

DES CHIFFRES ET DES LETTRES



MARTIN.OUWELAND@epfl.ch, DOMAINE IT

Le but de cet article est de présenter quelques statistiques concernant l'utilisation du courrier électronique à l'EPFL, pour l'édification et le divertissement de tous.

ADRESSES E-MAIL

Commençons par les adresses e-mail @epfl.ch: au moment où je rédige ces lignes, le 5 avril 2005, il y a 10484 adresses personnelles au format *prénom.nom@epfl.ch* en usage (4966 étudiants, 1392 assistants/doctorants, 2978 employés/affiliés et 1148 ex-membres qui verront leur adresse supprimée sous peu), 583 adresses abstraites couvrant des fonctions, des projets, etc. (p.ex. *fi@epfl.ch* ou *support-it@epfl.ch*), et enfin 7125 adresses de listes de distribution telles que *personnel.epfl@epfl.ch*, automatiquement générées à partir des unités, classes et rôles figurant dans les données de l'accréditation (voir <http://dinfo.epfl.ch/cgi-bin/listes.pl>), pour un total de 18192 adresses.

MESSAGES

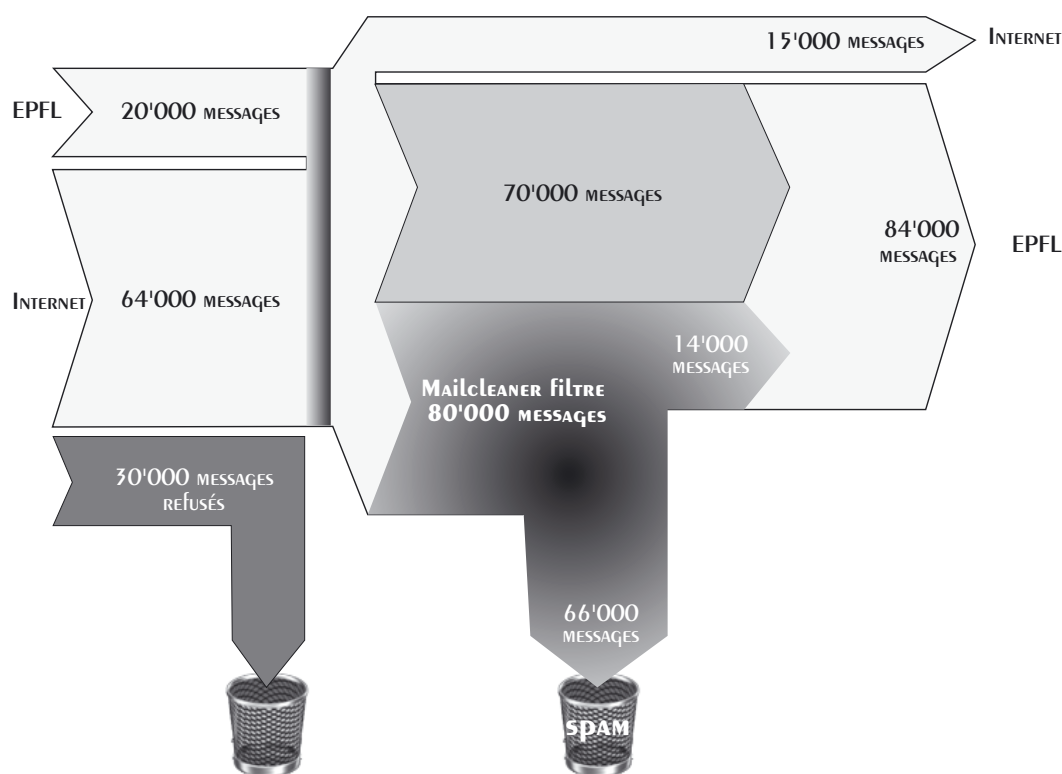
Passons au nombre de messages traités par les serveurs de messageries de l'EPFL. La figure de la page suivante résume la situation, indiquant diverses moyennes par jour sur tout le mois de mars 2005. Ainsi, sur la partie gauche on voit que nos serveurs reçoivent et acceptent de traiter chaque jour 84'000 messages, (64'000 externes + 20'000 internes) mais en refusent 30'000 pour diverses raisons, p.ex. parce que le destinataire n'existe pas ou parce qu'ils contiennent une annexe jugée dangereuse (environ 2500 cas, voir <http://mailwww.epfl.ch/danger.html>).

