

Groupe de travail Réseau
Request for Comments : 5064
 Catégorie : Sur la voie de la normalisation

M. Dürst, Aoyama Gakuin University
 décembre 2007
 Traduction Claude Brière de L'Isle

Champ d'en-tête de message Archived-At

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet sur la voie de la normalisation pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition en cours des "Protocoles officiels de l'Internet" (STD 1) pour voir l'état de normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Résumé

Le présent mémoire définit un nouveau champ d'en-tête de messagerie électronique, Archived-At., pour fournir une liaison directe à la forme archivée d'un message électronique individuel.

Table des Matières

1. Introduction.....	1
2. Définition de champ d'en-tête.....	2
2.1 Syntaxe.....	2
2.2 Champs d'en-tête Archived-At multiples.....	2
2.3 Interaction avec la fragmentation et le réassemblage de message.....	2
2.4 Extension de syntaxe pour les en-têtes de message internationalisés.....	2
2.5 Champ d'en-tête X-Archived-At.....	3
3. Considérations de mise en œuvre et d'usage.....	3
3.1 Formats de message archivé.....	3
3.2 Considérations de mise en œuvre.....	3
3.3 Considérations d'utilisation.....	3
4. Considérations sur la sécurité.....	4
5. Considérations relatives à l'IANA.....	4
5.1 Enregistrement du champ d'en-tête Archive-At.....	4
5.2 Enregistrement du champ d'en-tête X-Archived-At.....	5
6. Remerciements.....	5
7. Références.....	5
7.1 Références normatives.....	5
7.1 Références pour information.....	5
Adresse de l'auteur.....	6
Déclaration complète de droits de reproduction.....	6

1. Introduction

La [RFC2369] définit un certain nombre de champs d'en-tête qui peuvent être ajoutés à des messages Internet comme ceux envoyés par les listes de distribution de messagerie électronique ou dans les nouvelles du réseau (netnews) de la [RFC1036]. L'un d'eux est le champ d'en-tête List-Archive qui décrit comment accéder aux archives pour la liste. Cela permet l'accès aux archives comme un tout, mais pas à un message individuel.

Il y a souvent un besoin ou désir de se référer à la forme archivée d'un seul message. Pour des scénarios d'usage plus détaillés, voir le paragraphe 3.3. Le présent mémoire définit un nouvel en-tête, Archived-At, pour se référer à un seul message dans une localisation archivée. Cela fournit un accès rapide à la localisation d'un message d'une liste de diffusion dans l'archive de la liste. Il peut aussi être utilisé indépendamment des listes de diffusion, par exemple en connexion avec l'exigence légale d'archiver certains messages.

Les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDE", "PEUT", et "FACULTATIF" en majuscules dans ce document sont à interpréter comme décrit dans le BCP 14, [RFC2119].

2. Définition de champ d'en-tête

2.1 Syntaxe

Pour le champ d'en-tête Archived-At, le nom du champ est "Archived-At". Le corps du champ consiste en un URI [RFC3986] enclos entre des crochets angulaires ("<", ">"). L'URI PEUT contenir une espace blanche de pliage (FWS, *Folding WhiteSpace*) [RFC2822], qui est ignorée. Les agents de transfert de messagerie (MTA, *Mail Transfer Agent*) NE DOIVENT PAS insérer d'espace au sein des crochets angulaires, mais les applications de client DEVRAIENT ignorer toute espace qui pourrait avoir été insérée par des MTA au mauvais comportement. L'URI pointe sur une version archivée du message. Voir les détails au paragraphe 3.1.

Ce champ d'en-tête est sujet aux restrictions de codage et de caractères pour les en-têtes de messagerie, comme décrit dans la [RFC2822].

Plus formellement, le champ d'en-tête est défini comme suit en BNF augmenté (ABNF) conformément à la [RFC4234] :

```
archived-at = "Archived-At:" [FWS] "<" folded-URI ">" CRLF
folded-URI = <URI, mais libre insertion de FWS permise>
```

où URI est défini dans la [RFC3986], et CRLF et FWS sont définis dans la [RFC2822].

