Google Chrome et son utilisation du DNS

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 8 mai 2016. Dernière mise à jour le 10 mai 2016

https://www.bortzmeyer.org/google-le-dns-et-les-medias.html

Une tribune sensationnaliste <http://www.lemonde.fr/idees/article/2016/05/06/comment-google-a-c
4914648_3232.html> dans le Monde le 6 mai prétendait que « Google [avait] changé l'Internet » et portait une accusation précise : le navigateur Google Chrome utiliserait une racine DNS spécifique à Google. Passons sur le fond politique de l'article, est-ce qu'au moins les faits allégués sont exacts ?

Un petit mot au passage sur ce concept de « racine » (les explications dans la tribune publiée dans le Monde sont... floues). Un logiciel qui veut utiliser un nom de domaine (par exemple un navigateur Web qui veut aller voir https://www.laquadrature.net/fr) va devoir interroger le DNS, un système technique qui permet, à partir d'un nom de domaine, d'obtenir des informations techniques indispensables (comme l'adresse IP du serveur). Les noms de domaine sont organisés sous forme d'**arbre**, avec la racine représentée en haut (contrairement aux botanistes, les informaticiens mettent la racine des arbres en haut). L'interrogation du DNS est faite par un logiciel nommé **résolveur**, qui est en général indiqué automatiquement à la machine (c'est la plupart du temps une machine du FAI, mais on peut utiliser son propre résolveur). Le résolveur commence par la racine, d'où il est aiguillé vers les TLD puis de fil en aiguille vers tous les autres noms de domaine. On voit donc que qui contrôle la racine contrôle le DNS.

L'article publié dans le Monde dit « le navigateur Chrome de Google remplace sans avertissement l[Caractère Unicode non montré [1]]annuaire de l[Caractère Unicode non montré]ICANN par le sien ». Au milieu de toutes les vagues affirmations de ce texte, voici un fait précis, qu'on peut contrôler. Cela prend cinq minutes à n'importe quel technicien. Avec un outil d'usage courant, tcpdump, on peut voir le trafic DNS de la machine. Lançons Chrome, tentons de visiter et observons ce trafic. (Les lignes suivantes sont triviales à interpréter pour le technicien ordinaire dont je parlais.)

```
11:38:24.549321 IP6 2001:691:1348:7::bad:209.27121 > 2001:67c:217c:6::2.53: 58483% [1au] A? nic.la. (35)
11:38:24.551217 IP6 2001:67c:217c:6::2.53 > 2001:691:1348:7::bad:209.27121: 58483- 0/8/9 (685)
```

1. Car trop difficile à faire afficher par LaTeX

Que voit-on ? La machine d'adresse IP 2001:691:1348:7::bad:209 a fait une requête DNS au serveur 2001:67c:217c:6::2, qui est bien un serveur de la racine (leurs adresses IP sont publiques et connues).

Amusant détail, je n'utilise personnellement pas la « racine ICANN » mais une racine « alternative », Yeti <https://yeti-dns.org/>. Donc, même si on a fait un choix différent de la majorité, Chrome respecte ce choix (en fait, Chrome ne parle pas directement à la racine : comme tout bon logiciel, il parle au résolveur configuré par l'utilisateur, et ce résolveur a été configuré pour utiliser telle ou telle racine). Le complot s'effondre donc : le principal reproche fait à Google dans cet article ne repose sur rien. Alors que la vérification des faits prend quelques minutes à n'importe quel technicien, les auteurs de la tribune n'ont pas vérifié, ni demandé à un technicien de le faire.

Est-ce un accident isolé dans une tribune par ailleurs raisonnable ? Non, il y a beaucoup d'autres erreurs techniques. Par exemple, les auteurs affirment « [la racine de Google est] plus rapide puisque souvent délivré d'étapes de contrôle et de surveillance ». Enregistrer les requêtes (comme le fait la NSA) ne ralentit rien, cette affirmation est donc dénuée de sens.

