RFC 5782 : DNS Blacklists and Whitelists

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 19 février 2010. Dernière mise à jour le 25 février 2010

Date de publication du RFC : Février 2010

https://www.bortzmeyer.org/5782.html

Le spam est un tel problème pour le courrier électronique d'aujourd'hui, que de nombreuses techniques ont été développées pour le limiter. Comme toujours en sécurité informatique, il n'y a pas de solution idéale, juste des compromis entre trop de sécurité (et on bloque du courrier légitime) et pas assez (et on est noyés par le spam). Dans cet arsenal de techniques, une des plus courantes et des plus contestées est la liste noire, ou DNSBL ("*DNS black list*"). Bien que très largement déployée, elle n'avait jamais fait l'objet d'une documentation formelle, ce qui est désormais le cas, avec ce RFC.

Le principe d'une DNSBL est de distribuer, via le DNS, de l'information sur des machines (identifiées par leur adresse IP) qui ont envoyé du spam, ou bien sont susceptibles de le faire. Le problème n'est pas tant dans la technique utilisée (même si certains déplorent qu'on charge la barque du DNS en lui confiant de plus en plus de missions) que dans la gestion de ces listes. La grande majorité sont gérées de manière opaque, par des gens irresponsables, qui inscrivent ou désinscrivent un peu comme ça leur chante.

Ce RFC a été développé par le groupe de travail IRTF ASRG <http://asrg.sp.am/>. L'auteur du RFC, John Levine, est un des porte-paroles les plus fréquents des « éradicateurs », ceux qui sont prêts à faire beaucoup de dégâts collatéraux pour lutter contre le spam. Le RFC a le statut de « pour information », utilisé pour documenter les techniques qui sont utilisées et donc intéressantes à connaître, mais qui ne sont pas forcément approuvées par l'IETF. Il a failli être sur le chemin des normes, à la demande de l'ASRG, mais en a été retiré suite à une virulente discussion <http://www.ietf.org/mail-archive/web/ietf/current/msg53722.html> sur la liste principale de l'IETF.

Le document commence (section 1) par un peu d'histoire. En 1997, la première liste noire, RBL ("*Real-time blackhole list*") avait été créée par Paul Vixie et Dave Rand. Elle était distribuée avec BGP, le DNS n'est venu qu'après. (Tout le monde a un client DNS, ce qui n'est pas le cas de BGP.) Elle existe toujours sous le nom de MAPS même si elle n'a plus grand'chose à voir avec l'effort du début.

Le terme de RBL étant désormais une marque déposée, le RFC suggère de parler plutôt de DNSBL ("*DNS Black List*") et, pour celles qui sont utilisées pour autoriser plutôt que pour interdire, de DNSWL ("*DNS White List*"). Pour parler des deux simultanément, le RFC utilise DNSxL. (Notons que la différence entre les deux est uniquement dans l'usage que le client en fait, cf. section 2.2.)

Ce RFC décrit le fonctionnement technique des DNSxL et le protocole avec lequel elles communiquent leurs résultats. Il ne parle pas des questions de la maintenance de ces listes, ni de leur politique d'inscription/désinscription, qui font l'objet d'un autre document, le RFC 6471 [1].

La section 2 attaque la technique : une DNSxL est une zone du DNS, où les noms sont formés à partir de la clé d'accès à la liste. Cette clé est presque toujours une adresse IP (la section 3 traite des listes de noms) et le mécanisme d'encodage est inspiré de celui de `in-addr.arpa` (section 2.1) : on inverse les octets et on ajoute le nom de la liste. Ainsi, pour chercher si `192.0.2.129` est listé dans la DNSxL `list.example.com`, on fera une requête DNS pour `129.2.0.192.list.example.com`. Si `192.0.2.129` est listé, la zone doit contenir un enregistrement de type A (normalement, « adresse », mais utilisé dans un autre sens par les DNSxL) et peut contenir un enregistrement de type TXT qui stockera un message qu'on pourra afficher à l'utilisateur bloqué. Les données dans une DNBxL sont donc des données DNS ordinaires et on peut donc y accéder par des clients DNS comme dig, ici avec la liste `sbl-xbl.spamhaus.org` :

```
% dig A 129.2.0.192.sbl-xbl.spamhaus.org
...
;; ->>HEADER<<- opcode: QUERY, status: NXDOMAIN, id: 2015
...
```

Ici, on récupère le code NXDOMAIN ("*No Such Domain*") donc cette adresse IP n'est pas sur la liste. Essayons avec une adresse qui est sur la liste :

```
% dig 158.2.0.192.sbl-xbl.spamhaus.org
...
;; ANSWER SECTION:
158.2.0.192.sbl-xbl.spamhaus.org. 684 IN A     127.0.0.4
```

La valeur de l'enregistrement DNS est typiquement `127.0.0.2` mais d'autres sont possibles (section 2.3).

En IPv6, pas de grosses différences, mais notons que, ici, le RFC fait œuvre créative car il n'existait pas encore de DNSxL IPv6 (pour IPv4, le RFC documente un état de fait). La section 2.4 précise les détails. Depuis, la liste Virbl a ajouté IPv6 <`http://virbl.bit.nl/`>.

