RFC 5585 : DomainKeys Identified Mail (DKIM) Service Overview

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 8 juillet 2009

Date de publication du RFC : Juillet 2009

https://www.bortzmeyer.org/5585.html

Le système DKIM de signature numérique du courrier électronique est normalisé dans le RFC 6376 [1]. Ce n'est pas une norme simple et, comme le domaine dans lequel elle se situe est traditionnellement très délicat et voué aux polémiques vigoureuses, un effort d'explication était nécessaire. C'est ainsi qu'est né ce RFC 5585, qui donne une description de haut niveau de DKIM, en se focalisant sur le protocole (les aspects opérationnels ne sont pas évoqués).

Le principe de DKIM est de signer cryptographiquement les messages, en utilisant le DNS comme serveur de clés. La signature permet de lier, de manière fiable, un message à une organisation, celle qui gère le nom de domaine où a été trouvé la clé publique. Il y a très loin de cette liaison à la lutte contre le spam : DKIM n'est qu'une technique d'authentification, il ne peut donc pas résoudre tous les problèmes posés par les usages malveillants du courrier. Pour citer le RFC, « DKIM est une seule arme, dans ce qui doit être un vaste arsenal ».

La section 1 du RFC résume les principes de base de DKIM. Les attaques auxquelles DKIM permet de répondre ont été documentées dans le RFC 4686.

Une notion centrale est celle d'**identité**. Une personne ou une organisation a une identité et le but de DKIM est d'associer le message à une telle identité. Notons bien que, en soi, cela ne signifie pas que le message soit meilleur ou davantage valable, uniquement qu'il peut être rattaché à une personne ou une organisation. DKIM ne dit rien des qualités ou des défauts de cette personne ou de cette organisation, c'est une technique d'authentification, pas d'autorisation <https://www.bortzmeyer.org/authentifier-et-autoriser.html>.

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc6376.txt

Pour DKIM, l'identité est un nom de domaine comme `pape.va`, `enron.com`, `ministere-de-l-harmonie.cn` ou `enlargeyourpenis.biz`. Ce nom de domaine est désigné par le sigle SDID ("*Signing Domain IDentifier*"). De même que BGP transmet des informations entre AS, la confiance ne régnant qu'à l'intérieur d'un AS, DKIM signe du courrier entre **ADMD** ("*ADministrative Management Domain*"), un ADMD étant une organisation (parfois informelle) à l'intérieur de laquelle la confiance règne - alors qu'il n'y a évidemment pas, "*a priori*", de confiance entre deux ADMD (annexe A.2).

La section 1.1 rappelle ce que DKIM ne fait **pas** :
— DKIM ne garantit pas tout le message, uniquement ce qui est signé (ce qui peut ne pas inclure certaines en-têtes, ou une partie du corps),
— DKIM ne dit rien sur le caractère du signataire. Si un message proposant des modifications de l'anatomie prétend être émis par `enlargeyourpenis.biz`, DKIM permet de vérifier cette prétention, il ne teste pas l'efficacité du médicament promu,
— DKIM n'impose pas de politique quant au traitement à faire subir au message, après la vérification.

DKIM n'est pas, et de loin, la première technique mise au point pour augmenter la sécurité du courrier électronique. La section 1.2 examine les autres candidats (inutile de dire que la comparaison faite par les auteurs de DKIM est systématiquement défavorable à ces malheureux concurrents). Par exemple, SPF (RFC 7208) a des points communs avec DKIM, utilise également le nom de domaine comme identité (contrairement à ce que prétend ce RFC 5585) mais est également lié à l'adresse IP de l'émetteur, rendant difficile certaines délégations (par exemple, faire suivre un message).

Pas moins de quatre autres normes IETF utilisent une signature du message, l'ancêtre PEM (RFC 989), PGP (RFC 4880), MOSS (RFC 1848) et S/MIME (RFC 3851). Seuls PGP (plutôt apprécié dans le monde technique) et S/MIME (plutôt apprécié dans l'entreprise sérieuse et cravatée) ont connu un déploiement significatif mais qui, dans les deux cas, reste très loin du niveau de généralité qui serait nécessaire.

Outre le fait que la base installée était bien mince, le choix de DKIM de partir de zéro s'appuyait sur des arguments techniques. Ainsi, DKIM ne dépend pas, pour juger d'une clé, de signatures de cette clé (contrairement aux certificats X.509 de S/MIME ou au réseau de confiance de PGP) mais uniquement de sa disponibilité dans la DNS. DKIM n'a ainsi pas besoin d'une nouvelle infrastructure (comme le sont les serveurs de clés pour PGP).