Pour convertir un URI replié (*folded-URI*) en URI, on applique d'abord les règles standard de dépliage de la [RFC2822] (en remplaçant le FWS par un seul SP) et ensuite en supprimant tous les caractères restants non codés en SP.

Cette syntaxe est gardée simple en ce que seulement un URI par champ d'en-tête est permis. À cet égard, la syntaxe est différente de celle de la [RFC2369]. Aussi, les commentaires ne sont pas admis.

2.2 Champs d'en-tête Archived-At multiples

Chaque champ d'en-tête Archived-At contient seulement un URI. Si il est désiré de lister plusieurs URI où une copie archivée du message peut être trouvée, un champ Archived-At séparé par URI est requis. Plusieurs champs d'en-tête Archived-At avec le même URI DEVRAIT être évités. Un champ d'en-tête Archived-At DEVRAIT seulement être créé si le message est réellement disponible à l'URI donné dans le champ d'en-tête.

Si un message est transmis à partir d'une liste à une sous liste et si les deux listes prennent en charge l'ajout de champ d'en-tête Archived-At, alors la sous liste DEVRAIT ajouter un nouveau champ d'en-tête Archived-At sans supprimer le ou les existants, sauf si le champ d'en-tête est exactement le même qu'un déjà existant, auquel cas le nouveau champ d'en-tête NE DEVRAIT PAS être ajouté.

2.3 Interaction avec la fragmentation et le réassemblage de message

La [RFC2046] permet la fragmentation et le réassemblages des messages. Les champs d'en-tête Archived-At sont à traiter de la même façon que les champs d'en-tête Comments, c'est-à-dire, copiés au premier en-tête de fragment de message à la fragmentation et repris de là dans l'en-tête du message réassemblé.

Ce traitement a été choisi pour la compatibilité avec l'infrastructure existante. Il signifie que les champs d'en-tête Archived-At dans le premier fragment de message PEUVENT se référer à une version archivée du message entier, non fragmenté. Pour éviter toute confusion, les en-têtes Archived-At NE DEVRAIENT PAS être ajoutés aux fragments de messages.

2.4 Extension de syntaxe pour les en-têtes de message internationalisés

Il y a des efforts en cours pour permettre du texte non ASCII directement dans les corps de champ d'en-tête de messages. Dans ce contexte, l'URI non terminal dans la syntaxe définie au paragraphe 2.1 est à remplacer par un identifiant de ressource internationalisée (IRI, *Internationalized Resource Identifier*) comme défini dans la [RFC3987]. Les spécificités du codage de codets réel de l'IRI va suivre les règles du codage direct général de texte non ASCII. Pour la conversion entre les IRI et les URI, les procédures définies dans la [RFC3987] sont à appliquer.

2.5 Champ d'en-tête X-Archived-At

Pour la rétro-compatibilité, le présent document décrit aussi le champ d'en-tête X-Archived-At, un précurseur du champ d'en-tête Archived-At. Le champ d'en-tête X-Archived-At PEUT aussi être analysé, mais NE DEVRAIT PAS être généré.

La syntaxe du champ d'en-tête X-Archived-At en ABNF conformément à la [RFC4234] (qui définit aussi SP) est la suivante :

```
obs-archived-at = "X-Archived-At:" SP URI CRLF
```

Le champ d'en-tête X-Archived-At ne permet pas d'espace à l'intérieur de l'URI.

3. Considérations de mise en œuvre et d'usage

3.1 Formats de message archivé

Il n'y a pas de restriction sur le format utilisé pour servir le message archivé à partir de l'URI dans un champ d'en-tête Archived-At. On suppose que dans de nombreux cas, le message archivé va être servi comme (X)HTML, comme pur texte, ou dans sa forme originale comme message/rfc822 [RFC2046]. Certaines formes d'URI peuvent impliquer le format dans lequel le message archivé est servi, bien qu'on ne devrait pas trop y compter.

Si le protocole utilisé pour restituer le message permet une négociation de contenu, il est alors aussi possible de servir le message archivé dans plusieurs formats différents. Par exemple, un URI HTTP dans un en-tête Archived-At peut rendre possible de servir le message archivé comme text/html pour la consommation humaine dans un navigateur et comme message/rfc822 à l'usage de l'agent d'utilisateur de messagerie (MUA, *mail user agent*) sans perte d'information.

