

Équipe d'ingénierie de l'Internet (IETF)

J. Klensin

Request for Comments : 5891

août 2010

RFC rendue obsolète : 3490, 3491

RFC mise à jour : 3492

Catégorie : En cours de normalisation

ISSN: 2070-1721

Traduction Claude Brière de L'Isle

Noms de domaines internationalisés pour les applications (IDNA) : le protocole

Résumé

Le présent document est la définition du protocole révisé des noms de domaines internationalisés (IDN, *Internationalized Domain Names*). Les raisons de ces changements, les relations avec l'ancienne spécification, la terminologie importante, sont fournies dans d'autres documents. Le présent document spécifie le mécanisme du protocole, appelé noms de domaines internationalisés dans les applications (IDNA, *Internationalized Domain Names in Applications*) pour l'enregistrement et les recherches des IDN d'une façon qui n'exige pas de changement au DNS lui-même. IDNA est seulement destiné au traitement des noms de domaines, pas au texte libre.

Statut de ce mémoire

Ceci est un document de l'Internet en cours de normalisation.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG). Tous les documents approuvés par l'IESG ne sont pas candidats à devenir une norme de l'Internet ; voir la Section 2 de la RFC5741.

Les informations sur le statut actuel du présent document, tout errata, et comment fournir des réactions sur lui peuvent être obtenues à <http://www.rfc-editor.org/info/rfc5891>

Notice de droits de reproduction

Copyright (c) 2010 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (<http://trustee.ietf.org/license-info>) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document. Les composants de code extraits du présent document doivent inclure le texte de licence simplifié de BSD comme décrit au paragraphe 4.e des dispositions légales du Trust et sont fournis sans garantie comme décrit dans la licence de BSD simplifiée.

Le présent document peut contenir des matériaux provenant de documents de l'IETF ou de contributions à l'IETF publiées ou rendues disponibles au public avant le 10 novembre 2008. La ou les personnes qui ont le contrôle des droits de reproduction sur tout ou partie de ces matériaux peuvent n'avoir pas accordé à l'IETF Trust le droit de permettre des modifications de ces matériaux en dehors du processus de normalisation de l'IETF. Sans l'obtention d'une licence adéquate de la part de la ou des personnes qui ont le contrôle des droits de reproduction de ces matériaux, le présent document ne peut pas être modifié en dehors du processus de normalisation de l'IETF, et des travaux dérivés ne peuvent pas être créés en dehors du processus de normalisation de l'IETF, excepté pour le formater en vue de sa publication comme RFC ou pour le traduire dans une autre langue que l'anglais.

Table des matières

1. Introduction.....	2
2. Terminologie.....	2
3. Exigences et applicabilité.....	3
3.1 Exigences.....	3
3.2 Applicabilité.....	3
4. Protocole d'enregistrement.....	4
4.1 Entrée d'enregistrement IDNA.....	4
4.2 Caractère permis et validation d'étiquette.....	4

4.3 Restrictions de registre.....	5
4.4 Conversion Punycode.....	5
4.5 Insertion dans la zone.....	6
5. Protocole de recherche de nom de domaine.....	6
5.1 Entrée de chaîne d'étiquette.....	6
5.2 Conversion en Unicode.....	6
5.3. Entrée d'étiquette A.....	6
5.4 Validation et vérification de liste de caractères.....	7
5.5 Conversion Punycode.....	7
5.6 Résolution de nom DNS.....	7
6. Considérations sur la sécurité.....	8
7. Considérations relatives à l'IANA.....	8
8. Contributeurs.....	8
9. Remerciements.....	8
10. Références.....	8
10.1 Références normatives.....	8
10.2. Références pour information.....	9
Appendice A. Résumé des changements majeurs par rapport à IDNA2003.....	9
Adresse de l'auteur.....	10

1. Introduction

Le présent document fournit la définition du protocole pour les noms de domaines internationalisés dans les applications (IDNA, *Internationalized Domain Names in Applications*) dont la version spécifiée ici est appelée IDNA2008. Les définitions et la terminologie essentielles pour la compréhension de ce document et une description de la collection de documents qui constituent IDNA2008 apparaissent dans un document de définitions séparé [RFC5890]. L'Appendice A discute des relations entre la présente spécification et la version antérieure de IDNA (à laquelle on se réfère ici comme "IDNA2003"). La raison de ces changements, ainsi qu'un considérable matériel explicatif et des avis aux administrateurs de zone qui prennent en charge les IDN, sont fournis dans un autre document, appelé de façon informelle dans cette série le "document de raisons" [RFC5894].

