

RFC 6874 : Representing IPv6 Zone Identifiers in Address Literals and Uniform Resource Identifiers

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 février 2013

Date de publication du RFC : Février 2013

<https://www.bortzmeyer.org/6874.html>

Ce court RFC normalise la façon d'indiquer la **zone** dans un URI contenant une adresse IPv6 littérale.

Les zones sont définies et expliquées dans le RFC 4007¹. En gros, une adresse IPv6 n'est pas forcément globale, elle peut avoir une signification limitée à une seule zone, une zone étant un ensemble de sous-réseaux contigus, souvent composé d'un seul lien Ethernet. Les adresses locales au lien, par exemple, celles qui sont dans le préfixe `fe80::/10`, sont dans ce cas, le lien qui les relie forme une zone. Pour indiquer cette zone, la convention de mettre un %, puis l'identificateur de la zone, après l'adresse, est fréquemment utilisée. Mais elle n'était pas normalisée, dans le cas où l'adresse apparaissait dans un URI. Ce RFC répare ce manque.

Car il peut y avoir des adresses IP littérales dans un URI (section 3.2.2 du RFC 3986). Par exemple, `http://[2001:db8:1::23]/` est un URI valide. (Et je viens de tester avec mon Firefox qui l'a accepté sans problèmes.) Mais, si cette adresse est locale à une zone? D'habitude, on utilise la convention du % (décrite en section 11 du RFC 4007). Ici, sur une machine Linux récente (cela ne marchait pas autrefois, il fallait utiliser l'option `-I`) :

```
% ping6 -c3 fe80::21e:8cff:fe76:29b6%eth0
PING fe80::21e:8cff:fe76:29b6%eth0(fe80::21e:8cff:fe76:29b6) 56 data bytes
64 bytes from fe80::21e:8cff:fe76:29b6: icmp_seq=1 ttl=64 time=8.36 ms
64 bytes from fe80::21e:8cff:fe76:29b6: icmp_seq=2 ttl=64 time=4.97 ms
64 bytes from fe80::21e:8cff:fe76:29b6: icmp_seq=3 ttl=64 time=4.84 ms

--- fe80::21e:8cff:fe76:29b6%eth0 ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2003ms
rtt min/avg/max/mdev = 4.847/6.060/8.363/1.631 ms
```

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc4007.txt>

On teste la machine `fe80::1:23` reliée au réseau identifié par `eth0` (dans le cas de Linux, c'est le nom d'une interface réseau mais d'autres systèmes d'exploitation peuvent utiliser une syntaxe différente, par exemple identifier les zones par un simple numéro). Cette possibilité de désigner une zone spécifique est essentielle pour les activités de débogage (M. Toutlemonde n'a sans doute jamais besoin d'URI incluant des adresses IP). Mais si on utilise un navigateur Web pour cette activité, comment mettre la zone? Le RFC 4007 était muet à ce sujet. Certains navigateurs (par exemple certaines versions anciennes de Firefox) acceptaient des adresses suivies d'un pour-cent (alors que ce caractère est spécial dans les URI et que ces navigateurs violaient donc le RFC 3986).

Encore mieux, comme le RFC 4007 ne précise pas quels sont les caractères autorisés dans un identificateur de zone, on pourrait voir des zones avec un nom invalide dans un URI (comme / ou #).

La syntaxe standard normalisée par notre RFC est donc :

- Mettre l'identificateur de zone après un pour-cent encodé (donc, %25, ce qui était déjà la méthode déployée par Internet Explorer),
- Fortement conseiller de ne pas utiliser de caractères spéciaux dans les identificateurs de zone mais, si cela arrive, les encoder avec le système « pour-cent » du RFC 3986 (section 2.1),
- Suivre les recommandations du RFC 5952 pour représenter l'adresse IPv6.

Pour faciliter la vie de l'administrateur réseaux, notre RFC recommande que les navigateurs acceptent un URI incluant un % tout nu, s'il est suivi d'au moins deux caractères qui ne peuvent pas être des chiffres hexadécimaux. Cela permet de copier-coller une adresse comme `fe80::a%en1` dans le navigateur et que cela marche. (Par contre, `fe80::a%ee1` ne serait pas accepté car E peut être un chiffre hexadécimal et il y aurait une ambiguïté entre %ee et î, qui est encodé ainsi en notation pour-cent.) Cette opération de copier/coller sera sans doute le principal usage de ces URI comportant une adresse IPv6 et une zone : on ne voit pas bien dans quels cas ils se retrouveraient dans une page HTML, sauf peut-être produite automatiquement par des systèmes d'administration de réseaux (l'interface Web d'un routeur, par exemple).

Donc, désormais, `<http://[fe80::dead:beef%25en1]>` ou `<http://[fe80::1:2%25Ethernet%230]>` sont des URI légaux (dans le deuxième, le nom de la zone était `Ethernet/0`). Essayez-les sur votre navigateur!

Comme ces identificateurs de zone n'ont de signification que pour une machine donnée, les relais HTTP, par exemple, doivent les retirer des URI.

Voilà, c'est tout, les amateurs d'alternatives peuvent lire l'annexe A, consacrée aux raisons du choix qui a été fait. Il a l'avantage de ne pas changer le format des URI et les inconvénients qu'il n'est pas très joli, et que le copier/coller d'une console vers un navigateur ne marche pas toujours. Les autres possibilités considérées (les discussions ont été chaudes et longues) étaient :

- Ne rien normaliser, en considérant que ping ou telnet depuis une console étaient plus pratique que d'utiliser un navigateur Web. C'était le choix le plus simple.
- Entériner l'usage des % nus dans l'URI. Cela était cohérent avec le RFC 4007 et cela permettait le copier/coller mais cela produisait des URI illégaux.
- Trouver un autre caractère que le % pour séparer l'adresse de la zone, par exemple le -, comme dans `http://[fe80::a-ee1]` (au passage, pourquoi pas un tiret bas? parce que les navigateurs soulignent souvent automatiquement les URI, rendant le trait invisible). Mais une telle réforme du RFC 4007 aurait nécessité de changer tous les outils IPv6 et toutes les documentations.
- Se servir de la production `IPvFuture` de la section 3.2.2 du RFC 3986 pour écrire, par exemple `http://[v6.fe80::a-en1]`. Mais cela ne permettait pas le copier/coller. (Et j'ajoute qu'aucun navigateur ne gère cette possibilité `IPvFuture`.)