

RFC 6648 : Deprecating the X- Prefix and Similar Constructs in Application Protocols

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 25 juin 2012

Date de publication du RFC : Juin 2012

<https://www.bortzmeyer.org/6648.html>

Ce RFC met officiellement fin à une amusante tradition, qui était un des charmes des protocoles TCP/IP mais qui, en pratique, a causé quelques ennuis, justifiant son abandon. Cette tradition concernait le nommage des paramètres dans les protocoles. Normalement, ces noms étaient enregistrés à l'IANA pour éviter des collisions (par exemple, il ne doit y avoir qu'un seul champ nommé `User-Agent` : dans une requête HTTP). Comme cet enregistrement nécessitait une procédure, parfois lente et lourde, cela contrariait les gens qui voulaient expérimenter avec une nouvelle idée, qui impliquait un nouveau paramètre. L'habitude, avec l'appui de quelques RFC, s'était donc répandue de donner à ces paramètres expérimentaux, non officiels, un nom commençant par la lettre X suivie d'un tiret. Cette pratique a rencontré plusieurs problèmes et est donc désormais fortement déconseillée.

Le préfixe « X- » permettait de marquer clairement le nom de paramètre comme expérimental. Sans lui, certains développeurs auraient peut-être été tentés de s'auto-attribuer des noms, quitte à provoquer des collisions avec d'autres développeurs. Mais le problème de l'Internet est que les expériences durent : on invente un nouveau champ dans les requêtes HTTP, mettons `X-recipe` : pour indiquer sa recette favorite, on se dit que c'est juste un essai, qu'on ne va pas s'embêter à demander un nom officiel pour si peu, on l'utilise dans un "patch" de Firefox, dans un module Apache, des gens commencent à s'en servir, on se dit qu'il serait bon d'enregistrer un vrai nom et on s'aperçoit que la base installée d'utilisateurs est déjà trop grande pour abandonner le nom expérimental. C'est essentiellement pour éviter ce genre de blocage que l'IETF vient de décider de mettre fin à cette convention du « X- ».

Il y a d'autres protocoles que HTTP (d'où sont extraits les exemples précédents) qui utilisent de tels paramètres. Par exemple le format standard du courrier électronique (RFC 5322¹) ou bien les vCard

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc5322.txt>

(RFC 6350). Tous à un moment ou à un autre ont vu passer des paramètres dont les noms commençaient par « X- », X comme eXpérimental ou comme eXtension (SIP faisait un peu différemment avec des en-têtes dont le nom commençait par « P- », cf. RFC 5727). Une méta-information sur le paramètre, son statut expérimental, se trouvait ainsi stockée dans le nom. L'idée semblait tellement bonne qu'elle a même été codifiée par plusieurs RFC comme le RFC 822 (qui différençait les "*Extension fields*" et les "*User-defined fields*", section 4.1) ou comme le RFC 5545 (section 3.1). Par la suite, ces codifications ont souvent été retirées (pour le courrier, le RFC 2822 a annulé les dispositions du RFC 822).

Donc, ce nouveau RFC :

- Supprime la convention « X- », même pour les cas où elle était non-officielle et non documentée (cas de HTTP),
- Recommande à la place de nouvelles pratiques et de nouvelles conventions,
- N'interdit pas les noms de paramètres locaux ou expérimentaux : c'est la convention « X- » qui est en cause, pas l'idée de paramètres non enregistrés,
- Ne spécifie rien sur les paramètres « X- » existants, une question délicate qu'il faudra traiter au cas par cas.

Les programmeurs reçoivent (en section 2) la consigne de ne **pas** traiter un paramètre différemment juste sur la base de son nom. Qu'il commence par « X- » ou pas, le résultat doit être le même. (Faire autrement pourrait avoir des conséquences négatives sur la sécurité, par exemple si un programme considérait que les noms sans « X- » sont plus « sûrs » que les autres, comme expliqué en section 5.)

