RFC 7819 : Privacy considerations for DHCP

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 3 mai 2016

Date de publication du RFC : Avril 2016

https://www.bortzmeyer.org/7819.html

Le protocole DHCP est bien connu : la grande majorité des machines « client » qui se connectent à l'Internet l'utilisent pour récupérer des éléments de configuration indispensables, comme l'adresse IP à utiliser. Mais peu de gens sont conscients que DHCP peut être dangereux pour la vie privée : le client DHCP n'est en effet pas passif, il envoie au serveur des informations qui peuvent permettre de suivre à la trace une machine mobile.

DHCP pour IPv4 est normalisé dans le RFC 2131 [1]. (Le RFC 7824 traite le cas de DHCP pour IPv6, qui pose des problèmes similaires.) Son principe de fonctionnement est simple : le client DHCP (la machine de M. Michu) envoie à la cantonade une requête pour demander des informations de configuration réseau, le serveur DHCP se reconnait, il répond avec ces informations. Voici une transaction DHCP, vue par tcpdump :

```
  21:32:13.284690 IP (tos 0x0, ttl 64, id 960, offset 0, flags [none], proto UDP (17), length 377)
    0.0.0.0.68 > 255.255.255.255.67: [udp sum ok] BOOTP/DHCP, Request from b8:27:eb:84:35:e3, length 349, xid 0x
  Client-Ethernet-Address b8:27:eb:84:35:e3
  Vendor-rfc1048 Extensions
    Magic Cookie 0x63825363
    DHCP-Message Option 53, length 1: Request
    Client-ID Option 61, length 19: hardware-type 255, eb:84:35:e3:00:01:00:01:c7:92:bc:8a:b8:27:eb:ba:90:94
    Requested-IP Option 50, length 4: 192.168.2.9
    MSZ Option 57, length 2: 1500
    Vendor-Class Option 60, length 46: "dhcpcd-6.9.0:Linux-4.4.8-2-ARCH:armv6l:BCM2708"
    Hostname Option 12, length 5: "amour"
    T145 Option 145, length 1: 1
    Parameter-Request Option 55, length 14:
      Subnet-Mask, Classless-Static-Route, Static-Route, Default-Gateway
      Domain-Name-Server, Hostname, Domain-Name, BR
      NTP, Lease-Time, Server-ID, RN
```

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc2131.txt

```
      RB, Option 119
    END Option 255, length 0
21:32:13.299825 IP (tos 0x0, ttl 64, id 0, offset 0, flags [DF], proto UDP (17), length 576)
    192.168.2.254.67 > 192.168.2.9.68: [udp sum ok] BOOTP/DHCP, Reply, length 548, xid 0x4feaaa6f, Flags [no
  Your-IP 192.168.2.9
  Client-Ethernet-Address b8:27:eb:84:35:e3
  Vendor-rfc1048 Extensions
    Magic Cookie 0x63825363
    DHCP-Message Option 53, length 1: ACK
    Server-ID Option 54, length 4: 192.168.2.254
    Lease-Time Option 51, length 4: 43200
    Subnet-Mask Option 1, length 4: 255.255.255.0
    Default-Gateway Option 3, length 4: 192.168.2.254
    Domain-Name-Server Option 6, length 8: 192.168.2.254,149.20.64.21
    BR Option 28, length 4: 192.168.2.255
    END Option 255, length 0
    PAD Option 0, length 0, occurs 264
```

Malheureusement, dans la requête du client se trouvent plein de détails sur la machine demanderesse (section 1 du RFC). Compte-tenu de la sensibilité aux problèmes de vie privée (RFC 6973) et de l'ampleur de la surveillance de masse exercée par les États (RFC 7258), il était nécessaire de limiter cette fuite d'informations. Ce premier RFC va décrire le problème, et proposer quelques pratiques qui le limitent. Le RFC 7844 détaille un profil DHCP qui limite sérieusement la fuite. Dans une prochaine étape, il y aura peut-être des modifications au protocole DHCP mais ce n'est pas encore fait.

Pour cette analyse de sécurité, la section 2 de notre RFC introduit la notion d'**identificateur stable**. Un identificateur stable ("*stable identifier*") est une information envoyée par le client DHCP qui change peu ou pas dans le temps (et qui peut donc permettre de tracer une machine mobile). La stabilité peut dépendre de la mise en œuvre logicielle utilisée. Ainsi, une adresse MAC est typiquement un identificateur stable mais, si macchanger <http://www.gnu.org/software/macchanger/> est utilisé, ce n'est plus le cas. Un identificateur stable n'est pas forcément mondialement unique.

Le gros de ce RFC est la section 3, qui liste les identificateurs envoyés par le client DHCP. Le plus évident, car il est prévu pour cela, est l'option DHCP "*Client Identifier*" (RFC 2131, section 2, et RFC 2132, section 9.14). Il est en général stable (le RFC 1533, prédécesseur du RFC 2132, recommandait même d'utiliser une adresse MAC, mais on voit parfois un nom de domaine, ou bien un DUID - décrit dans le RFC 4361). Même si on utilise un logiciel comme macchanger <http://www.gnu.org/software/macchanger/> pour changer d'adresse MAC, pas mal de mises en œuvre de DHCP utiliseront la valeur initiale et la stockeront... pour toujours.

Moins spectaculaire, plusieurs champs de la requête transportent des identificateurs fondés sur une adresse. C'est le cas de `yiadrr`, qui indique l'adresse IP actuelle du client ou `chaddr` qui indique l'adresse MAC. Plusieurs options font de même comme la "*Requested IP Address*" (qui permet d'indiquer l'adresse IP qu'on souhaiterait recevoir).