Après ce tour d'horizon, le RFC expose les services que rend DKIM, en section 2. Il y a deux services importants, vérifier une identité et l'évaluer, juger de sa crédibilité et de son sérieux. DKIM fournit le premier service (section 2.1) et permet le second, dont il est un pré-requis.

Cette évaluation (section 2.2) n'est pas directement faite par DKIM. Ce dernier ne peut pas répondre à la question « Est-ce qu'un message venu de `enron.com` ou `france-soir.fr` mérite d'être délivré ? ». Tout ce que DKIM pourra faire est de garantir que cette identité est authentique et que le message n'a pas été modifié en route (section 2.3). Mais, comme le dit le RFC, « si le message était mensonger au début, il le sera toujours après la signature, et DKIM permettra de garantir que le mensonge a été transmis fidèlement ».

À noter que l'absence d'une signature DKIM peut aussi bien signifier que le domaine (le SDID) en question ne signe pas, que d'indiquer une attaque active où le méchant a retiré la signature. La protection contre ces attaques sera assurée par la future norme sur les règles de signature ("*Signing Practices*"), qui permettra aux gérants d'un domaine d'indiquer dans le DNS si les messages émis par leur organisation sont systématiquement signés ou pas.

La section 3 résume le cahier des charges que suivait DKIM. Parmi les buts de ce dernier, pour ce qui concerne le protocole (section 3.1) :
— Une vérification du seul nom de domaine, pas de l'adresse entière (comme le fait PGP),

— La possibilité de signer à n'importe quel endroit de la chaîne, pas uniquement dans le MUA (pour tous ces termes techniques, l'annexe A du RFC rappelle leur sens) de l'émetteur (là encore, contrairement à PGP),
— Permettre de déléguer la signature à des tiers.

Et parmi les buts plus opérationnels (section 3.2) :

— Ne pas être intrusif pour les lecteurs qui ne gèrent pas DKIM : concentrée dans les en-têtes, la signature n'est guère visible pour eux,
— Permettre un déploiement progressif, sans exiger de tout le monde qu'il adopte le nouveau gadget immédiatement.

Comment DKIM atteint-il ces buts ? La section 4 expose les fonctions qu'il assure. Le principe de base est que le signeur choisit le SDID (le nom de domaine qui identifie l'origine du message), un sélecteur (une chaîne de caractères arbitraire) et signe le message, mettant la signature dans un en-tête `DKIM-Signature:`. Le récepteur du message trouve la clé publique en interrogeant le DNS pour un nom de domaine qui est formé par l'ajout du sélecteur au SDID. Il peut alors vérifier la signature.

Ainsi, si je reçois un message d'un utilisateur de Google Mail contenant :

```
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
        d=gmail.com; s=gamma;
        ...
```

je sais que ce message vient de `gmail.com` (quoi que puissent dire les autres en-têtes) et que, comme le sélecteur est `gamma`, je peux trouver la clé publique en interrogeant `gamma._domainkey.gmail.com`.

Pour mieux comprendre où se situent ces différentes fonctions, on peut regarder le joli diagramme de la section 5, qui représente graphiquement l'architecture de DKIM (il vaut la peine de lire également l'annexe A, qui résume l'architecture du courrier électronique). Un autre point important de cette section est qu'actuellement, l'**absence** d'une signature dans un message ne prouve rien. Elle peut indiquer que l'émetteur ne signe pas, mais aussi qu'un méchant a modifié un message (ou inséré un faux). Dans l'état actuel du déploiement de DKIM, il est donc difficile de prendre des décisions lorsqu'un message n'est pas signé. Le problème sera traité par le futur RFC sur les règles de signature ("*Signing Practices*"), qui permettra à un domaine de publier les règles qu'il a choisi et qu'il suit en matière de signature DKIM. Un domaine pourra ainsi annoncer au monde qu'il signe systématiquement, ce qui permettra de rejeter les messages non signés mais prétendant venir de ce domaine.

Un article récent de Cisco sur le déploiement de DKIM <`http://blogs.cisco.com/news/comments/domainkeys_identified_mail_dkim_grows_significantly/`> indiquait une croissance rapide de son utilisation. Enfin, on peut trouver davantage d'informations sur le site officiel <`http://dkim.org`>.