IDNA fonctionne en permettant aux applications d'utiliser certaines étiquettes de chaîne ASCII [ASCII] (commençant par un préfixe spécial) de représenter des étiquettes de noms non ASCII. Les protocoles de couche inférieure n'ont pas besoin de connaître cela ; donc, IDNA ne change aucune infrastructure. En particulier, IDNA ne dépend pas de changements au serveurs, résolveurs ou éléments de protocole du DNS, parce que le service de noms ASCII fourni par le DNS existant peut être utilisé pour IDNA.

IDNA ne s'applique qu'à un sous ensemble spécifique des étiquettes du DNS. Les normes de base du DNS [RFC1034], [RFC1035], et leurs diverses mises à jour spécifient comment combiner les étiquettes en noms de domaines pleinement qualifiés et comment analyser les étiquettes à partir de ces noms.

Le présent document décrit deux protocoles distincts, un pour l'enregistrement IDN (Section 4) et un pour les recherches d'IDN (Section 5). Ces deux protocoles partagent la terminologie, les données de référence, et les opérations.

2. Terminologie

Comme mentionné ci-dessus, la terminologie utilisée au titre de la définition de IDNA apparaît dans le document de définitions [RFC5890]. Il vaut de noter qu'une partie de cette terminologie se chevauche, tout en restant cohérente avec elle, avec celle utilisée dans Unicode ou les autres jeux de caractères standard et dans le DNS. Les lecteurs du document sont supposé familiarisés avec le document de définitions associé avec la terminologie spécifique du DNS de la [RFC1034].

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].

3. Exigences et applicabilité

3.1 Exigences

IDNA pose les exigences suivantes :

1. Chaque fois qu'un nom de domaine est mis dans un créneau de nom de domaines qui n'est pas à capacité IDNA (voir le paragraphe 2.3.2.6 du document de définitions [RFC5890]), il DOIT contenir seulement des caractères ASCII (c'est-à-dire, ses étiquettes doivent être soit des étiquettes A, soit des étiquettes LDH-NR) sauf si l'application DNS n'est pas soumise aux recommandations historiques des noms de style "nom d'hôte" (voir la [RFC1034] et le paragraphe 3.2.1).
2. Les étiquettes DOIVENT être comparées en utilisant des formes équivalentes : soit toutes deux des formes d'étiquette A, soit toutes deux des formes d'étiquette U. Comme les étiquettes A et les étiquettes U peuvent être transformées de l'une dans l'autre sans perte d'information, ces comparaisons sont équivalentes (cependant, en pratique, la comparaison des étiquettes U exige d'abord de vérifier qu'elles sont effectivement des étiquettes U et pas juste des chaînes Unicode). Une paire d'étiquettes A DOIT être comparée comme de l'ASCII insensible à la casse (comme avec toutes les comparaisons d'étiquettes ASCII du DNS). Les étiquettes U DOIVENT être comparées telles qu'elles, sans changement de casse ni autres étapes intermédiaires. Bien qu'il ne soit pas nécessaire de valider les étiquettes dans l'ordre où on les compare, une comparaison réussie n'implique pas la validité. Dans de nombreux cas, et pas seulement des comparaisons, la validation peut être importante pour d'autres raisons et DEVRAIT être effectuée.
3. Les étiquettes à enregistrer DOIVENT se conformer aux exigences de la Section 4. Les étiquettes qui font l'objet d'une recherche et le processus de recherche DOIVENT se conformer aux exigences de la Section 5.

3.2 Applicabilité

IDNA s'applique à tous les noms de domaines dans tous les créneaux de noms de domaines des protocoles excepté lorsque il est explicitement exclu. Il ne s'applique pas aux créneaux de noms de domaines qui n'utilisent pas les règles de la syntaxe de LDH telles que décrites par le document de définitions [RFC5890].

Parce qu'il utilise le DNS, IDNA s'applique à de nombreux protocoles qui ont été spécifiés avant sa conception. Les IDN occupant un créneau de noms de domaines dans de plus anciens protocoles DOIVENT être en forme d'étiquette A jusqu'à ce que ces protocoles et leurs mises en œuvre soient explicitement mises à niveau pour avoir la capacité IDN et accepter la forme d'étiquette U. Les IDN qui apparaissent effectivement dans les interrogations ou réponses du DNS DOIVENT être des étiquettes A.