Autre option qui envoie un identificateur stable, et souvent unique, l'option "*Client FQDN*" (RFC 4702) qui transmet au serveur le FQDN.

Après les adresses et les noms de domaine, un autre danger se présente avec les options qui permettent au client d'indiquer le logiciel qu'il utilise. C'est le cas de l'option "*Vendor class*" (RFC 2132, section 9.13, une sorte d'équivalent du `User-Agent:` de HTTP, dans l'exemple plus haut, elle indique une machine ARM sous Linux), du "*Vendor-Identifying*" du RFC 3925, et de toutes les options "*Vendor-specific information*" (RFC 2132, section 8.4), qui peuvent indiquer le numéro de version du logiciel utilisé, sa configuration spécifique, etc. Certes, elles ne sont pas uniques (elles ne désignent pas une machine

particulière) mais elles font quand même fuiter de l'information sur le client et, combinées avec d'autres informations, elles peuvent mener à une identification unique. Une option analogue est "*Client System Architecture Type*" (RFC 4578) qui indique le type exact d'architecture pour les clients DHCP qui vont démarrer sur le réseau (par exemple avec PXE), en téléchargeant une version particulière du système d'exploitation. Si l'architecture utilisée est un peu rare, cette option donne des informations précieuses à un observateur.

En lisant cette liste, le paranoïaque peut se dire que la NSA a envoyé des gens à l'IETF pour faire normaliser le plus grand nombre possible d'extensions indiscrètes, de façon à être sûr d'identifier tous les clients DHCP observés. Il y a même une option pour indiquer l'adresse (civile, dans le monde physique, pas sur le réseau), "*Civic Location*", dans le RFC 4776. Il est vrai que, contrairement à la plupart des options listées ici, elle est fournie par le serveur et pas par le client, et ne permet donc pas de suivre un client à la trace. Mais elle peut permettre à un observateur de savoir où, dans le monde physique, se situe le client.

Outre tous ces champs et ces options par lesquels une information souvent excessive est transmise, DHCP dispose de certains mécanismes qui peuvent être utilisés pour compromettre la vie privée (section 4 du RFC). Par exemple, l'option "*Client FQDN*" du RFC 4702, citée plus haut, sert souvent à faire des mises à jour dynamiques du DNS (RFC 2136), et, dans ce cas, l'adresse IP du client DHCP (qui peut indiquer sa localisation) et son nom, identificateur stable, sont publiés dans le DNS, que tout le monde peut consulter. On peut connaitre les déplacements d'une machine simplement en consultant ce DNS public, sans avoir besoin d'espionner des milliers de réseaux. L'observateur peut être très discret et, en toute légalité, vous suivre.

Autre mécanisme dangereux, la stratégie d'allocation du serveur DHCP. Lorsqu'un client DHCP réclame une adresse IP, le serveur peut la choisir de plusieurs façons, et certaines ont des conséquences pour la vie privée :

— Allocation itérative : le serveur alloue les adresses IP dans l'ordre de la plage d'adresses dont il dispose. Elle est très simple et très rapide mais fournit des adresses IP qui sont très prévisibles. En outre, avec ce système, les premières adresses de la plage seront plus souvent utilisées, ce qui rendra encore plus facile des activités comme la reconnaissance d'un réseau avant une attaque.
— Allocation fondée sur un identificateur : le serveur a une table et y cherche un des identificateurs transmis par le client, allouant l'adresse trouvée dans la table. C'est très pratique pour l'administrateur système, car une machine donnée aura toujours la même adresse IP. Mais c'est nettement moins bon pour la vie privée du client, qui sera ainsi trivialement pistable, par son adresse IP fixe.
— Allocation aléatoire : le serveur DHCP prend au hasard une adresse IP libre. C'est sans doute la meilleure méthode, du point de vue de la protection de la vie privée.

La taille très restreinte de l'espace d'adressage IPv4 complique le problème, en limitant les possibilités du serveur d'utiliser certaines stratégies d'allocation, comme l'allocation fondée sur un identificateur (le serveur DHCP n'a pas forcément assez d'adresses IP pour tous ses clients potentiels.)

Bon, OK, le client DHCP envoie des tas d'identificateurs stables. Mais en quoi est-ce dangereux ? Qu'est-ce qu'un observateur indiscret peut en faire (section 5 du RFC) ? Déjà, il peut découvrir le type de machine qu'est le client (directement, via des options comme "*Vendor Class*", ou indirectement, via l'adresse MAC (dont le préfixe, l'OUI, dépend en général du matériel utilisé). L'espion peut aussi trouver le système d'exploitation utilisé.

Il peut aussi apprendre les réseaux que le client avait visité précédemment. Le client met en effet dans l'option "*Requested IP Address*" la précédente adresse IP qu'il avait obtenue. Si c'est une adresse publique (pas issue du RFC 1918), cela renseigne sur le précédent réseau utilisé.

L'observateur peut facilement trouver un identificateur stable à partir d'une requête DHCP (en combinant les options comme "*Client FQDN*"). Cela lui permet, s'il a accès au trafic de plusieurs serveurs

DHCP, de suivre les déplacements d'une machine (c'est ce qui se produit dans les cas de surveillance massive, cf. RFC 7258.)

Dans certains cas, l'observateur n'a même pas besoin d'être présent sur le réseau du serveur DHCP : si on fait des mises à jour dynamiques du DNS, il lui suffit de faire des requêtes DNS pour suivre les changements d'adresse IP (et donc de réseau) du client.

Bref, un client DHCP en dit en général trop à son serveur, et cela permet aux machines mobiles d'être facilement pistées. Notre RFC ne fournit pas de solutions immédiates, une solution est décrite dans le RFC 7844, d'autres feront l'objet d'un futur travail.