Lutte anti-spam concrète et pratique avec du logiciel libre

Stéphane Bortzmeyer
Afnic

Pierre David
Centre Réseau Communication - Université Louis Pasteur

Mots clefs

Spam, sendmail, postfix, greylisting, bayes

Résumé

Le spam, par la consommation de ressources matérielles et humaines qu'il nécessite, par le temps qu'il fait perdre à ses lecteurs involontaires, par l'agacement ou les craintes qu'il suscite, par le danger de certains messages qu'il transporte (escroqueries, parexemple) est un des principaux fléaux de l'Internet aujourd'hui.

Il existe d'innombrables colloques et commissions sur le spam. Il y a également beaucoup de produits commerciaux payants pour le filtrer, la plupart n'expliquant pas du tout leur fonctionnement et donc leurs limites ou leurs risques, alors qu'aucune solution technique n'est parfaite, que toutes ont leurs problèmes.

Mais il existe peu de documentations **pratiques** et **holistiques** sur la lutte anti-spam telle que doit la pratiquer l'ingénieur système moyen, qui ne peut pas attendre la signature d'un Traité International ou le résultat d'un procès, et qui doit ralentir le flot aujourd'hui.

Chaque logiciel anti-spam libre, et il y en a beaucoup, vient avec sa documentation mais les documentations transversales, sur la combinaison de ces logiciels, sur leur choix, sont beaucoup plus rares. Etant donné qu'il est largement acquis que la lutte contre le spam ne peut pas reposer sur une seule technique, ce tutoriel va tenter de présenter l'installation et la configuration d'un ensemble de techniques efficaces pour lutter contre le spam **entrant**.

Il ne tentera pas de faire le tour de toutes les méthodes, seulement de celles qui sont raisonnables pour l'administrateur système d'aujourd'hui. Cette tâche de filtrage des solutions est une composante importante de ce tutoriel car le monde de la lutte contre le spam comprend plus de rebouteux que d'experts.

Toutes les techniques seront illustrées par des exemples concrets avec les serveurs Postfix et sendmail, les deux logiciels libres de loin les plus répandus pour gérer le courrier.

Le tutoriel s'adresse à des administrateurs système en charge d'un serveur de messagerie Unix. Une compétence minimum en configuration de Postfix ou bien de sendmail est nécessaire.

Le tutoriel se composera donc des parties suivantes :

Rappel rapide de l'architecture du courrier, de l'adressage, de SMTP et du format des messages, notamment les en-têtes « d'identification » comme From ou Sender.

Comment fonctionne techniquement le spam (avec notamment le rôle des botnets, réseaux de machines zombies, et l'utilisation de logiciels spécialement développés pour le spam, non compatibles avec les RFC).

Rappel rapide de quelques principes de la sécurité informatique, notamment l'analyse coût / bénéfice : toute technique anti-spam laisse passer des spams (faux négatifs) et arrête des messages légitimes (faux positifs). De plus, toute technique a des coûts directs et indirects et qui doivent être pris en compte lors du choix.

Les critères techniques des solutions de filtrage (taux de faux positifs, taux de faux négatifs, ...).

Panorama très rapide de toutes les techniques de lutte.

L'authentification "faible" du courrier électronique avec SPF (Sender Policy Framwork). Principe et déploiement. Bien choisir son enregistrement SPF, le tester.

Le "greylisting". Principe et déploiement.

Les filtres à heuristiques comme SpamAssassin. Principe, déploiement et réglages.

Les filtres bayésiens. Principe et déploiement. Comme on envisage ici le cas d'un serveur collectif, pas d'un logiciel de lecture individuel, on mettra l'accent sur les procédures de mise à jour du vocabulaire.

Conclusion rapide : la posture à adopter pour les prochaines années. Techniquement, adaptation permanente et politiquement /juridiquement / socialement, quelles perspectives ?

Une référence qui colle à l'esprit de ce tutoriel :

http://www.freesoftwaremagazine.com/free_issues/issue_02/focus_spam_postfix/