Autre exemple du manque de rigueur de cette tribune, la mention du .google comme étant « la racine de Google » alors qu'il s'agit d'un banal TLD... Est-ce une distinction importante <https://www.bortzmeyer.org/parties-nom-domaine.html> ou bien est-ce que je pinaille excessivement ? Oui, c'est important car cette erreur permet de juger le sérieux d'un texte. Imaginez un article portant sur la production laitière. L'article confond à un moment vaches et chèvres. Sans importance, puisqu'après tout ce sont bien deux mammifères produisant du lait ? Mais si un article sur ce sujet comportait une telle erreur, lui accorderait-on la moindre crédibilité ? C'est pareil ici : confondre TLD et racine, dans un article sur le DNS, est un signe d'extrême légèreté.

On notera aussi que des affirmations tout aussi fausses (« Cependant Google (et Chrome) utilise leur propre racine (copie de la racine ICANN) dont une est l[Caractère Unicode non montré]adresse IP 8.8.8.8 » se trouvent depuis longtemps sur le serveur Web de la société commerciale Open-Root <http://www.open-root.eu/decouvrir-open-root/les-racines-ouvertes/>, pour laquelle travaillent deux des auteurs. (8.8.8.8, alias Google Public DNS, est un résolveur DNS, qui n'a donc rien à voir avec une racine. Vaches et chèvres, encore.)

Ces erreurs techniques sont d'autant plus curieuses que, si on souhaite dire du mal de Google et mettre en garde contre le pouvoir dont jouit cette société, les exemples ne manquent pas. Google **a** un poids important, par sa part de marché, **peut** influencer ce que font les internautes, et ce poids important est certainement une question politique à traiter, comme toutes les fois où une boîte privée a trop de pouvoir. Alors, pourquoi rajouter des accusations fantaisistes ? Est-ce parce qu'il y a un risque que Chrome, un jour, fasse ce qui lui est reproché dans cette tribune ? Tout est possible mais ce ne serait pas très malin de la part de Google. En effet, l'utilisation d'une autre racine se voit (comme montré avec tcpdump plus haut). Ce ne serait pas très discret de la part de Google (la tribune parle de « substitution invisible », ce qui n'a pas de sens, vu cette visibilité du trafic réseau). D'autre part, pas mal d'entreprises utilisent en interne des TLD non-existants comme .corp, qui ne sont connus que de leurs résolveurs. Si Chrome n'utilisait pas ces résolveurs, ces TLD ne fonctionneraient pas.

Autre chose frappante, ce texte a largement circulé sur les réseaux sociaux, étant repris sans nuances, sans aucune vérification. On l'a par exemple vu cité sur le site de l'émission Arrêt sur Images <https://www.arretsurimages.net/breves/2016-05-06/Google-vers-un-annuaire-parallele-d-Interne
Apparemment, le fait d'être publié dans le Monde a un poids qui suffit à débrayer toute critique. (Alors qu'il s'agissait d'une tribune et pas d'un article, ce qui indique que le Monde n'a pas exercé de contrôle trop serré sur le contenu.)

Depuis, il y a eu une réponse à ce texte <http://www.open-root.eu/la-doc/realites-et-affabulations/he-oui-google-a-son-propre-dns/>. Elle a été annoncée sur Twitter en utilisant le raccourcisseur d'URL de Google (goo.gl). Le site d'Open-Root utilise d'ailleurs le "*spyware*" Google Analytics. Cela indique leur niveau de cohérence (mais c'est courant chez les souverainistes : de grands discours anti-GAFA à la tribune, mais l'utilisation de tous leurs outils en pratique.) Sans compter la paranoïa « nous sommes attaqués de toute part », de la part de gens qui passent dans le Monde ou à la télé quand ils veulent, et qui sont invités à tous les colloques ministériels...

Bon, je critique cette réponse sur la forme, mais sur le fond ? Sur le fond, il n'y a rien à dire, parce qu'il n'y a pas d'éléments nouveaux : la réponse cite comme preuve que Google a son propre DNS... l'existence de Google Public DNS <https://www.bortzmeyer.org/google-dns.html>. Confondre un résolveur DNS public avec une racine, c'est comme l'expert en élevage cité plus haut qui confondrait vache et chèvre... C'est tellement désespérant que c'est au delà de toute discussion.

Pour les reste, pas de détails, pas de tcpdump montrant Chrome faire des requêtes à la soi-disant racine Google, rien.

---

https://www.bortzmeyer.org/google-le-dns-et-les-medias.html