QUIC et le suivi des utilisateurs par le serveur

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 12 juin 2021

https://www.bortzmeyer.org/quic-tracking.html

Suite à mes articles sur le protocole QUIC <https://www.bortzmeyer.org/quic.html>, on a parfois attiré mon attention sur un problème potentiel de vie privée : certes, QUIC protège bien contre la surveillance exercée par un acteur extérieur à la communication, grâce à son chiffrement systématique. Mais qu'en est-il de la surveillance exercée par le serveur sur lequel on se connecte ? Penchons-nous sur la complexité de la conception des protocoles Internet, quand on est soucieux de vie privée.

Au contraire du cas de la surveillance par un tiers, très longuement traité dans les RFC sur QUIC (voir par exemple le RFC 9000 [1], section 9.5), le cas du suivi d'un utilisateur par le serveur auquel il se connecte est absent. Et ce alors que les grosses entreprises capitalistes qui forment un partie très visible du Web d'aujourd'hui sont connues pour pratiquer la surveillance de masse. Mais qu'en est-il exactement ?

Voyons d'abord le principe de l'éventuel problème de suivi d'un utilisateur par le serveur. Les connexions QUIC peuvent être longues, peut-être plusieurs heures, voire jours, et elles survivent même aux changements d'adresses IP, QUIC permettant la migration d'une adresse à une autre. Le serveur peut donc facilement déterminer que la demande de `/truc.html` à 9 h est faite par le même utilisateur que la demande de `/machin.html` à 16 h, puisque c'est la même connexion QUIC. Indiscutablement, ce problème de suivi de l'utilisateur existe. Mais est-ce spécifique à QUIC et est-ce un vrai problème en pratique ?

D'abord, le problème est ancien. Si le vieil HTTP original n'envoyait qu'une requête par connexion, cette limitation a disparu il y a longtemps. Ainsi, HTTP/2 (RFC 7540) privilégiait déjà les connexions de longue durée, posant les mêmes problèmes. Toutefois, QUIC, avec sa capacité de survivre aux changements d'adresse IP, étend encore la durée de ces connexions, ce qui peut être vu comme aggravant le problème. (Des techniques assez rares, comme "*multipath TCP*", RFC 8684, fonctionnaient également à travers les changements d'adresses IP.)

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc9000.txt`

Mais surtout, dans l'utilisation typique du Web aujourd'hui, il existe bien d'autres méthodes de suivi de l'utilisateur par le serveur. Il y a évidemment les "*cookies*" du RFC 6265. Même si on n'est pas connecté à un service comme YouTube, des "*cookies*" sont placés. Et ces cookies, contrairement à la connexion de longue durée de QUIC, permettent un suivi inter-serveurs, via Google Analytics et les boutons de partage des GAFA que tant de webmestres mettent sur leurs pages sans réfléchir. Et il n'y a pas que les "*cookies*", le "*fingerprinting*" du navigateur peut également permettre d'identifier un visiteur unique, par toutes les informations que le très bavard HTTP transmet, comme le montre bien le test de l'EFF <https://coveryourtracks.eff.org/>. Bref, à l'heure actuelle, le serveur indiscret qui veut pister ses utilisateurs a bien des moyens plus puissants à sa disposition.

En revanche, si on utilise un système tout orienté vie privée, tel le "*Tor Browser*", qui débraye beaucoup de services du Web trop indiscrets, et fait tout passer par Tor, alors, la durée des connexions QUIC pourrait devenir le maillon faible de la vie privée.

Pour Tor, le problème est à l'heure actuelle purement théorique puisque Tor ne transmet que TCP et que QUIC utilise UDP. Mais il pourrait se poser dans le futur, le projet Tor a d'ailleurs déjà réfléchi à cela dans le contexte de HTTP/2 <https://gitlab.torproject.org/legacy/trac/-/issues/4100> (qui s'appelait SPDY à ses débuts).

Un client QUIC soucieux de ne pas être suivi à la trace peut donc, une fois qu'il a géré les problèmes bien plus énormes que posent les "*cookies*" et le "*fingerprinting*", envisager des solutions comme de ne pas laisser les connexions QUIC durer trop longtemps et surtout ne pas utiliser la migration (qui permet de maintenir la connexion lorsque l'adresse IP change). Cela peut se faire en raccrochant délibérement la connexion, ou simplement en ne prévoyant pas de réserve de "*connection IDs*" (RFC 9000, section 5.1.1). Ceci dit, c'est plus facile à dire qu'à faire car une application n'est pas forcément informée rapidement d'un changement d'adresse IP de la machine. Et, évidemment, cela aura un impact négatif sur les performances totales.

La longue durée des connexions QUIC n'est pas le seul mécanisme par lequel un serveur pourrait suivre à la trace un client. QUIC permet à un client de mémoriser des informations qui lui permettront de se reconnecter au serveur plus vite (ce qu'on nomme le « 0-RTT »). Ces informations (qui fonctionnent exactement comme un "*cookie*" HTTP) permettent évidemment également au serveur de reconnaitre un client passé. Cette possibilité et ses conséquences parfois néfastes sont détaillées dans le RFC 9001, sections 4.5 et 9.1. Notez que cela existe également avec juste TLS (ce qu'on nomme le "*session resumption*", RFC 8446, section 2.2) et avec le TCP Fast Open (RFC 7413), avec les mêmes conséquences sur la possibilité de suivi d'un client par le serveur. Le client QUIC qui voudrait protéger sa vie privée doit donc faire attention, quand il démarre une nouvelle connexion, à ne pas utiliser ces possibilités, qui le trahiraient (mais qui diminuent la latence <https://www.bortzmeyer.org/latence.html> ; toujours le compromis).

Comme souvent en sécurité, on est donc face à un compromis. Si on ne pensait qu'à la vie privée, on utiliserait Tor tout le temps…Les navigateurs Web, par exemple, optimisent clairement pour la vitesse, pas pour la vie privée.