Les protocoles et mises en œuvre à capacité IDNA PEUVENT accepter des étiquettes U, des étiquettes A, ou les deux, selon ce que leurs protocoles particuliers spécifient. IDNA n'est pas défini pour des types d'étiquettes étendus (voir la Section 3 de la [RFC2671]).

3.2.1 Enregistrements de ressource du DNS

IDNA ne s'applique qu'aux noms de domaines dans les champs NAME et RDATA des enregistrements de ressource du DNS dont la CLASS est IN. Voir dans la spécification du DNS [RFC1035] la définition précise de ces termes.

L'application de IDNA aux enregistrements de ressource du DNS dépend entièrement de la classe (*CLASS*) de l'enregistrement, et non du type (*TYPE*) sauf comme noté ci-dessous. Cela va rester vrai, même lorsque de nouveaux types sont définis, sauf si un nouveau type définit des règles spécifiques du type. Des conventions de désignation particulières pour les enregistrements SRV (et d'une façon plus générale pour les "étiquettes à souligné") sont incompatibles avec le codage IDNA comme expliqué dans le document des définitions [RFC5890], en particulier au paragraphe 2.3.2.3. Bien sûr, les étiquettes à souligné peuvent faire partie d'un domaine qui utilise des étiquettes IDN à des niveaux plus élevés de l'arborescence.

3.2.2 Types de données non nom de domaine mémorisées dans le DNS

Bien que IDNA permette la représentation de caractères non ASCII dans les noms de domaines, cela n'implique pas que IDNA permet la représentation de caractères non ASCII dans les autres types de données mémorisées dans les noms de domaines, en particulier dans le champ RDATA pour les types qui ont un format RDATA structuré. Par exemple, une partie locale d'adresse de messagerie électronique est mémorisée dans un nom de domaine dans le champ RNAME au titre des RDATA d'un enregistrement SOA (par exemple, `hostmaster@exemple.com` serait représenté par `hostmaster.exemple.com`). IDNA ne met pas à jour les normes existantes de messagerie, qui permettent seulement des

caractères ASCII dans les parties locales. Bien que des travaux soient en cours pour définir l'internationalisation des adresses de messagerie électronique [RFC4952], des changements à la partie adresse de messagerie électronique des RDATA de SOA exigeraient des actions, ou des mises à jour, sur d'autres normes, en particulier celles qui spécifient le format de l'enregistrement de ressource SOA.

4. Protocole d'enregistrement

Cette section définit le modèle pour enregistrer un IDN. Le modèle est indépendant de la mise en œuvre ; toute séquence d'étapes qui produit exactement le même résultat pour toutes les étiquettes est considérée comme une mise en œuvre valide.

Noter que, alors que les protocoles d'enregistrement (cette section) et de recherche (Section 5) sont très similaires sous la plupart des aspects, ils ne sont pas identiques, et les développeurs devraient suivre scrupuleusement les étapes décrites dans la présente spécification.

4.1 Entrée d'enregistrement IDNA

Les processus d'enregistrement, et particulièrement par des entités (souvent appelées "registraires") qui traitent des enregistrements avant que la demande atteigne en fait le gestionnaire de zone (le "registre") sortent du domaine d'application de cette définition et peuvent différer significativement en fonction des besoins locaux. Au moment où une chaîne entre dans le processus d'enregistrement IDNA comme décrit dans la présente spécification, elle DOIT être en Unicode et en format de normalisation C (NFC, *Normalisation Form C*) [UAX15]. Les entités responsables des fichiers de zone ("registres") DOIVENT n'accepter que les chaînes exactes pour lesquelles l'enregistrement est demandé, sans aucune transposition ni ajustement local. Elles PEUVENT accepter cette entrée dans une des trois formes suivantes :

1. comme paire d'étiquettes A et U,
2. comme étiquette A seulement,
3. comme étiquette U seulement.

Les deux premières de ces formes sont RECOMMANDÉES parce que l'utilisation des étiquettes A évite toute possibilité d'ambiguïté. La première est normalement préférée à la seconde parce que elle permet une vérification ultérieure de l'intention de l'utilisateur (voir le paragraphe 4.2.1).

4.2 Caractère permis et validation d'étiquette

4.2.1 Format d'entrée

Si les deux formes d'étiquette U et A sont disponibles, le registre DOIT s'assurer que la forme d'étiquette A est en minuscules, effectuer une conversion en étiquette U, effectuer les étapes et vérifications décrites ci-dessous sur cette étiquette U, puis vérifier que l'étiquette A produite par les étapes du paragraphe 4.4 correspond à celle produite en entrée. De plus, l'étiquette U qui a été fournie en entrée et celle obtenue par la conversion de l'étiquette A DOIVENT correspondre exactement. Si, pour une raison quelconque, ces essais échouent, l'enregistrement DOIT être rejeté.