3.2 Considérations de mise en œuvre

Les expanseurs de liste de diffusion et les archives de messagerie électronique sont souvent des éléments de logiciel séparés. Il peut donc être difficile de créer un champ d'en-tête Archived-At dans le logiciel d'expanseur de liste de diffusion.

Une façon de régler cette difficulté est de faire que le logiciel d'expanseur de liste de diffusion génère un URI sans ambiguïté, par exemple, un URI fondé sur l'identifiant de message du message entrant, et de régler le système d'archivage de façon à ce qu'il redirige les demandes à de tels URI sur les messages réels. Si la messagerie électronique ne contient pas d'identifiant de message, un identifiant unique peut être généré.

Un tel système a été mis en œuvre est en usage au W3C. Par exemple, l'URI "<http://www.w3.org/mid/015U00G08DFGCR@mailsj-v1.corp.adobe.com>", contenant la partie significative de l'identifiant de message "<015U00G08DFGCR@mailsj-v1.corp.adobe.com>", est redirigé sur l'URI de ce message dans l'archive de liste de diffusion du W3C à <http://lists.w3.org/Archives/Public/uri/2004Oct/0017.html>.

Le code de source pour cette mise en œuvre est disponible à <http://dev.w3.org/cvsweb/search/>, en particulier <http://dev.w3.org/cvsweb/search/cgi/mid.pl> et <http://dev.w3.org/cvsweb/search/bin/msgid-db.pl>. Ces localisations peuvent être sujettes à des changements.

Quand on utilise l'identifiant de message pour créer une adresse pour le message archivé, il faut faire attention aux caractères d'échappement dans l'identifiant de message qui ne sont pas permis dans l'URI, ou les retirer, comme on le fait ci-dessus pour les délimiteurs "<" et ">".

Les mises en œuvre comme celle décrite ci-dessus peuvent introduire un problème de sécurité. Quelqu'un pourrait délibérément réutiliser un identifiant de message pour casser le lien à un message. Cela peut être traité en confrontant les identifiants de message entrants à ceux des messages déjà dans l'archive et en éliminant les dupliqués entrants, en vérifiant le contenu des dupliqués entrants et en les éliminant si ils sont significativement différents de celui du premier message, en offrant plusieurs choix dans la réponse à l'URI, ou en utilisant un mécanisme d'authentification sur les messages entrants.

3.3 Considérations d'utilisation

Il peut à première vue sembler étrange d'avoir un pointeur sur une forme archivée d'un message dans un champ d'en-tête du

même message. Après tout, si on a le message, pourquoi aurait on besoin d'un pointeur sur lui ? Il se trouve que de tels pointeurs peuvent être extrêmement utiles. Ce paragraphe décrit certains scénarios de leur utilisation.

Un utilisateur peut vouloir se référer aux messages dans un contexte qui n'est pas de message, comme sur une page de la Toile, dans un message instantané, ou dans une conversation téléphonique. Dans ce cas, l'utilisateur peut extraire l'URI du champ d'en-tête Archived-At, évitant la recherche du message correct dans l'archive.

Un utilisateur peut vouloir se référer aux autres messages dans un contexte de message. Se référer à un seul message est souvent fait en répondant à ce message. Cependant, quand on se réfère à plus d'un message, fournir des pointeurs sur les messages archivés est une pratique largement répandue. Le champ d'en-tête Archived-At rend plus facile de fournir ces pointeurs.

Un utilisateur peut vouloir trouver des messages relatifs à un message en cours. L'utilisateur peut n'avoir pas reçu les messages concernés, et a donc besoin d'utiliser une archive. L'utilisateur peut aussi préférer trouver les messages concernés dans l'archive plutôt que dans son MUA, parce que les messages dans les archives peuvent être reliés de façons non fournies par le MUA. Le champ d'en-tête Archived-At fournit un lien au point de départ dans l'archive à partir duquel trouver les messages concernés.

