

RFC 6783 : Mailing Lists and non-ASCII Addresses

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 novembre 2012

Date de publication du RFC : Novembre 2012

<https://www.bortzmeyer.org/6783.html>

Depuis la sortie du RFC 6530¹ et de ses petits camarades, il existe désormais une solution normalisée pour avoir du courrier électronique complètement internationalisé <<https://www.bortzmeyer.org/courrier-entierement-internationalise.html>>, y compris les adresses (arôme@café.fr...). Mais comment est-ce que cela s'applique aux listes de diffusion? Peut-on avoir des listes où certains membres accepteraient les nouvelles adresses et d'autres non? Ce RFC, successeur du RFC 5983, se penche sur la question. Pour l'instant, la réponse est pessimiste : il n'y a pas de moyen pratique de faire coexister anciens et nouveaux sur la même liste.

Une liste de diffusion prend un message et le redistribue à N destinataires. Certains de ces N destinataires seront derrière un serveur de messagerie récent, qui gère les adresses internationalisés. Une partie de ceux-ci, en outre, auront des adresses internationalisées. Même si le destinataire, lui, n'a pas une adresse internationalisée, si l'émetteur en a une, ses messages ne pourront pas être remis aux gens qui ont les vieux logiciels.

Il y a donc un problème technique, le fait que des adresses internationalisées soient refusées par certains. Et il y a un problème humain : en communication bilatérale, après quelques essais, on peut toujours trouver une solution technique qui permette la communication. Sur une liste, ce n'est plus possible, l'émetteur ne connaît pas les destinataires et réciproquement. Il n'y a pas d'ajustement possible.

Autre problème : idéalement, la liste ne modifie pas le message, à part par l'ajout de certains en-têtes spécifiques aux listes (comme `List-Unsubscribe:` ou `List-Id:`, normalisés dans le RFC 2369 et RFC 2919). Toutefois, un certain nombre de listes vont plus loin et, bien que cela soit une source d'ennuis connue, modifient d'autres en-têtes comme `Reply-To:` voire, pire, `From:`.

Pour synthétiser, il y a trois services différents qui peuvent marcher ou pas :

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

- Avoir une liste dont l'adresse est internationalisée, par exemple `évolution-langue@académie-française`
- Avoir une liste qui accepte l'abonnement de personnes ayant des adresses internationalisées, par exemple `étienne@massé.fr`,
- Avoir une liste qui accepte les messages émis depuis des adresses internationalisées.

Ces trois capacités sont relativement indépendantes (la deuxième et la troisième impliquent des logiciels différents donc l'une peut marcher et pas l'autre).

Le plus simple pour une liste qui veut gérer parfaitement l'internationalisation serait de n'avoir que des abonnés qui eux-même ont la gestion complète des adresses en Unicode (option `SMTPUTF8` de SMTP, cf. RFC 6531). Pour cela, le mieux est de tester lors de l'inscription, et de refuser l'abonnement de ceux qui ont le vieux logiciel. Pour ce test, si le candidat à l'abonnement s'abonne avec une adresse Unicode, c'est facile : s'il reçoit le message de confirmation, c'est bon. Et s'il s'abonne avec une adresse ASCII ? Alors, le mieux est de tester en tentant de leur envoyer le message « répondez pour confirmer » **depuis** une adresse Unicode.

Mais pour la majorité des listes, qui voudront accepter des abonnés `SMTPUTF8` et d'autres qui ne le sont pas ? (Fin 2012, très peu de serveurs de messagerie, à part en Extrême-Orient, gèrent `SMTPUTF8`.) Il n'existe pas (ou plus, cf. RFC 5504) de mécanisme de repli automatique d'un message aux adresses Unicode vers un message traditionnel. C'est donc au MLM ("*Mail List Manager*", le logiciel de gestion de la liste, par exemple Mailman) de faire ce repli.

D'abord, le logiciel de gestion de la liste doit déterminer quels abonnés sont `SMTPUTF8` et lesquels ne le sont pas. Il y a plusieurs méthodes :

- Décider que les abonnés avec une adresse Unicode peuvent recevoir du courrier entièrement internationalisé et que les abonnés avec une adresse ASCII ont un logiciel traditionnel (en pratique, cela privera certains de ces abonnés ASCII des bénéfices du courrier internationalisé),
- Tester, en envoyant un message depuis une adresse Unicode et en demandant confirmation,
- Demander à l'utilisateur de se classer lui-même comme, aujourd'hui, on lui demande s'il veut recevoir du texte ou de l'HTML.