Si seulement une étiquette A a été fournie et si la conversion en étiquette U n'est pas effectuée, le registre DOIT quand même vérifier que l'étiquette A est superficiellement valide, c'est-à-dire, qu'elle ne viole aucune des règles du codage Punycode [RFC3492] comme la prohibition d'un tiret/signé moins en queue, l'exigence que tous les caractères soient de l'ASCII, et ainsi de suite. Les chaînes qui apparaissent comme des étiquettes A (par exemple, qui commencent par "xn--") et les chaînes qui sont fournies au registre dans un contexte réservé aux étiquettes A (comme un champ dans un formulaire à remplir) mais ne sont pas des étiquettes A valides comme décrit dans ce paragraphe, NE DOIVENT PAS être placées dans les zones du DNS qui prennent en charge IDNA.

Si seule une étiquette A est fournie, la conversion en étiquette U n'est pas effectuée, mais les vérifications superficielles décrites au paragraphe précédent sont effectuées, les procédures d'enregistrement PEUVENT, et généralement vont, court-circuiter les essais et actions dans le choix entre le paragraphe 4.2 et les paragraphes 4.3 et 4.4.

4.2.2 Rejet des caractères qui ne sont pas permis

La chaîne candidate Unicode NE DOIT PAS contenir de caractères qui apparaissent dans les listes "NON PERMIS" et "NON ALLOUÉS" spécifiées dans le document Tableaux [RFC5892].

4.2.3 Validation d'étiquette

L'étiquette proposée (sous la forme d'une chaîne Unicode, c'est-à-dire, une chaîne qui apparaît au moins superficiellement comme une étiquette U) est alors examinée en utilisant les essais qui exigent l'examen de plus d'un caractère. L'ordre des caractères est considéré comme étant celui des octets du réseau. Cet ordre peut n'être pas le même que celui de l'affichage.

4.2.3.1 Restrictions sur les tirets

La chaîne Unicode NE DOIT PAS contenir "--" (deux tirets consécutifs) dans les troisièmes et quatrièmes positions de caractères et NE DOIT PAS commencer ou se terminer par un tiret "-".

4.2.3.2 Marques de combinaison en tête

La chaîne Unicode NE DOIT PAS commencer par une marque de combinaison ou un caractère de combinaison (voir la dans la norme Unicode, paragraphe 2.11 [Unicode] la définition exacte).

4.2.3.3 Règles contextuelles

La chaîne Unicode NE DOIT PAS contenir de caractères dont la validité dépend du contexte, sauf si la validité est positivement confirmée par une règle contextuelle. Pour vérifier cela, chaque codet identifié comme CONTEXTJ ou CONTEXTO dans le document Tableaux [RFC5892] DOIT avoir une règle non nulle. Si ce codet n'a pas de règle, l'étiquette est invalide. Si la règle existe mais si le résultat de l'application de la règle est négatif ou non concluant, l'étiquette proposée est invalide.

4.2.3.4 Étiquettes contenant des caractères écrits de droite à gauche

Si l'étiquette proposée contient des caractères provenant de scripts qui sont écrits de droite à gauche, elle DOIT satisfaire aux critères Bidi de la [RFC5893].

4.2.4 Exigences de validation d'enregistrement

Les chaînes qui contiennent au moins un caractère non ASCII, ont été produites par les étapes ci-dessus, dont le contenu a réussi à toutes les vérifications du paragraphe 4.2.3, et font 63 caractères ou moins en codage de forme compatible ASCII (ACE) (voir le paragraphe 4.4) sont des étiquettes U.

Pour résumer, les essais sont faits au paragraphe 4.2 pour les caractères invalides, les combinaisons de caractères invalides, pour les étiquettes qui sont invalides même si les caractères qu'elles contiennent sont individuellement valides, et pour les étiquettes qui ne se conforment pas aux restrictions sur les chaînes contenant des caractères de droite à gauche.

4.3 Restrictions de registre

En plus des règles et vérifications ci-dessus, il y a de nombreuses raisons pour qu'un registre rejette une étiquette. Les registres à tous les niveaux du DNS, pas seulement au niveau supérieur, sont supposés établir des politiques d'enregistrement des étiquettes. Les politiques seront probablement informées des langages locaux et des scripts qui sont utilisés pour les écrire et qui peuvent dépendre de nombreux facteurs incluant quels caractères sont dans l'étiquette (par exemple, une étiquette peut être rejetée sur la base des autres étiquettes déjà enregistrées). Voir au paragraphe 3. 2 du document Raisons [RFC5894] pour plus d'explications et des recommandations sur les politiques de registre.

