RFC 6686 : Resolution of The SPF and Sender ID Experiments

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 21 juillet 2012

Date de publication du RFC : Juillet 2012

https://www.bortzmeyer.org/6686.html

L'IETF avait créé en 2004 un groupe de travail nommé MARID <http://tools.ietf.org/wg/marid/>, qui avait pour tâche de normaliser un mécanisme d'**authentification faible** du courrier électronique par le biais de la publication dans le DNS des serveurs autorisés à envoyer du courrier pour un domaine. Deux propositions étaient sur la table, SPF et Sender ID. Microsoft avait pratiqué une intense obstruction contre SPF, pour promouvoir sa propre solution, Sender ID. En 2006, l'IETF avait cédé devant cette obstruction, en renonçant à normaliser SPF, déjà très répandu, et en publiant les deux protocoles avec le statut Expérimental : RFC 4408 [1] pour SPF et RFC 4406 pour Sender ID. Cette solution peu courageuse avait été contestée, notamment pour le risque que la coexistence de deux protocoles incompatibles faisait peser sur la bonne délivrance du courrier. Officiellement, la retraite de l'IETF devait être provisoire : une période de deux ans d'observation était prévue, suite à laquelle des conclusions pourraient être tirées sur l'expérience. En fait, c'est seulement six ans après qu'est publié notre RFC 6686 qui conclut enfin officiellement que Sender ID est un échec total et que seul SPF est déployé et utilisé.

Le groupe de travail spfbis <http://tools.ietf.org/wg/spfbis> avait été constitué en février 2012 pour documenter les preuves du déploiement de SPF et pour réviser la spécification. Ce document est son premier RFC, et il n'a pas été facile à produire, tant les controverses étaient vives. D'une certaine façon, c'est la création de ce groupe qui avait marqué le choix définitif en faveur de SPF. Mais, en 2006, le consensus n'était pas évident, notamment face à l'opposition acharnée de Microsoft, qui tenait à pousser sa solution. Il était pourtant déjà clair à l'époque que SPF était largement déployé et que Sender ID n'était pas utilisé, même par les services appartenant à Microsoft comme Hotmail. Mais la politique politicienne l'avait emporté et l'IETF avait donc fait semblant de traiter les deux protocoles sur un pied d'égalité.

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc4408.txt

Avantage de ce délai : il a permis de récolter beaucoup plus d'informations pratiques. C'est l'une des forces de ce RFC 6686 que de présenter de nombreuses données issues de plusieurs campagnes de mesure.

Notez aussi que ce RFC ne fait pas de comparaison technique des deux protocoles. Il considère que le choix a déjà été fait par la communauté et il se contente donc d'observer ce qui tourne réellement dans la nature.

Avant les chiffres, qui constituent l'essentiel de ce document, un petit rappel : le RFC 4408 introduisait un nouveau type de données DNS, le type `SPF` (code 99). Historiquement, SPF utilisait le type générique `TXT` (code 16) et le nouveau type `SPF` avait été ajouté alors que le déploiement était déjà largement commencé. Une question secondaire à laquelle notre RFC répond est donc « pour ceux qui utilisent SPF, avec quel type DNS ? ».

La section 3, le cœur de notre RFC, présente donc les résultats de différentes mesures. D'abord, trois séries de mesures actives (faites en interrogeant le DNS sur son contenu, comme peut le faire un logiciel comme DNSdelve <`http://www.dnsdelve.net/`>) ont mesuré le pourcentage de domaines qui publiaient des enregistrements SPF (commençant par `v=spf1`) ou Sender ID (commençant par `spf2.0/`, oui, Microsoft était allé jusqu'à détourner le terme « SPF » pour brouiller davantage la discussion ; il était prévu de critiquer ce point dans le RFC mais le texte a dû être retiré pour ne fâcher personne).

La première étude avait été faite par Cisco à partir du million de domaines les plus populaires sur le Web selon Alexa. 39,8 % ont un enregistrement TXT d'authentification du courrier électronique. Seuls 1,3 % de ces enregistrements sont pour Sender ID.

