RFC 9549 : Internationalization Updates to RFC 5280

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 22 novembre 2024

Date de publication du RFC : Mars 2024

https://www.bortzmeyer.org/9549.html

Ce court RFC ajoute aux certificats PKIX du RFC 5280[1] la possibilité de contenir des adresses de courrier électronique dont la partie locale est en Unicode. Et il modifie légèrement les règles pour les noms de domaine en Unicode dans les certificats. Il remplace le RFC 8399, avec peu de changements.

Les certificats sur l'Internet sont normalisés dans le RFC 5280, qui décrit un profil de X.509 nommé PKIX (définir un profil était nécessaire car la norme X.509 est bien trop riche et complexe). Ce RFC 5280 permettait des noms de domaine en Unicode (sections 4.2.1.10 et 7 du RFC 5280) mais il suivait l'ancienne norme IDN, celle des RFC 3490 et suivants. Depuis, les IDN sont normalisés dans le RFC 5890 et suivants, et notre nouveau RFC 9549 modifie très légèrement le RFC 5280 pour s'adapter à cette nouvelle norme de noms de domaines en Unicode. Les noms de domaine dans un certificat peuvent être présents dans les champs Sujet (titulaire du certificat) et Émetteur (AC ayant signé le certificat) mais aussi dans les contraintes sur le nom (une autorité de certification peut être limitée à des noms se terminant en `example.com`, par exemple).

Notez que, comme avant, ces noms sont exprimés dans le certificat en Punycode (RFC 3492, `xn--caf-dma.fr` au lieu de `café.fr`), appelé "*A-label*" dans notre RFC. C'est un bon exemple du fait que les limites qui rendaient difficiles d'utiliser des noms de domaine en Unicode n'avaient rien à voir avec le DNS (qui n'a jamais été limité à ASCII <https://www.bortzmeyer.org/pourquoi-idn-et-pas-un-dns-unicode.html>, contrairement à ce qu'affirme une légende courante). En fait, le problème venait des applications (comme PKIX), qui ne s'attendaient pas à des noms en Unicode. Un logiciel qui traite des certificats aurait été bien étonné de voir des noms de domaines non-ASCII, et aurait peut-être planté. D'où ce choix du Punycode ("*A-label*") par opposition à la forme plus lisible du "*U-label*".

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc5280.txt`

Nouveauté plus importante du RFC 8399, mis à jour par notre RFC 9549, les adresses de courrier électronique en Unicode (EAI pour "*Email Address Internationalization*"). Elles étaient déjà permises par la section 7.5 du RFC 5280, mais seulement pour la partie domaine (à droite du @). Désormais, elles sont également possibles dans la partie locale (à gauche du @). Le RFC 8398 donne tous les détails sur ce sujet.

Reste à savoir quelles AC acceptent Unicode. En 2018, j'avais testé avec Let's Encrypt (avec le client Dehydrated <https://dehydrated.io/>, en mettant le Punycode dans `domains.txt`) et ça marchait, regardez le certificat de . Le voici, affiché par GnuTLS :

```
% gnutls-cli www.potamochère.fr
...
- subject 'CN=www.xn--potamochre-66a.fr', issuer 'CN=R10,O=Let's Encrypt,C=US', serial 0x03a34cfc017b4311da0
```

D'autres AC acceptent ces noms en Unicode : Gandi le fait aussi, par exemple. On notera que le célèbre service de test de la qualité des configurations TLS, SSLlabs <https://www.ssllabs.com/>, gère bien les IDN :

Enfin, le registre du `.ru` a participé au développement de logiciels <https://cctld.ru/en/media/news/kc/22372/> pour traiter l'Unicode dans les certificats.

La section 1.2 résume les changements depuis le RFC 8399. Notamment :

— Le RFC 8399 ne les mentionnait pas mais les emojis ne sont pas autorisés dans les noms de domaine <https://www.afnic.fr/observatoire-ressources/papier-expert/peut-on-mettre-d> (ils appartiennent à la catégorie Unicode So, « Autres symboles », que le RFC 5890 n'inclut pas). Comme le rappelle notre RFC 9549, un nom comme `[Caractère Unicode non montré`[2]`].example` ne serait pas légal. Ce point est désormais explicite.

— Si le RFC 8399 prescrivait certains traitements à faire sur la forme Unicode ("*U-label*") des noms, désormais, tout se fait sur la forme en Punycode ("*A-label*", donc `xn--potamochre-66a.fr` et pas `potamochère.fr`). Cela permet d'éviter de déclencher certaines failles de sécurité <https://securitylabs.datadoghq.com/articles/openssl-november-1-vulnerabilities/>.

2. Car trop difficile à faire afficher par LaTeX