La chaîne produite par les étapes du paragraphe 4.2 est vérifiée et traitée comme approprié par rapport aux restrictions du registre local. L'application de ces restrictions de registre peut résulter en le rejet de certaines étiquettes ou en l'application de restrictions particulières aux autres.

4.4 Conversion Punycode

Les étiquettes U résultantes sont converties en une étiquette A (définie au paragraphe 2.3.2.1 du document de définitions [RFC5890]). L'étiquette A est le codage de l'étiquette U conformément à l'algorithme Punycode [RFC3492] avec le préfixe ACE "xn--" ajouté au début de la chaîne. La chaîne résultante doit, bien sûr, se conformer aux limites de longueur imposées par le DNS. Le présent document ne met ni à jour ni n'altère d'aucune façon l'algorithme Punycode spécifié dans la RFC 3492. La RFC 3492 fait une référence non normative aux informations sur la valeur et la construction du préfixe ACE qui apparaît dans la RFC 3490 ou dans Nameprep [RFC3491]. Pour les besoins de cohérence et l'agrément du lecteur,

IDNA2008 met effectivement à jour cette référence pour pointer sur le présent document. Ce changement n'altère pas le préfixe lui-même. Le préfixe "xn--" est le même dans les deux ensembles de documents.

À l'exception de la vérification de la longueur maximum de chaîne sur le résultat de Punycode, les conditions d'échec identifiées dans la procédure de codage Punycode ne peuvent pas se produire si l'entrée est une étiquette U comme déterminé par les étapes des paragraphes 4.1 à 4.3 ci-dessus.

4.5 Insertion dans la zone

L'étiquette est enregistrée dans le DNS en insérant l'étiquette A dans une zone.

5. Protocole de recherche de nom de domaine

La recherche est différente de l'enregistrement et des vérifications différentes sont appliquées au client. Bien que certaines vérifications de validité soient nécessaires pour éviter de sérieux problèmes avec le protocole, les essais en matière de recherche sont moins permissifs et s'appuient sur l'hypothèse que les noms qui sont présents dans le DNS sont valides. Cette hypothèse est, cependant, assez faible parce que la présence de caractères génériques dans le DNS peut être cause qu'une chaîne qui n'est en fait pas enregistrée dans le DNS fasse néanmoins l'objet d'une recherche réussie.

5.1 Entrée de chaîne d'étiquette

L'utilisateur fournit une chaîne dans le jeu de caractères local, par exemple, en la tapant, en cliquant dessus, ou en la copiant et en la passant à partir d'un identifiant de ressource, par exemple, un identifiant de ressource universel (URI) [RFC3986] ou un identifiant de ressource internationalisée (IRI) [RFC3987], à partir duquel le nom de domaine est extrait. Autrement, un processus n'impliquant pas directement l'utilisateur peut lire la chaîne à partir d'un fichier ou l'obtenir de quelque autre façon. Le traitement de cette étape et de celle spécifiée au paragraphe 5.2 sont des affaires locales, à accomplir avant l'invocation réelle de IDNA.

5.2 Conversion en Unicode

La chaîne est convertie du jeu de caractères local en Unicode, si elle ne l'est pas déjà. Selon les besoins locaux, cette conversion peut impliquer de transposer des caractères en d'autres caractères ainsi que des conversions de codage. Ces problèmes sont discutés dans les sections relatives à la transposition (paragraphes 4.2, 4.4, 6, et 7.3) du document des raisons [RFC5894] et dans le document séparé sur les transpositions [RFC5895]. Le résultat DOIT être une chaîne Unicode en forme NFC.

5.3. Entrée d'étiquette A

Si l'entrée à cette procédure apparaît être une étiquette A (c'est-à-dire, si elle commence par un "xn--", interprétée sans égard à la casse) l'application de recherche PEUT tenter de la convertir en une étiquette U, en s'assurant d'abord que l'étiquette A est entièrement en minuscules (la convertir en minuscules si nécessaire) et en appliquant les essais du paragraphe 5.4 et la conversion du paragraphe 5.5 en cette forme. Si l'étiquette est convertie en Unicode (c'est-à-dire, en forme d'étiquette U) en utilisant l'algorithme de décodage Punycode, alors le traitement spécifié dans ces deux paragraphes DOIT être effectué, et l'étiquette DOIT être rejetée si l'étiquette résultante n'est pas identique à l'original. Voir au paragraphe 8.1 du document de raisons [RFC5894] une discussion supplémentaire sur ce sujet.