Quant aux créateurs de nouveaux paramètres (« Eh, je trouve qu'il faudrait pouvoir indiquer la boisson favorite de l'utilisateur, y a-t-il un en-tête `Favorite-Drink: ?` »), la section 3 leur dit que :

- Il faut s'attendre à ce que tout paramètre créé devienne un jour commun, largement déployé, non modifiable (c'est le grand problème avec la convention « X- » : les paramètres censés être expérimentaux deviennent permanents et non-modifiables),
- Il faut utiliser des noms parlants et non déjà utilisés,
- Et ne commençant pas par « X- ».

Pour limiter les risques de collision, si le paramètre n'est pas enregistré, cette section recommande de le préfixer avec le nom de l'organisation, à la Java, par exemple si l'organisation se nomme Example International et a le domaine `example.com`, utiliser `Example-International-Favorite-Drink:` ou `Com-Example-Favorite-Drink:` pour l'en-tête `experimental`.

Quant aux concepteurs des futurs protocoles, la section 4 leur dit :

- N'oubliez pas de créer un **registre** des valeurs possibles pour les paramètres textuels, registre qui permettra d'éviter les collisions entre noms de paramètres. C'est le cas des noms de champs, utilisés entre autres par HTTP `<https://www.iana.org/assignments/message-headers/perm-headers.html>`, par exemple.
- Idéalement, les procédures d'enregistrement dans ledit registre devraient être suffisamment simples et claires pour que les créateurs de paramètres (section précédente) ne soient pas tentés de choisir des noms de manière unilatérale,
- Ne faites pas de règles spécifiques pour les noms commençant par « X- ». Par exemple, ne spécifiez pas que ces noms sont interdits, ou qu'ils indiquent un type particulier de paramètre.

Reléguée en annexe A, mais essentielle, se trouve l'analyse historique de la convention « X- ». Il semble qu'elle ait été mentionnée la première fois par Brian Harvey en 1975, pour les paramètres FTP du RFC 691 (« "*an initial letter X be used for really local idiosyncracies*" »). Les RFC 737, RFC 743, RFC 775, RFC 822 ont ensuite repris l'idée, à laquelle les RFC 1123 et RFC 1154 ont donné une forme de reconnaissance officielle, en notant toutefois déjà le problème que posent des paramètres qui sont expérimentaux au début et qui deviennent normalisés sans le « X- » plus tard. On ne peut pas citer les nombreux RFC qui ont ensuite mentionné l'idée, sans forcément l'encourager. Mais même LDAP s'y était mis (RFC 4512). Par contre, cette convention n'a jamais reçu de validation au niveau de tout l'IETF : ni le RFC 2026, ni le RFC 5226 n'en parlent. Cette convention est donc (théoriquement) restée cantonnée à certains protocoles.

Les problèmes que pose cette convention « X- » ont quand même été notés assez tôt et, dès le RFC 2822, le courrier électronique arrêta officiellement de l'utiliser. Idem pour SIP avec le RFC 5727.

Si vous voulez tout savoir des inconvénients de la convention « X- », la partie du RFC à lire est l'annexe B, qui énumère les problèmes. Le principal est que les paramètres ainsi nommés, normalement locaux et/ou temporaires, tendent à fuir vers l'extérieur et que, s'ils sont des succès, ils deviennent rapidement des « standards de fait » (et les programmes doivent le traiter comme tel). Il faut alors se décider entre une migration depuis le nom en « X- » et le nom officiel sans « X- », ou bien une officialisation du nom avec « X- ». Dans les deux cas, la convention « X- » n'aura pas aidé. Elle ne servirait que si les paramètres ainsi nommés étaient tous des échecs qui ne restaient pas longtemps. Mais ce n'est pas le cas. Un exemple typique est donné par les en-têtes HTTP `x-gzip` et `x-compress`, qui sont devenus des standards de fait, au point que le RFC 2068 (section 3.5) conseillait de les accepter, comme synonymes aux en-têtes standard ! (Son successeur, le RFC 2616, a gardé cette recommandation.)

