RFC 7844 : Anonymity profile for DHCP clients

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 29 mai 2016

Date de publication du RFC : Mai 2016

https://www.bortzmeyer.org/7844.html

Vous vous connectez à un réseau IP, vous obtenez la configuration via un serveur DHCP, mais vous voulez rester relativement anonyme, et vous n'avez pas envie que le serveur, ou que des indiscrets qui écoutent, sachent qui vous êtes ? C'est là que ce RFC est utile en proposant un profil de requête DHCP qui envoie le moins d'informations possibles.

C'est que DHCP, par défaut, est trop bavard. Comme l'expliquent bien les RFC 7819[1] et RFC 7824, une requête DHCP comporte pas mal de champs et d'options qui peuvent contenir des identificateurs stables, qui permettront au serveur DHCP, ou à des espions qui écoutent le trafic, de reconnaitre une machine. Un exemple d'une telle surveillance est celle effectuée dans les aéroports canadiens <http://www.cbc.ca/news/politics/csec-used-airport-wi-fi-to-track-canadian-travellers-edward-snowden-2517881> mais ce n'est certainement pas la seule. Un des buts de cette surveillance peut être, pour quelqu'un qui peut écouter en plusieurs endroits, de suivre à la trace les utilisateurs en déplacement, en reconnaissant les requêtes spécifiques de leur ordinateur portable ou "*smartphone*". Bien sûr, DHCP n'est pas le seul outil pour faire cela. L'identificateur stable le plus évident est l'adresse MAC, mais c'est aussi le plus connu : les utilisateurs soucieux de leur vie privée prennent souvent des mesures pour la changer de temps en temps. Mais peu de gens pensent aux risques pour la vie privée associés à DHCP. Ceux-ci sont désormais bien établis, et documentés dans les RFC 7819 et RFC 7824.

Mais pourquoi en faire un RFC, une norme ? Après tout, on pourrait se contenter d'attendre que les différents fournisseurs de clients DHCP minimisent les informations envoyées. Le problème est qu'ils risquent de le faire légèrement différemment (certains supprimant une option de la requête que d'autres garderont, par exemple). Cette différence permettra une différenciation des utilisateurs, justement ce qu'on veut éviter (cf. la section 2.4). Au contraire, si tout le monde met en œuvre ce profil standard, tous les clients seront pareils, ce qui protège nettement mieux leur anonymat.

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc7819.txt

La section 2 de notre RFC détaille certains problèmes concrets. J'ai dit plus haut que l'identificateur le plus évident pour suivre une machine à la trace était l'adresse MAC. En Wi-Fi, n'importe qui peut la voir, même si le réseau utilise le chiffrement, puisqu'elle est dans la partie en clair du paquet. En la corrélant avec d'autres informations qu'on retrouve ailleurs, on peut remonter à l'individu qui utilise la machine en question. Le danger est donc réel.

La contre-mesure évidente est de changer l'adresse MAC, en la remplaçant par une valeur aléatoire, par exemple avec macchanger <http://www.gnu.org/software/macchanger>. Cela a fait d'ailleurs l'objets de tests en grand à la réunion IETF de Honolulu <http://www.ietf.org/blog/2014/11/mac-privacy/>, ainsi que dans d'autres réunions <http://www.it.uc3m.es/cjbc/papers/pdf/2015_bernardos_cscn_privacy.pdf>. Il n'y a pas de norme technique décrivant cette « aléatoirisation ». Parmi les choix possibles :

— Un tirage au hasard à chaque connexion à un réseau,
— Un tirage au hasard pour chaque **nouveau** réseau mais on garde toujours la même adresse MAC pour un réseau donné,
— Des changements de temps en temps, si on reste connecté sur des longues périodes. (Cela ne marchera pas en Wi-Fi où le changement d'adresse MAC nécessitera déconnexion/reconnexion donc peut-être changement d'adresse IP donc coupure des connexions TCP en cours.) Cette astuce n'est même pas forcément efficace : un observateur peut facilement noter le départ d'une machine et l'arrivée immédiate d'une autre, corrélant ces deux événements.