La deuxième étude a été faite par le Trusted Domain Project <`http://www.trusteddomain.org/`>, en utilisant comme liste de domaines des extraits des journaux de leurs serveurs de messagerie. Elle trouve 56,4 % de TXT d'authentification du courrier (dont 4,6 % pour Sender ID). Cette étude a également observé que certains serveurs DNS cassés répondent pour une demande de type TXT mais pas pour une demande de type SPF (ce point est couvert plus en détail par la suite, dans l'annexe A).

La troisième étude a été faite par Hotmail, à partir des domaines qui leur envoient du courrier. 46,2 % de TXT avec authentification. Hotmail, filiale de Microsoft, n'a pas publié le pourcentage de Sender ID parmi eux.

En étudiant le contenu de la zone `.fr` (non mentionnée dans le RFC, étude faite à l'AFNIC avec le logiciel DNSdelve <`http://www.dnsdelve.net/`>), on trouve 20 % de domaines ayant un enregistrement SPF. Et aucun avec Sender ID... (la mesure a été faite en échantillonant la zone donc une poignée de domaines ayant Sender ID a pu échapper à la mesure). 70 % de ces domaines SPF sont dus à un seul hébergeur de courrier. À noter que cette étude n'a pas fait de différence entre les enregistrements DNS `TXT` et `SPF`.

Ces trois études portaient sur les enregistrements dans le DNS. Et pour les requêtes DNS, que voit-on ? Des mesures passives ont été faites (mais le RFC ne contient guère de détails), montrant très peu de requêtes de type SPF (99) et depuis un très petit nombre d'acteurs. À noter qu'on ne peut en général pas savoir, en envoyant ou en recevant du courrier, si le pair SMTP utilise SPF et/ou Sender ID.

Le serveur de noms de `.fr` `d.nic.fr` voit, sur ses différentes instances "*anycast*", 0,7 % de requêtes `TXT`, et 0,1 % de requêtes `SPF` (données analysées par DNSmezzo <`http://www.dnsmezzo.net/`>). `SPF` est quand même loin devant un type comme `NAPTR` mais il est nettement dominé par `TXT` (toutes

les requêtes TXT ne sont pas forcément pour SPF). L'abandon du type DNS SPF a ensuite été repris dans la nouvelle norme SPF, le RFC 7208.

Reste la question des programmes qui mettent en œuvre SPF et/ou Sender ID (section 3.2). L'offre logicielle pour Sender ID est recensée en <http://www.microsoft.com/mscorp/safety/technologies/senderid/support.mspx>. Aucun logiciel libre n'y est cité. Parmi les treize opérateurs mentionnés, au moins un n'utilise plus Sender ID (la dernière mise à jour date de 2007).

Celle de SPF est en <http://www.openspf.net>. Six bibliothèques sont citées, vingt-deux MTA (à la fois en logiciel libre et en privateur), ainsi que des "*patches*" pour les autres.

Il y a quand même un moyen de tester la présence de Sender ID dans un MTA, via l'extension SMTP SUBMITTER (RFC 4405). Elle est utilisée pour optimiser l'algorithme PRA du RFC 4407 en annonçant au récepteur l'identité de l'expéditeur. Deux logiciels la mettent en œuvre, Santronics <http://www.santronics.com/> WinServer et McAfee MxLogic. Les journaux des fournisseurs montrent qu'environ 11 % des sessions SMTP utilisent cette extension, mais sans qu'on sache s'il s'agit d'un logiciel très répandu ou de plusieurs logiciels moins populaires.

Le Trusted Domain Project a donc fait des mesures actives, se connectant à de nombreux MTA pour leur demander les extensions qu'ils géraient. 4,7 % annonçaient l'extension SUBMITTER et presque tous étaient du MxLogic.