La conversion à partir de l'étiquette A et la vérification que le résultat est une étiquette U DEVRAIT être effectué si le nom de domaine va plus tard être présenté à l'utilisateur sous sa forme de caractère native (cela exige que l'application de recherche soit à capacité IDNA). Si ces étapes ne sont pas effectuées, le processus de recherche DEVRAIT au moins vérifier que la chaîne est bien une étiquette A, en examinant si elle a un des formats invalides spécifiés dans la spécification de décodage Punycode. Les applications qui n'ont pas la capacité IDNA vont évidemment omettre cette vérification ; les autres PEUVENT traiter la chaîne comme opaque pour éviter le traitement supplémentaire au prix d'une protection moindre et de moins d'informations pour l'utilisateur.

5.4 Validation et vérification de liste de caractères

Comme avec la procédure d'enregistrement décrite à la Section 4, la chaîne Unicode est vérifiée pour voir si tous les caractères qui y apparaissent sont des entrées valides au traitement de recherche IDNA. Comme discuté ci-dessus et dans le document des raisons [RFC5894], la vérification de recherche est plus libérale que celle d'enregistrement. Les étiquettes qui n'ont pas été complètement évaluées quant à leur conformité aux règles applicables sont appelées étiquettes "putatives" comme expliqué au paragraphe 2.3.2.1 du document de définitions [RFC5890]. Les étiquettes U putatives avec une des caractéristiques suivantes DOIVENT être rejetées avant la recherche dans le DNS :

- o Étiquettes qui ne sont pas en NFC [UAX15].
- o Étiquettes qui contiennent "--" (deux tirets consécutifs) dans les positions de caractère trois et quatre.
- o Étiquettes dont le premier caractère est une marque de combinaison (voir le paragraphe 2.11 de la norme [Unicode]).
- o Étiquettes qui contiennent des codets prohibés, c'est-à-dire, ceux qui appartiennent à la catégorie "NON PERMIS" du document Tableaux [RFC5892].
- o Étiquettes qui contiennent des codets qui sont identifiés dans le document Tableaux comme "CONTEXTJ", c'est-à-dire, qui exigent un traitement de règle contextuelle exceptionnelle sur la recherche, mais ne se conforment pas à ces règles. Noter que cela implique qu'une règle doit être définie, non nulle : un caractère qui exige une règle contextuelle mais pour lequel la règle est nulle est traité dans cette étape comme ayant échoué à se conformer à la règle.
- o Étiquettes qui contiennent des codets qui sont identifiés dans le document Tableaux comme "CONTEXTO", mais pour lesquels aucune de ces règles n'apparaît dans le tableau des règles. Les applications qui résolvent les noms du DNS ou effectuent des opérations équivalentes ne sont pas obligées de vérifier les règles contextuelles pour les caractères "CONTEXTO", seulement de vérifier qu'une règle est définie (bien qu'elles PUISSENT faire de telles vérifications pour fournir une meilleure protection ou donner de meilleures informations à l'utilisateur).
- o Étiquettes qui contiennent des codets qui ne sont pas alloués dans la version de Unicode utilisée par l'application, c'est-à-dire, dans la catégorie NON ALLOUÉ du document Tableaux.

Cette exigence signifie que l'application doit utiliser une liste de caractères non alloués qui corresponde à la version de Unicode utilisée pour les autres exigences de ce paragraphe. Il n'est pas exigé que l'application sache quelle version de Unicode est utilisée ; cette information peut faire partie de l'environnement de fonctionnement de l'application.

De plus, l'application DEVRAIT appliquer la vérification suivante :

- o Vérification que la chaîne est conforme aux exigences des caractères de droite à gauche spécifiées dans le document Bidi [RFC5893].
Cet essai peut être omis dans des circonstances particulières, comme quand l'application de recherche sait que les conditions sont appliquées ailleurs, parce que la tentative de recherche et de résolution de telles chaînes va presque toujours conduire à un échec de la recherche dans le DNS sauf quand des caractères génériques sont présents dans la zone. Cependant, appliquer la vérification va probablement donner de bien meilleures informations sur la raison d'un échec de la recherche -- informations qui peuvent être utilement passées à l'utilisateur quand c'est faisable -- plutôt qu'une information d'échec de résolution du DNS seule.