En tout cas, un changement de l'adresse MAC est nécessaire, sinon, le profil DHCP « anonyme » de ce RFC aura une utilité limitée. Mais, inversement, attention à ne pas se contenter de changer un seul identificateur. Si on change d'adresse MAC en gardant tout le reste, la tâche d'un surveillant est trop facile, et il déduira sans problème la nouvelle adresse MAC. Il faut changer l'adresse MAC, l'adresse IP et les identificateurs DHCP en même temps.

Notre RFC mentionne qu'il existe des techniques de plus bas niveau (couche 1) qui peuvent défaire tout le travail de protection de la vie privée. Par exemple, certains récepteurs radio permettent d'identifier des particularités d'un émetteur physique et, en observant le trafic Wi-Fi, on peut ainsi identifier une machine donnée quels que soient ses efforts pour ne pas laisser de traces (cf. « "*Wireless Device Identification with Radiometric Signatures*" <http://www.winlab.rutgers.edu/~gruteser/papers/brik_paradis.pdf> »). Heureusement, tous les surveillants n'ont pas les moyens de placer ce genre de récepteurs. S'ils doivent se contenter de capter le trafic, sans pouvoir analyser le signal physique, les précautions recommandées plus loin dans ce RFC vont suffire.

Le profil d'« anonymat » décrit dans ce RFC n'est pas explicite. On ne peut pas distinguer les clients DHCP qui appliquent ce profil de ceux qui n'ont simplement pas mis en œuvre les options trop révélatrices. Il y avait eu une discussion à l'IETF sur une mention explicite, un bit dans la requête disant « je veux être anonyme et ne vous étonnez donc pas si je ne publie pas grand'chose ». Cette idée avait été rejetée avec l'idée que proclamer publiquement qu'on veut être anonyme peut être au contraire un excellent moyen de se distinguer, comme de sortir dans la rue avec un masque de Guy Fawkes. Au contraire, le vrai anonymat s'obtient en ne faisant rien de trop extraordinaire.

Enfin, comme toute mesure de sécurité, celles qui préservent la vie privée ont des inconvénients. La plus évidente est que l'« anonymat » va rendre plus difficile la gestion des réseaux. Si l'administrateur réseaux suit les machines sur le réseau local en les identifiant par leurs adresses MAC, il devra s'adapter. D'autre part, il y a un conflit entre la protection de la vie privée, qui requiert des changements fréquents d'adresse MAC, et le bon fonctionnement du réseau, qui demande qu'on limite le trafic DHCP (mais aussi ARP et NDP). Attention donc à ne pas changer d'adresse MAC « trop souvent ».

Autre inconvénient, le changement de l'identificateur DHCP par le client, s'il permet davantage de vie privée, empêchera d'obtenir des paramètres de connexion stables (comme l'adresse IP, cf. RFC 7618) dans le temps. On n'a rien sans rien !

Et la vie privée du serveur DHCP, on en parle ? On ne s'est préoccupé jusqu'à présent que du client. C'est parce que le RFC considère que c'est le client qui est aujourd'hui suivi à la trace, et c'est son cas qui nécessite des solutions urgentes. Pour le serveur, cela peut attendre (par exemple parce qu'un serveur qui veut être discret peut ne répondre qu'aux clients authentifiés au niveau 2).

Bien, assez de préliminaires, voyons maintenant le profil proposé. D'abord, pour IPv4 (section 3). Donc, le client qui veut minimiser la surveillance doit donc :

— Dans le message `DHCPDISCOVER`, mettre les options « identificateur de client » et « paramètres demandés » mais rien d'autre. (Voir comme contre-exemple cette bogue `<https://labs.riseup.net/code/issues/7688>` de Tails.)
— Dans un `DHCPREQUEST`, ajouter éventuellement la demande d'une adresse IP précise.