Restait une question délicate à étudier : le refus de l'IETF de trancher entre les deux propositions SPF et Sender ID laissait ouverte la possibilité que ces deux protocoles donnent des résultats différents pour un même message. Par exemple que l'un l'accepte et que l'autre le refuse. Cela serait certainement gênant, l'acceptation ou le rejet d'un message dépendant du choix technique fait par le récepteur. D'abord, deux études par Hotmail et par Trusted Domain Project ont montré que, dans environ 50 % des cas, le PRA (l'identité utilisée par Sender ID, cf. RFC 4407) et le MAIL FROM SMTP (l'identité utilisée par SPF) sont les mêmes.

Ensuite, le Trusted Domain Project a analysé 150 000 messages reçus et abouti à la conclusion que, dans 95 % des cas, SPF et Sender ID auraient formé le même diagnostic (acceptation ou rejet). Une analyse analogue faite par Hotmail sur un échantillon plus large (des millions de messages) trouvait moins : 80 %.

Bon, en pratique, on a rarement vu de rapports comme quoi des ingénieurs système avaient été perturbés par les différences restantes.

La section 5 analyse tous ces résultats et synthétise :
— Les enregistrements DNS de type SPF (99) n'ont pas été adoptés et des problèmes techniques à l'utilisation d'enregistrements d'un type « nouveau » demeurent.
— Les enregistrements TXT publiés ne le sont que rarement pour Sender ID.
— Il n'y a pas de données indiquant qu'en pratique, un des deux protocoles donnerait des résultats bien meilleurs qu'un autre.
— Les mises en œuvre de SPF sont plus nombreuses et bien mieux mantenues. L'extension SMTP SUBMITTER est rare.

En conclusion, la section 6 estime que :
— On a assez de données issues de l'expérience pour arriver à une conclusion.
— Le type d'enregistrement DNS numéro 99, SPF, ne sert à rien.
— Presque personne n'a déployé Sender ID,

— Au contraire de SPF, très répandu. Malgré son statut officiel « Expérimental », il est bien plus déployé que certains protocoles normalisés.

Le problème des types de données DNS, déjà mentionné plusieurs fois ci-dessus, est traité en détail dans l'annexe A. SPF avait été créé en dehors de l'IETF et ses concepteurs n'avaient pas suivi la voie recommandée, qui consistait à utiliser un type d'enregistrement DNS spécifique. Au lieu de cela, ils s'étaient servis d'un type existant, et perçu comme généraliste, `TXT` (code 16). Il faut dire qu'à l'époque (le RFC ne le dit pas franchement), obtenir un nouveau type d'enregistrement DNS à l'IANA était très difficile.

Lorsque SPF est arrivé à l'IETF, il y avait déjà une forte base installée, et donc une grande difficulté à migrer vers un nouveau type d'enregistrement. Ce nouveau type, `SPF` (code 99), était accompagné lors de sa création d'un plan de migration, qui a été un échec.

Plusieurs raisons à cet échec sont identifiées par notre RFC 6686 :

— Comme déjà signalé dans la section 3, des pare-feux configurés avec les pieds par un stagiaire refusent parfois les types DNS inconnus (donc le `SPF`), en violation du RFC 3597,
— Le DNS n'est pas composé que de serveurs de noms, il y a aussi tout un ensemble de logiciels autour (par exemple les interfaces Web de gestion du contenu des zones, ou bien les bases de données qui stockaient les zones) qui auraient dû être modifiés pour adopter le nouveau type,
— La base installée utilisait `TXT` et ne voyait pas de raison impérative de migrer (« faire plus propre » n'est pas une raison impérative),
— Le plan était lui-même erroné : permettant au serveur de ne publier qu'un seul des deux types, et au client de n'en demander qu'un seul, il faisait que deux logiciels parfaitement conformes au RFC pouvaient être dans l'impossibilité de communiquer.

Depuis, des choses ont changé. L'allocation de nouveaux types, très pénible, a été considérablement libéralisée par le RFC 6195, suivant les recommandations du RFC 5507. Mais le déploiement effectif des nouveaux types reste un problème difficile dans l'Internet d'aujourd'hui.