Dans le contexte du courrier électronique, on voit un phénomène équivalent avec le conseil du RFC 5064 d'accepter `X-archived-at` : comme équivalent à `archived-at` : (section 2.5). Excellent exemple d'une expérimentation qui a tellement bien réussi qu'elle devient ensuite une gêne. Quant au conseil de considérer les deux noms comme équivalents, il est dangereux : les deux noms ont été définis séparément et peuvent avoir des propriétés subtilement différentes (le nom officiellement enregistré a souvent des règles de sécurité plus strictes sur son usage, par exemple).

Bien sûr, il y a des gens qui défendent la convention « X- ». D'abord, il peut y avoir des cas où elle est légitime :

— Si on est absolument sûr que le paramètre ne sera jamais standardisé (l'expérience invite à la prudence...) Auquel cas l'annexe B recommande, pour éviter toute collision avec un nom officiel, de préfixer le nom du paramètre par celui de l'organisation.

— Quant un paramètre existe déjà et qu'il a le nom qu'on voudrait. Le RFC estime que c'est une mauvaise idée. Si on a un paramètre `priority` et qu'on est tenté de créer `x-priority`, il vaut mieux utiliser un synonyme (comme `urgency`) ou bien être créatif (`get-it-there-fast`).

— Quant les noms doivent être très courts (c'est le cas des étiquettes de langue du RFC 5646 où on peut utiliser RFC Mais là, il vaut encore mieux utiliser des numéros philosophiques : Les critères étant un signal clair du statut du paramètre (mais, en pratique, des paramètres « X- » sont devenus standardisés de fait).

— Les collisions entre noms sont plus facilement évitées s'il existe une ségrégation, avec un préfixe dédié pour les noms locaux et expérimentaux. Mais la plupart des espaces de nommage sont de grande taille : rien n'empêche de nommer un en-tête `Very-Long-Parameter-With-Useless-Information-Ju` pour être sûr qu'on ne le confonde pas avec un autre. (Les étiquettes de langue citées plus haut sont un contre-exemple, où l'espace de nommage n'est pas énorme.)

— En parlant d'économie, il existe un RFC, le RFC 3692 qui recommande d'utiliser des valeurs spéciales pour les expériences ou les cas locaux. Mais cela concerne surtout les paramètres numériques, où l'espace manque souvent. Si vous avez un seul octet pour stocker une valeur, enregistrer des paramètres expérimentaux risque d'épuiser rapidement cet octet. Un exemple est celui de l'algorithme de cryptographie utilisé par DNSSEC. Il n'y a que huit bits pour l'indiquer et il faut donc allouer des valeurs `<https://www.iana.org/assignments/dns-sec-alg-numbers/dns-sec-alg-numbers.xml>` prudemment. C'est pour cela que les valeurs 253 et 254 sont réservées pour les expérimentations et l'usage privé. Les paramètres alphabétiques n'ont en général pas ce problème : on n'est pas près de tomber à court de noms pour les en-têtes HTTP!

Si les gens utilisent la convention « X- », au lieu d'enregistrer le nom de leur paramètre, ce n'est pas uniquement pour semer des chausse-trappes sous les pieds des futurs programmeurs. C'est aussi parce que l'enregistrement d'un nom a souvent été considéré (souvent à juste titre) comme excessivement compliquée et bureaucratique. Si on veut mettre effectivement fin à la convention « X- », la bonne méthode est donc aussi de libéraliser l'enregistrement, ce qu'avaient fait les RFC 3864 ou RFC 4288. Plusieurs participants à l'IETF avaient fait remarquer lors des discussions sur ce RFC que le vrai enjeu était de normaliser plus rapidement, pour éviter que les concepteurs de nouveaux paramètres ne soient tentés de s'allouer des noms de manière sauvage.

Ce RFC ne s'applique qu'aux protocoles utilisant des paramètres textuels. Lorsque le paramètre est, au contraire, identifié par sa position ou par un nombre, il n'est évidemment pas concerné.