

RFC 4952 : Overview and Framework for Internationalized Email

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 25 juillet 2007. Dernière mise à jour le 7 septembre 2008

Date de publication du RFC : Juillet 2007

<https://www.bortzmeyer.org/4952.html>

Parmi les protocoles utilisés sur Internet, presque tous sont internationalisés depuis longtemps. Le dernier gros manque concernait les adresses de courrier électronique, obligées de s'en tenir à `stephane@internet-en-coopération.fr` alors qu'on voudrait pouvoir écrire à `stéphane@internet-en-coopération.fr`. Désormais, nous avons une solution, "*Internationalized Email Addresses*" <<https://www.bortzmeyer.org/courrier-entierement-internationalise.html>>. Expérimentale à l'époque de la parution de ce RFC, cette solution est devenue une norme complète par la suite avec le RFC 6530¹, qui a remplacé ce RFC.

Bien sûr, pour l'exemple ci-dessus, le gain n'est pas évident. Mais il l'est beaucoup plus si vous voulez écrire votre adresse avec l'écriture arabe ou chinoise.

Le **contenu** des courriers était internationalisé depuis longtemps, au moins depuis MIME (RFC 1341 à l'origine, en juin 1992). Mais les adresses ne l'étaient pas encore. Notre RFC décrit le cadre général de l'internationalisation des adresses de courrier.

Il y a deux parties à l'adresse (décrites dans le RFC 2822) : la partie locale, à gauche du @ et le nom de domaine à droite. Si ce nom de domaine, depuis la sortie du RFC 3490, en mars 2003, peut s'écrire en Unicode (norme IDN), le courrier exigeait toujours un nom de domaine en ASCII donc un message de non-délivrance, par exemple, allait afficher la forme Punycode du nom. Ce n'était pas une vraie internationalisation.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc6530.txt>

Le cadre d'internationalisation du courrier couvre plusieurs aspects, couverts chacun par un RFC <<https://www.bortzmeyer.org/courrier-entierement-internationalise.html>>. Ce premier RFC expose l'architecture générale.

SMTP est modifié pour permettre une nouvelle extension, indiquant le support du courrier internationalisé (RFC 5336).

Le format des messages (RFC 2822) est modifié pour permettre d'inclure de l'Unicode, encodé en UTF-8, directement dans les en-têtes, comme le champ `To`: (RFC 5335).

Un mécanisme de **repli** ("*downgrading*") était prévu pour le cas où un émetteur internationalisé rencontrerait un récepteur qui ne l'est pas, nécessitant une modification des adresses en ASCII (RFC 5504). En effet, contrairement à HTTP, le courrier ne fonctionne pas de bout en bout et le message est relayé par plusieurs serveurs, qu'on ne connaît pas forcément a priori. Ce mécanisme s'est révélé bien compliqué pour pas grand'chose et a finalement été abandonné avec le RFC 6530.

De même, les protocoles POP et IMAP doivent être modifiés pour que le courrier ainsi internationalisé puisse être récupéré... et lu.

Le repli est certainement la partie la plus complexe et fera l'objet d'un RFC propre : modifier un message est toujours une opération délicate, notamment en présence de signatures cryptographiques, que le repli ne doit pas invalider !

Notons (section 6.3 du RFC) que les adresses de courrier sont utilisés en de nombreux endroits, par exemple comme identificateurs sur certains sites Web de commerce électronique, comme Amazon et que ces utilisateurs vont également devoir s'adapter. Quant on voit le nombre de sites, supposés professionnels, qui continuent à interdire des adresses de courrier légales <<https://www.bortzmeyer.org/arreter-d-interdire-des-adresses-legales.html>>, on mesure l'ampleur du travail <http://www.circleid.com/posts/20091120_idn_and_email_the_harsh_reality/>. La sortie en février 2012 de la norme RFC 6530, qui remplace notre RFC, accélérera-t-elle les choses ?