L'option « identificateur de client » ("*Client Identifier*", RFC 2132, code 61) est, comme son nom l'indique, dangereuse pour la vie privée. Elle permet un suivi très efficace d'une machine, précisement ce qu'on voudrait empêcher. Ne pas la mettre n'est pas une solution car tous les clients DHCP l'incluent. Ne pas le faire serait se singulariser (voir la remarque sur Guy Fawkes plus haut). Notre RFC demande donc qu'on envoie cette option mais qu'on lui donne comme valeur l'adresse MAC (qui est déjà dans l'entête couche 2 du paquet). Cela contredit le RFC 4361, section 6.1, mais ce changement de politique reflète l'accent désormais mis sur la préservation de la vie privée. Si on n'a pas d'adresse MAC (cas de certaines liaisons point à point), on doit tirer au hasard (après avoir lu le RFC 4086 et la section 5 du RFC 7217) un identificateur.

Et l'autre option acceptable, la liste des paramètres demandés ("*Parameter Request List*", RFC 2132, code 55) ? La liste doit être réduite au minimum, et ordonnée au hasard, pour éviter qu'on reconnaisse le type de client DHCP utilisé.

L'adresse MAC est indispensable, on ne peut pas ne pas la mettre dans le paquet, et, donc, il ne sert à rien de l'omettre du champ correspondant dans la requête DHCP, mais il est recommandé de la changer (et que le client DHCP utilise cette adresse dans son paquet, qu'il ne garde surtout pas l'ancienne).

Le champ « adresse IP du client » ne pose guère de problème de vie privée (puisque la même adresse IP figure dans l'en-tête IP du paquet) **sauf** si le client est nouveau sur ce réseau, l'adresse IP indiquée étant alors celle du précédent réseau. Donc : pensez à effacer l'adresse IP lorsque le matériel signale qu'on vient de changer de réseau. Pour la même raison, le profil anonyme ne comprend pas l'option « adresse IP souhaitée » (RFC 2132, code 50). Elle est certes bien pratique (permettant de conserver son adresse IP) mais très révélatrice.

Parmi les autres options, le nom de la machine (code 12) est très révélateur. Même si ce n'est pas un FQDN, il peut être très spécifique (« `pc-de-jean-dupont` »). Il ne doit pas être utilisé dans ce profil. Si on doit quand même le faire, le nom doit être choisi aléatoirement. Une méthode possible est de condenser un secret et l'adresse MAC, puis prendre la représentation en hexadécimal de six premiers octets (quelque chose comme `echo -n "$SECRET" "$MACADDRESS" | sha256sum | cut -c 1-12` en shell Unix).

L'option "*Client FQDN*" (RFC 4702, code 81) est évidemment également à proscrire. Ou alors, il faut l'« anonymiser » avec la même méthode qu'au paragraphe précédent.

Le RFC 4578 décrit une option UUID, prévue pour PXE, et identifiant de manière unique un client DHCP. Normalement, on n'utilise PXE que dans des environnements contrôlés et sécurisés. Dans un cyber-café inconnu, télécharger son système d'exploitation serait une très mauvaise idée. Il ne faut donc utiliser cette option que dans un réseau que l'on sait sûr.

Enfin, les options indiquant le type de logiciel utilisé (comme "*Vendor Specific Information*", code 43, ou "*Vendor Class, code 60*" sont évidemment à rejeter.

Pour éviter le "*fingerprinting*", le client ne devrait pas ajouter d'autres options, mêmes si celles-ci sont inoffensives du point de vue de la vie privée. En effet, si certains clients mettent telle option, quel que soit son contenu, ils diffusent une information. Et, si on envoie des options DHCP, leur ordre doit être tiré au hasard avant l'envoi, là encore, pour éviter de révéler le type de client utilisé (DHCP permet pas mal de choix, et le choix facilite le "*fingerprinting*"). Si on n'a pas de bon générateur de nombres aléatoires, un repli possible est d'envoyer toujours les options dans l'ordre de leur code : ainsi, tous les logiciels (qui sont dans ce cas) feront pareil.