Pour toutes les autres chaînes, l'application de recherche DOIT s'appuyer sur la présence ou l'absence d'étiquettes dans le DNS pour déterminer la validité de ces étiquettes et la validité des caractères qu'elles contiennent. Si elles sont enregistrées, elles sont présumées valides ; si elles ne le sont pas, leur possible validité n'est pas pertinente. Alors qu'une application de recherche peut raisonnablement produire des avertissements sur les chaînes qu'elle estime être problématiques, les applications qui refusent de traiter une chaîne qui se conforme aux règles ci-dessus (c'est-à-dire, qui ne cherche pas dans le DNS) ne sont pas en conformité avec ce protocole.

5.5 Conversion Punycode

La chaîne qui a maintenant été validée pour la recherche est convertie en forme ACE en appliquant l'algorithme Punycode à la chaîne et ensuite en ajoutant le préfixe ACE ("xn--").

5.6 Résolution de nom DNS

L'étiquette A qui résulte de la conversion du paragraphe 5.5 ou qui est fournie directement (voir au paragraphe 5.3) est combinée avec les autres étiquettes comme nécessaire pour former un nom de domaine pleinement qualifié qui va être cherché dans le DNS, en utilisant les procédures normales de résolveur du DNS. La recherche peut évidemment soit réussir (en retournant les informations) soit échouer.

6. Considérations sur la sécurité

Les considérations sur la sécurité pour cette version de IDNA sont décrites dans le document des définitions [RFC5890], sauf les questions particulières associées aux scripts et caractères de droite à gauche. Ces dernières sont discutées dans le document Bidi [RFC5893].

Afin d'éviter des attaques intentionnelles ou accidentelles à partir d'étiquettes qui pourraient être confondues avec d'autres, des problèmes particuliers de rendu, et ainsi de suite, le modèle IDNA exige que les registres fassent preuve d'attention et de pénétration quant aux étiquettes qu'ils choisissent de permettre. Ce problème est discuté au paragraphe 4.3 du présent document qui, à son tour, pointe sur une discussion un peu plus développée dans le document des raisons [RFC5894].

7. Considérations relatives à l'IANA

Les actions de l'IANA pour cette version d'IDNA sont spécifiées dans le document des tableaux [RFC5892] et discutées de façon informelle dans le document des raisons [RFC5894]. Les composants de IDNA décrits dans le présent document n'exigent aucune action de l'IANA.

8. Contributeurs

Bien que l'éditeur mentionné ait tenu la plume, les versions originales de ce document représentent le travail conjoint et les conclusions d'une équipe de conception ad hoc consistant en l'éditeur et, par ordre alphabétique, Harald Alvestrand, Tina Dam, Patrik Falstrom, et Cary Karp. Le présent document a fait des emprunts significatifs à la version originale de IDNA [RFC3490] aussi bien pour les concepts que pour le texte spécifique. Cette version de seconde génération n'aurait pas été possible sans le travail fait sur la première version et en particulier les contributions de ses auteurs Patrik Falstrom, Paul Hoffman, et Adam Costello. Alors que Falstrom a été activement impliqué dans la création de cette version, Hoffman et Costello ne l'ont pas été et ne devraient pas être tenus pour responsables de ses erreurs ou omissions.

9. Remerciements

Cette révision de IDNA aurait été impossible sans l'expérience accumulée depuis la publication de la RFC 3490 et les commentaires et plaintes résultant de nombreuses personnes de l'IETF, ICANN, et autres communautés (trop nombreux pour en faire la liste). Elle n'aurait pas été non plus possible sans la RFC 3490 elle-même et les efforts du groupe de travail qui l'a définie. Ces personnes dont les contributions sont reconnues dans les [RFC3490], [RFC4690], et [RFC5894] ont été particulièrement importantes.

Des changements textuels spécifiques ont été incorporés dans le présent document après les suggestions des autres contributeurs, Stephane Bortzmeyer, Vint Cerf, Lisa Dusseault, Paul Hoffman, Kent Karlsson, James Mitchell, Erik van der Poel, Marcos Sanz, Andrew Sullivan, Wil Tan, Ken Whistler, Chris Wright, et autres participants et relecteurs du groupe de travail parmi lesquels Martin Duerst, James Mitchell, Subramanian Moonesamy, Peter Saint-Andre, Margaret Wasserman, et Dan Winship qui ont relevé des erreurs spécifiques et recommandé des corrections. Des remerciements tout particuliers sont dus à Paul Hoffman pour sa permission d'extraire des matériaux pour former la base de l'Appendice A d'un projet de document qu'il avait préparé.