Noter que dans les scénarios d'usage ci-dessus, c'est surtout le lecteur humain, plutôt que le logiciel de client de messagerie électronique, qui utilise l'URI dans l'en-tête Archived-At. Cependant, cela n'exclut pas l'utilisation de l'URI dans l'en-tête Archived-At par le client de messagerie électronique ou autre logiciel si un tel usage est utile.

4. Considérations sur la sécurité

Il y a de nombreux problèmes potentiels de sécurité quand on active et déréférence un URI. Pour plus de détails, incluant certaines contre-mesures, voir la [RFC3986]. Dans le contexte de la présente proposition, ce qui suit est particulièrement pertinent : un intrus peut obtenir l'accès à la transmission du message et être capable d'insérer un URI pointant sur un contenu malveillant. Cela peut être traité en utilisant une façon sécurisée de transmettre le message. Aussi, quelqu'un peut être capable de construire un message qui soit sans dommage quand il est reçu directement, mais qui produit des problèmes quand on y accède via l'URI. Une raison pour cela peut être le format utilisé dans l'archive, où certains contenus n'ont pas été échappés de façon adéquate. Cela peut être traité en utilisant un échappement adéquat.

Le champ d'en-tête Archived-At pointe sur une forme archivée du message lui-même. Elle peut à son tour contenir le champ Archived-At. Cela crée un potentiel d'attaque de déni de service sur le serveur pointé par l'URI dans le champ d'en-tête Archived-At. Les conditions sont que la forme archivée du message est téléchargée automatiquement, et que d'autres URI dans ce message sont suivis et téléchargés de façon récurrente sans vérification des ressources déjà téléchargées. Cependant, cette sorte de scénario peut être facilement évitée par les mises en œuvre. D'abord, l'URI dans le champ d'en-tête Archived-At ne devrait pas être déréférencé automatiquement. Ensuite, des mesures appropriées de détection de boucle devraient être utilisées.

Au paragraphe 3.2, une attaque est décrite qui peut casser un URI sur un message en introduisant un nouveau message avec le même identifiant de message. Des contre-mesures possibles sont aussi discutées.

5. Considérations relatives à l'IANA

5.1 Enregistrement du champ d'en-tête Archive-At

L'IANA a enregistré le champ d'en-tête Archived-At dans le registre des champs d'en-tête de message [RFC3864] comme suit :

Nom de champ d'en-tête : Archived-At

Protocole applicable : messagerie (RFC 2822) et netnews (RFC 1036)

Statut : standard

Auteur/Contrôleur des changements : IETF

Spécification : RFC 5064

Informations en rapport : aucune

5.2 Enregistrement du champ d'en-tête X-Archived-At

Ce paragraphe n'est pas normatif (précisément, une mise en œuvre qui ignore ce paragraphe reste conforme à la présente spécification).

L'IANA a enregistré le champ d'en-tête X-Archived-At dans le registre des champs d'en-tête de message [RFC3864] comme suit :

Nom de champ d'en-tête : X-Archived-At

Protocole applicable : messagerie (RFC 2822) et netnews (RFC 1036)

Statut : déconseillé

Auteur/Contrôleur des changements : IETF

Spécification : RFC 5064

Informations en rapport : aucune

6. Remerciements

Les membres de l'équipe système du W3C, en particulier Gerald Oskoboiny, Olivier Thereaux, Jose Kahan, et Eric Prud'hommeaux, ont créé le système de recherche d'archive de messagerie électronique mid-based et la forme expérimentale de l'en-tête Archived-At. Pete Resnik a fourni la motivation pour la rédaction du présent mémoire. La discussion sur la liste de diffusion ietf-822@imc.org, en particulier les contributions de Frank Ellermann, Arnt Gulbrandsen, Graham Klyne, Bruce Lilly, Charles Lindsey, et Keith Moore, ont conduit à d'autres améliorations de la proposition. Chris Newman, Chris Lonvick, Stephane Borzmeyer, Vijay K. Gurbani, et S. Moonesamy ont fourni des commentaires supplémentaires précieux.