Ensuite, il faut décider, pour chaque message entrant vers la liste, s'il est internationalisé ou pas. Cela peut se faire en considérant que, si l'option `SMTPUTF8` a été utilisée dans la session, le message sera marqué comme « Unicode ». Ou bien on peut examiner les en-têtes du message, à la recherche d'adresses en Unicode.

Les logiciels de gestion de liste réécrivent en général l'émetteur dans l'enveloppe SMTP du message. Ils mettent leur propre adresse (par exemple `ima-bounces@ietf.org` pour la liste du groupe de travail EAI, ex-IMA) ou même ils mettent une adresse d'origine différente pour chaque abonné, pour mieux gérer les messages de non-remise (technique VERP). Dans les deux cas, l'adresse d'émetteur, qui recevra les messages de non-remise, contient en général le nom de la liste. Dans ce cas, elle sera une adresse Unicode et les serveurs ne connaissant pas le courrier complètement internationalisé ne pourront donc pas envoyer ces avis de non-remise ("*bounces*").

La section 3 contient les recommandations concrètes. D'abord, pour les en-têtes ajoutés comme `List-Id:`, la recommandation est de les mettre sous une forme purement ASCII, pour limiter les risques, même si cet en-tête prend comme valeur un URI (comme le fait `List-Unsubscribe:`) et que cet URI a un plan ("*scheme*") qui permet l'Unicode. Il faut dans ce cas utiliser l'« encodage pour-cent ». Donc, pour une liste `pause-café@bavardage.fr`, on aura :

```
List-Id: Liste de discussion en Unicode <pause-café%3%A9@bavardage.fr>
List-Unsubscribe: <mailto:pause-café%3%A9-requ%3%AAt@bavardage.fr?subject=unsubscribe>
List-Archive: <http://www.example.net/listes/pause-café%3%A9/archives>
```

`List-Id:`, notamment, est très souvent utilisé dans des systèmes de traitement automatique du courrier reçu comme les filtres Sieve ou procmail. Ces filtres ne réagissent pas forcément correctement à de l'Unicode (par exemple, ils peuvent échouer à comparer la forme UTF-8 et une forme encodée différemment). D'où la recommandation, vue l'importance de cet en-tête pour le bon fonctionnement de la liste, qu'on se limite à l'ASCII.

Dans certains cas, cela ne donnera pas un résultat lisible (par exemple le `List-Id:` d'une liste en chinois encodé en ASCII ne va pas être tellement utile aux lecteurs...) Mais cet inconvénient semble, **aujourd'hui**, moins grave que le risque de faire arriver de l'Unicode à des gens qui ne s'y attendent pas. Une meilleure solution devra attendre une nouvelle version des RFC 2369 et RFC 2919.

Le RFC n'a pas de recommandation pour le cas où un auteur d'un message a une adresse Unicode et où le gestionnaire de listes essaie de diffuser le message à un abonné n'ayant que des logiciels anciens, ne gérant pas SMTPUTF8. Il n'y a pas de bonne solution technique, juste des petits arrangements : réserver la liste aux gens pouvant recevoir du courrier internationalisé (cf. plus haut la discussion sur de telles listes), proposer aux abonnés d'enregistrer une adresse ASCII et remplacer le `From:` par cette adresse avant de transmettre ou enfin remplacer le `From:` par une adresse de la liste. Reste à programmer cela...

Je ne connais pas encore de « grand » gestionnaire de liste de diffusion qui ait des plans précis pour la mise en œuvre des adresses de courrier internationalisées. Sympa, par exemple, en est toujours au stade de la réflexion sur ce point.

Notre RFC ne contient pas de liste des changements depuis le RFC 5983. Il y a eu réécriture complète puisque le repli automatique et les adresses de secours (`alt-address`) ont été retirées de la norme finale du courrier internationalisé <<https://www.bortzmeyer.org/courrier-entierement-internationalise.html>>.