Sur la partie à droite figurent les messages envoyés par les serveurs, avec un effet de multiplication (les 165'000 messages en sortie sont à peu près deux fois plus nombreux que les messages entrant) qui n'est que partiellement représenté avec les échelles choisies. L'explication en est simplement qu'un message peut avoir plusieurs destina-

taires, un cas extrême étant représenté par un envoi à *etudiants.epfl@epfl.ch* qui générera 4783 message en sortie.

Sur les 150'000 messages destinés à l'EPFL, 80'000 sont d'abord soumis au filtre anti-spam MailCleaner (voir FI,

adresses personnelles:	10484
■ étudiants/doctorants/postgrade	4966
■ assistants	1392
■ employés	2978
■ "en partance"	1148
adresses de fonctions (aliases)	583
adresses listes (834 unités, 57 fonctions)	7125
adresses "@epfl.ch"	18192



STATISTIQUES DU COURRIER ÉLECTRONIQUE (MOYENNES JOURNALIÈRES SUR LE MOIS DE MARS 2005)

http://dit.epfl.ch/publications-spi/article.php3?id_article=182) qui en retient 66'000 comme spam (les abonnés à ce service témoigneront que son taux d'erreur ne dépasse pas quelques pour cent). Les autres 70'000 messages ne sont pas filtrés soit parce qu'ils proviennent de l'EPFL soit parce qu'ils sont destinés à des personnes qui ne sont pas abonnées à MailCleaner. Avec les 14'000 messages que MailCleaner laisse passer, nous obtenons 84'000 messages parvenant finalement dans les boîtes aux lettres des membres de l'EPFL. On peut donc dire que le courrier électronique est à l'EPFL plutôt un outil de travail interne, bien que les 15'000 messages envoyés depuis l'EPFL vers l'extérieur ne soient pas négligeables.

Alors, combien de spam ?

Il n'est pas surprenant de constater qu'à cause du spam il y a proportionnellement plus de messages destinés aux abonnés à MailCleaner: alors qu'ils sont un peu moins de 20 % (1980 abonnés sur 10484 membres de l'EPFL), les 80'000 messages qui leur sont destinés représentent plus de la moitié du trafic à destination interne. La proportion de spam détecté après la multiplication des messages vers leurs divers destinataires, soit 44 % (66'000 messages sur 150'000), est une estimation assez fiable du spam à ce stade-là (elle en représente même une borne inférieure puisque certaines personnes qui ne sont pas abonnées à MailCleaner reçoivent sûrement un peu de spam), mais il est par contre beaucoup plus difficile de répondre autrement que par de grossières estimations à la question qui est sur toutes les lèvres: *combien de spams l'EPFL reçoit-elle chaque jour ? si on entend par là quelle proportion des 64'000 messages externes entrant est-elle du spam ?* Il est probable qu'en admettant simplement la proportionnalité, soit 28'000 spams quotidiens (44 % de 64'000),

on en surestime le nombre réel parce que les spams ont en général de nombreux destinataires et je me contenterai de cette fourchette: l'EPFL reçoit entre 10'000 et 30'000 spams par jour (donc entre 16 % et 47 % du flux entrant).

Taille des messages

Pour conclure, quelques chiffres sur la quantité de données que représentent tous ces messages, toujours en moyenne journalière sur le mois de mars 2005. Tout d'abord en sortie, les 84'000 messages à destination interne représentent 5,3 Gigaoctets, et les messages de l'EPFL vers l'extérieur 2 Gigaoctets. Si on sait que les 80'000 messages soumis à MailCleaner ne font *que* 1,3 Gigaoctets, on voit que le spam a au moins cette vertu d'être de taille plus modeste que le reste des messages. Ceci explique aussi pourquoi, en entrée, les messages provenant de l'EPFL représentent au total 2,2 Gigaoctets et les messages externes 2,5 Gigaoctets, alors que ces derniers sont plus de trois fois plus nombreux: les messages internes sont donc en moyenne près de trois fois plus gros que les messages externes. Un autre facteur contribuant à cette différence est le type d'utilisation de la messagerie électronique: en interne, on hésitera moins à se communiquer des versions provisoires de gros documents en ébauche.

Ces chiffres, de l'ordre du Gigaoctet, sont à comparer avec le trafic journalier global de l'EPFL, qui se compte en centaines de Gigaoctets (voir <http://mathe.epfl.ch/>). On peut donc mentionner, si on considère sa part du trafic global, que le spam n'entraîne sans doute pas comme on l'entend parfois dire des coûts importants, qui se situent plutôt dans son filtrage, automatique (MailCleaner) ou *manuel* (le temps consacré par chacun d'entre nous pour s'en débarrasser). ■