C'est tout pour le principe. Parmi les détails, la section 5 couvre le cas où une DNSxL arrête de fonctionner correctement et ne contient plus aucun enregistrement (ou au contraire répond systématiquement quelle que soit l'adresse, cas qui s'est déjà produit). La liste doit contenir une entrée pour `127.0.0.2` et ne **pas** en contenir pour `127.0.0.1`. Tester ces deux clés permet de voir si le fonctionnement technique de la liste est correct.

Comment utilise t-on une DNSxl ? La section 6 couvre la configuration des clients. Par exemple, avec le MTA Postfix, la configuration se fait ainsi :

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc6471.txt`

```
smtpd_recipient_restrictions = permit_mynetworks, reject_unauth_destination,
            reject_rbl_client sbl-xbl.spamhaus.org, ...
```

et l'examen du journal permet de voir les rejets de certains émetteurs :

```
Jan 20 21:28:02 lilith postfix/smtpd[7442]: NOQUEUE: reject: RCPT from unknown[192.0.2.158]: 554 5.7.1 Service u
```

La configuration ci-dessus est de tout ou rien. Si l'adresse IP du client SMTP est dans la DNSBL, il est rejeté. Comme le rappelle le RFC, une autre utilisation est de tester plusieurs listes noires et de faire un calcul de pondération entre elles, ne rejetant le message que si l'émetteur se trouve dans plusieurs listes. C'est ce que fait SpamAssassin, par exemple. Cela protège contre les listes excessivement zélées (ou contre les listes mortes <http://spamlinks.net/filter-dnsbl-dead.htm>, qui peuvent répondre positivement pour toutes les adresses soumises, entraînant des rejets systématiques).

Et comment savoir si on est dans une liste noire ou pas ? Il existe des outils pratiques et j'en recommande deux : rblcheck <http://rblcheck.sourceforge.net/> qui permet d'interroger plusieurs listes noires d'un coup et check-auth@verifier.port25.com, une adresse de courrier à laquelle on peut écrire, et qui vous renvoie automatiquement un rapport détaillé sur votre message et les machines qui l'ont transmis, y compris leur éventuelle présence dans les listes noires. Mais on peut aussi citer (accès Web uniquement) MultiRBL <http://multirbl.valli.org/lookup/> (qui permet de construire des URL REST comme http://multirbl.valli.org/lookup/198.51.100.1.html), Robtex <http://www.robtex.com/> (même remarque sur les URL) ou MXtoolbox <http://www.mxtoolbox.com/blacklists.aspx> ou encore Blacklist check <http://whatismyipaddress.com/blacklist-check>.

La section 7, enfin, couvre le problème de sécurité général. Comme le résume bien le RFC, utiliser une liste noire externe à votre organisation, c'est sous-traiter la sécurité à un tiers. Celui-ci peut faire un meilleur travail que vous (la plupart des listes noires maintenues localement sont statiques et jamais mises à jour <https://www.bortzmeyer.org/pas-de-listes-noires-statiques.html>) ou au contraire un bien pire. Il faut donc être bien conscient de ses responsabilités. Contrairement à ce que prétendent les porte-paroles officiels des gros FAI, les listes noires ne sont pas parfaites : elles comprennent des faux négatifs (spammeurs non listés) et des faux positifs (innocents listés). Tenir à jour ces listes est un gros travail et personne ne peut croire à la vérité des affirmations d'un gros FAI, dont un porte-parole à l'IETF prétend lire et traiter tous les rapports de faux positifs envoyés par ses clients.

Idéalement, l'administrateur système qui configure son serveur de messagerie pour utiliser une liste noire devrait passer du temps à s'informer sur la ou les listes envisagées, étudier leur politique et leur pratique, et surtout suivre son évolution dans le temps puisque les listes changent. Mais très peu le font.

Ne nous faisons pas d'illusion : l'Internet n'est pas un monde de gentils Bisounours où tout le monde coopère pour le bien commun, c'est un espace de concurrence et la polémique à l'IETF reflétait simplement la différence entre d'une part les gros FAI, plutôt éradicateurs, qui connaissent et regardent les autres de haut (la remarque « vous ne gérez pas des millions de boîtes aux lettres, donc taisez-vous » était la plus fréquente lors de la discussion) et d'autre part les petits FAI et les utilisateurs, bien forcés de subir <http://www.ietf.org/mail-archive/web/ietf/current/msg54453.html>.

Un rapport intéressant, dû à Bruno Rasle et Frédéric Aoun, avait bien montré le manque de sérieux de tant de listes noires : « Blacklists anti-Spam : plus de la moitié des entreprises indexées <http://www.

halte-au-spam.com/ddm-blacklists.htm> ». Plus récemment, on peut aussi lire un intéressant article sur la mésaventure survenue à Gandi avec SORBS <http://www.lebardegandi.net/post/2010/03/05/Blackliste-par-SORBS-un-jour-dans-la-vie-de-Gandinet>.

Notons pour terminer que les DNSxL posent d'autres problèmes de sécurité que les faux positifs : comme elles sont consultées à chaque message, le gérant de la DNSxL peut avoir une idée de qui sont vos correspondants, et le DNS n'étant pas très sûr, les DNSxL ne le sont pas non plus (aujourd'hui, aucune n'est signée avec DNSSEC).

Merci à Olivier Fauveaux et Serge Aumont pour leurs remarques et corrections.