La section 4 du RFC fait la même chose, mais pour IPv6. DHCP v6 a d'autres possibilités (comme l'existence de deux modes, avec ou sans état). Le profil est donc un peu plus compliqué mais bien des conseils sont identiques (par exemple ceux sur les options "*Client FQDN*", sections 3.8 et 4.9), et je ne les répéterai donc pas ici. Dans le mode sans état, les clients configurent l'adresse avec d'autres mécanismes, comme le SLAAC et ne se servent de DHCP que pour récupérer des informations générales. Le choix entre les deux modes dépend d'options publiées dans les "*Router Advertisement*" (RFC 4861, section 4.2, les bits M et O).

Parmi les messages spécifiques à DHCP v6, `Confirm` (RFC 8415, section 16.5) est dangereux, car il contient de l'information sur le précédent réseau où on s'est connecté. En mode « anonyme », il ne faut pas l'envoyer, sauf si on est certains qu'on est toujours sur le même réseau (ce qui est une vérification pas toujours triviale, le RFC recommande d'utiliser 802.1X).

L'option "*Client Identifier*" (code 1) en DHCP v6 contient un identificateur unique, le DUID ("*DHCP Unique Identifier*", RFC 8415, section 11). Elle rend donc le suivi d'un client très facile. Pour empêcher ce suivi, le profil demande d'utiliser le format 3, « adresse MAC », du DUID (RFC 8415, section 11.4), avec une adresses MAC changée comme indiquée plus haut. Dans les cas où cette option n'est pas exigée par le protocole (comme les "*Information-Request*", RFC 8415, section 18.2.6, qui sont sans état), elle ne doit pas être envoyée.

DHCP v6 permet évidemment d'obtenir des adresses IPv6 (mais ce n'est pas le seul moyen, contrairement à ce qui se passe en IPv4). Une option semble intéressante pour la vie privée, la demande d'adresses temporaires (RFC 8415, section 13.2). Mais elle n'est pas si utile que ça : car peu de serveurs la mettent en œuvre (le client ne peut donc pas se reposer dessus), le renouvellement de ces adresses n'est pas spécifié, et le fait que le client doive la demander est déjà une fuite d'informations (« je veux être anonyme »). Elle est donc déconseillée.

Notre RFC met également en garde contre des options qui peuvent être amusantes mais qui sont rares : les utiliser revient, pour le client DHCP, à « sortir du lot ». C'est le cas par exemple de la demande de délégation de préfixe (RFC 8415), que le PC typique n'utilise pas.

Le RFC se termine par quelques considérations opérationnelles (section 5). La plus évidente, et la plus « gênante » est que l'application de ce profil va cache l'identité du client au serveur DHCP. C'est le but. Mais cela peut être ennuyeux. Des fonctions comme la publication de l'adresses du client dans le DNS, ou comme la relative stabilité des adresses IP (redonner la même adresse lorsqu'un client revient) ne seront plus possibles. On n'a rien sans rien.

D'autre part, les clients utilisant ce profil consommeront davantage de ressources. Par exemple, à chaque changement au niveau 2, ils réclameront une nouvelle adresse IP, alors que l'ancienne reste réservée, jusqu'à la fin du bail.

Et certains serveurs DHCP fascistes ne donnent une adresse IP qu'aux clients connus, interdisant ainsi l'anonymat. Pour ces raisons, le RFC recommande aux programmeurs de permettre la configuration du client DHCP, configuration allant jusqu'à permettre de débrayer l'anonymisation.

Ce profil a été mis en œuvre dans un client Microsoft, ce qui a donné lieu à un compte-rendu d'expérience <https://www.ietf.org/proceedings/93/slides/slides-93-dhc-1.pdf> à une réunion IETF.