7. Références

7.1 Références normatives

- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC2822] P. Resnick, "[Format de message Internet](#)", avril 2001. (Remplace la [RFC0822](#), STD 11, Remplacée par [RFC5322](#))
- [RFC3864] G. Klyne, M. Nottingham, J. Mogul, "Procédures d'[enregistrement pour les champs d'en-tête de message](#)", septembre 2004. ([BCP0090](#) ; MàJ par [RFC9110](#))
- [RFC3986] T. Berners-Lee, R. Fielding et L. Masinter, "[Identifiant de ressource uniforme](#) (URI) : Syntaxe générique", STD 66, janvier 2005. (P.S. ; MàJ par [RFC8820](#))
- [RFC3987] M. Duerst et M. Suignard, "[Identifiant de ressource internationalisé](#) (IRI)", janvier 2005.
- [RFC4234] D. Crocker et P. Overell, "[BNF augmenté pour les spécifications de syntaxe](#) : ABNF", octobre 2005. (Remplace [RFC2234](#), remplacée par [RFC5234](#))

7.1 Références pour information

- [RFC1036] M. Horton et R. Adams, "Norme pour l'[échange de messages](#) USENET", décembre 1987. (Remplacée par [RFC5536](#))
- [RFC2046] N. Freed et N. Borenstein, "[Extensions de messagerie Internet](#) multi-objets (MIME) Partie 2 : Types de prendre en charge", novembre 1996. (D. S., MàJ par [2646](#), [3798](#), [5147](#), [6657](#), [8098](#))
- [RFC2369] G. Neufeld, J. Baer, "Utilisation des [URL comme méta syntaxe](#) pour les commandes centrales de liste de messagerie et leur transport à travers les champs d'en-tête de message", juillet 1998. (P.S.)

Adresse de l'auteur

Martin Dürst
Aoyama Gakuin University
5-10-1 Fuchinobe
Sagamihara, Kanagawa 229-8558
Japan

téléphone : +81 42 759 6329
mél : duerst@it.aoyama.ac.jp
URI : <http://www.sw.it.aoyama.ac.jp/D%C3%BCrst/>

Déclaration complète de droits de reproduction

Copyright (C) The IETF Trust (2007)

Le présent document est soumis aux droits, licences et restrictions contenus dans le BCP 78, et sauf pour ce qui est mentionné ci-après, les auteurs conservent tous leurs droits.

Le présent document et les informations contenues sont fournis sur une base "EN L'ÉTAT" et le contributeur, l'organisation qu'il ou elle représente ou qui le/la finance (s'il en est), la INTERNET SOCIETY, le IETF TRUST et la INTERNET ENGINEERING TASK FORCE déclinent toutes garanties, exprimées ou implicites, y compris mais non limitées à toute garantie que l'utilisation des informations encloses ne viole aucun droit ou aucune garantie implicite de commercialisation ou d'aptitude à un objet particulier.

Propriété intellectuelle

L'IETF ne prend pas position sur la validité et la portée de tout droit de propriété intellectuelle ou autres droits qui pourraient être revendiqués au titre de la mise en œuvre ou l'utilisation de la technologie décrite dans le présent document ou sur la mesure dans laquelle toute licence sur de tels droits pourrait être ou n'être pas disponible ; pas plus qu'elle ne prétend avoir accompli aucun effort pour identifier de tels droits. Les informations sur les procédures de l'ISOC au sujet des droits dans les documents de l'ISOC figurent dans les BCP 78 et BCP 79.

Des copies des dépôts d'IPR faites au secrétariat de l'IETF et toutes assurances de disponibilité de licences, ou le résultat de tentatives faites pour obtenir une licence ou permission générale d'utilisation de tels droits de propriété par ceux qui mettent en œuvre ou utilisent la présente spécification peuvent être obtenues sur le répertoire en ligne des IPR de l'IETF à <http://www.ietf.org/ipr>.

L'IETF invite toute partie intéressée à porter son attention sur tous copyrights, licences ou applications de licence, ou autres droits de propriété qui pourraient couvrir les technologies qui peuvent être nécessaires pour mettre en œuvre la présente norme. Prière d'adresser les informations à l'IETF à ietf-ipr@ietf.org.