout à plat et où personne n'est d'accord à 100 % mais qui disait plein de choses justes, comme le fait que nous serons probablement amenés à dépendre de plus en plus de TCP dans le futur.

Et mon opinion ? Je pense qu'en effet, les défauts d'UDP (notamment en cas d'attaques par réflexion) deviendront de plus en plus insupportables avec le temps et que les problèmes de performance de TCP doivent être relativisés : aujourd'hui, avec l'expérience des serveurs HTTP, faire des serveurs qui encaissent des dizaines de milliers de connexion par seconde n'est plus de la magie noire... Experts TCP, il y aura donc peut-être bientôt du travail pour vous dans le monde DNS.

Vous pouvez aussi lire une bonne étude sur TCP pour le DNS <http://www.circleid.com/posts/20130820_a_question_of_dns_protocols/>, un exemple de ce qui se passe si on bloque TCP <https://x-cli.eu/free/free-notcp.txt> et l'exposé de Huston <https://indico.dns-oarc.net/indico/materialDisplay.py?contribId=13&materialId=slides&confId=1> sur le pourcentage de résolveurs DNS qui savent faire du TCP.
