

# RFC 5983 : Mailing Lists and Internationalized Email Addresses

Stéphane Bortzmeyer  
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 5 octobre 2010

Date de publication du RFC : Octobre 2010

<https://www.bortzmeyer.org/5983.html>

---

L'arrivée des normes sur le courrier électronique entièrement internationalisé <<https://www.bortzmeyer.org/courrier-entierement-internationalise.html>> a suscité beaucoup d'intérêt et d'expériences. Un des points qui avait été laissé dans l'ombre au début est le cas de listes de diffusion. Ce RFC comblait le manque, explique le problème et donnait des recommandations. Il a depuis été remplacé par le RFC 6783<sup>1</sup>.

Une liste de diffusion est un mécanisme par lequel un message, bien qu'envoyé à une seule adresse, est distribué à plusieurs destinataires. L'agent logiciel qui reçoit le message et le redistribue modifie l'adresse de retour indiquée dans l'enveloppe, de façon à ce que ce soit lui, et non pas l'expéditeur originel, qui reçoive les éventuels avis de non-remise. Dans le cas d'une « vraie » liste de diffusion (par exemple, pas un simple alias Postfix), le message ainsi modifié est réinjecté par une soumission normale.

Le message subit d'autres changements, comme par exemple l'ajout d'en-têtes spécialisés dans la gestion de listes (RFC 2369 et RFC 2919) comme `List-Id:`, qui contient une adresse, celle de la liste, ou `List-Unsubscribe:`, qui contient un URI, typiquement de plan `mailto:`.

Idéalement, cela s'arrêterait là. Mais certains MLM ("*Mailing List Manager*", logiciels de gestion de listes) vont plus loin et modifient le message, par exemple pour ajouter un en-tête `Reply-To:`, ou pour ajouter le nom de la liste dans le sujet, ou bien pour ajouter du texte à la fin du message. Pire, certains modifient des en-têtes comme `From:`.

Il y a trois endroits où l'arrivée des adresses de courrier en Unicode a un impact sur les listes de diffusion :

---

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc6783.txt>

- Le transport des messages,
- Les en-têtes mis dans les messages,
- Les politiques de gestion de la liste.

Les questions soulevées par ces nouvelles adresses peuvent être purement techniques (un MLM pourrait par exemple refuser les adresses non-ASCII) ou plus sémantiques, liées aux utilisateurs (si la liste elle-même a une adresse ASCII, et qu'un de ses membres a une adresse Unicode, l'expéditeur pourrait être surpris si cette adresse Unicode venait à sa connaissance, même si toute la partie technique a bien fonctionné).

Certains de ces problèmes peuvent aussi arriver en communication directe, sans l'intermédiaire de la liste de diffusion. Mais ils se règlent plus difficilement dans le cas d'une liste car l'expéditeur d'un message, par exemple, ne peut évidemment pas s'adapter à chaque membre de la liste.

Bref, après les trois endroits mentionnés plus haut, il faut se poser les trois questions concrètes suivantes :

- La liste peut-elle avoir elle-même une adresse en Unicode ? Par exemple `évolution-langue@académie-fra`
- La liste accepte t-elle des abonnés « Unicode » comme `étienne@massé.fr` ?
- La liste accepte t-elle les messages délivrés entièrement en Unicode (RFC 5336) ?

Avant d'étudier ces trois questions plus en détail, la section 3 rappelle un scénario particulièrement délicat : si le message originel nécessite UTF8SMTP, par exemple parce que l'adresse de son expéditeur comprend de l'Unicode, et que certains des destinataires de la liste ont des serveurs SMTP sans cette option, le message devra se rabattre sur de l'ASCII seul, suivant le RFC 5504, sans que le serveur SMTP de l'expéditeur original n'ait pu se rendre compte de quelque chose. Ce problème devrait disparaître avec le temps, le repli sur l'ASCII seul étant largement considéré aujourd'hui comme une mauvaise idée, qui a été exclue de la normalisation définitive du courrier électronique avec adresses Unicode.

Après ce petit détour, revenons aux trois questions. En théorie, le système de gestion de la liste peut répondre OUI à n'importe laquelle des questions indépendamment des autres. En pratique, toutefois, on ne voit pas bien l'intérêt d'avoir une adresse en Unicode si le serveur SMTP frontal ne gère pas UTFSMTP (et réciproquement, d'ailleurs). La section 4 discute ces trois services séparément mais, logiquement, les MLM les fourniront tous ou aucun.

Donc, premier service, que la liste ait une adresse de soumission en Unicode, comme `pause-café@bavardage.f`. Le RFC recommande que la liste ait aussi une adresse alternative en ASCII (`alt-address` du RFC 5336, section 3.4), mettons `pause-cafe@bavardage.fr`. À noter que la liste a aussi une autre adresse, utilisée comme expéditeur dans l'enveloppe des messages sortants, et qui reçoit donc les messages d'erreur. Pour ce cas, la liste peut très bien se contenter d'une adresse en ASCII seul et, si cette adresse est en Unicode, le RFC impose (c'est la seule obligation nouvelle de ce RFC) qu'il y ait une `alt-address` en ASCII, pour garantir la bonne réception des messages d'erreur.

Et pour les adresses des abonnés ? Il n'y a pas de recommandation stricte, juste le rappel qu'une liste peut, si elle le souhaite, imposer que les abonnés fournissent une adresse alternative en ASCII.

Enfin, pour UTF8SMTP, le gestionnaire de la liste doit s'assurer que tous les serveurs sur le chemin gèrent bien cette option.

Le courrier entièrement internationalisé pose un autre problème, c'est celui des en-têtes `List*` : tels que normalisés par les RFC 2369 et RFC 2919. Pour reprendre l'exemple de `pause-café@bavardage.fr`, à quoi doivent ressembler, par exemple, ces en-têtes :

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

Pour les deux derniers en-têtes, qui contiennent des URI, tout dépend du plan (ici, `mailto:` et `http:`). Même si ce dernier autorise les IRI (ce qui n'est pas le cas de `mailto:` aujourd'hui), le RFC demande que ce soit leur forme encodée en URI qui soit utilisée. (Notons que les IRI sont actuellement en cours de réforme à l'IETF, et qu'un des points en discussion est l'encodage du nom de domaine dans l'IRI, problème qui ne se pose pas dans cet exemple.) On aura donc, par exemple :

```
List-Unsubscribe: <mailto:pause-caf%C3%A9-requ%C3%Aate@bavardage.fr?subject=unsubscribe>  
List-Archive: <http://www.example.net/listes/pause-caf%C3%A9/archives>
```

Ces en-têtes `List*` : étant essentiels pour le bon fonctionnement des listes, le RFC insiste sur le fait qu'avoir également une alternative en ASCII pur est très souhaitable. Les deux adresses (Unicode et ASCII) sont indiquées dans deux URI différents, séparés par des virgules (RFC 2369, section 2).

L'en-tête `List-Id` (qui identifie une liste de manière unique), lui, ne prend pas un URI comme valeur et cette discussion ne s'applique pas à lui. En revanche, il est très souvent utilisé dans des systèmes de traitement automatique du courrier reçu comme les filtres Sieve. 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). Le RFC ne fournit toutefois pas de solution à ce problème.

Je ne connais pas encore de 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 au stade de la réflexion sur ce point. Le RFC 6783, version actuelle, traite le problème très différemment, notamment en raison de la suppression du repli automatique (d'un message internationalisé vers un message ASCII).