10. Références

10.1 Références normatives

[RFC1034] P. Mockapetris, "Noms de domaines - [Concepts et facilités](#)", STD 13, novembre 1987. (MàJ par [RFC1101](#), [1183](#), [1348](#), [1876](#), [1982](#), [2065](#), [2181](#), [2308](#), [2535](#), [4033](#), [4034](#), [4035](#), [4343](#), [4035](#), [4592](#), [5936](#), [8020](#), [8482](#), [8767](#))

[RFC1035] P. Mockapetris, "Noms de domaines - [Mise en œuvre](#) et spécification", STD 13, novembre 1987. (MàJ par [RFC1101](#), [1183](#), [1348](#), [1876](#), [1982](#), [1995](#), [1996](#), [2065](#), [2136](#), [2181](#), [2137](#), [2308](#), [2535](#), [2673](#), [2845](#), [3425](#), [3658](#), [4033](#), [4034](#), [4035](#), [4343](#), [5936](#), [5966](#), [6604](#), [7766](#), [8482](#), [8767](#))

- [RFC2119] S. Bradner, "[Mots clés à utiliser](#) dans les RFC pour indiquer les niveaux d'exigence", BCP 14, mars 1997. (MàJ par [RFC8174](#))
- [RFC3492] A. Costello, "[Punycode : Codage Bootstring d'Unicode](#) pour les noms de domaine internationalisés dans les applications (IDNA)", mars 2003. (P.S.)
- [RFC5890] J. Klensin, "[Noms de domaine internationalisés pour les applications](#) (IDNA) : Définitions et cadre documentaire", août 2010. (Remplace [RFC3490](#)) (P.S.)
- [RFC5892] P. Faltstrom, "[Codets Unicode et noms de domaine](#) internationalisés pour les applications (IDNA)", août 2010. (P.S. ; MàJ par [RFC8753](#))
- [RFC5893] H. Alvestrand, C. Karp, "[Écritures de droite à gauche](#) pour les noms de domaine internationalisés pour les applications (IDNA)", août 2010. (P.S.)
- [UAX15] The Unicode Consortium, "Unicode Standard Annex #15: Unicode Normalization Forms, Revision 31", septembre 2009, < <http://www.unicode.org/reports/tr15/tr15-31.html> >.

10.2. Références pour information

- [ASCII] American National Standards Institute, "USA Code for Information Interchange", ANSI X3.4-1968. ANSI X3.4-1968 a été remplacé par des versions plus récentes avec de légères modifications, mais la version 1968 reste celle de référence pour l'Internet.
- [RFC2671] P. Vixie, "Mécanismes d'[extension pour le DNS](#) (EDNS0)", août 1999. (P.S.) (Remplacée par [RFC6891](#))
- [RFC3490] P. Faltstrom et autres, "Internationalisation des noms de domaine dans les applications (IDNA)", mars 2003. (Remplacée par les [RFC5890](#) et [5891](#), P.S.)
- [RFC3491] P. Hoffman et M. Blanchet, "[Nameprep : Profil Stringprep](#) pour les noms de domaine internationalisés (IDN)", mars 2003. (Remplacée par la [RFC5891](#), P.S.)
- [RFC3986] T. Berners-Lee, R. Fielding et L. Masinter, "[Identifiant de ressource uniforme](#) (URI) : Syntaxe générique", STD 66, janvier 2005.
- [RFC3987] M. Duerst et M. Suignard, "[Identifiant de ressource internationalisé](#) (IRI)", janvier 2005.
- [RFC4690] J. Klensin et autres, "Révisions et recommandations pour les noms de domaines internationalisés (IDN)", septembre 2006. (Information)
- [RFC4952] J. Klensin, Y. Ko, "Généralités et cadre pour la messagerie électronique internationalisée", juillet 2007. (Remplacée par la [RFC6530](#)) (Information)
- [RFC5894] J. Klensin, "[Noms de domaine internationalisés](#) pour les applications (IDNA) : Fondements, explication, et motivations", août 2010. (Information)
- [RFC5895] P. Resnick, P. Hoffman, "[Transposition de caractères](#) pour les noms de domaine internationalisés pour les applications (IDNA)", septembre 2010. (Information)
- [Unicode] The Unicode Consortium, "The Unicode Standard, Version 5.0", 2007. Boston, MA, USA: Addison-Wesley. ISBN 0-321-48091-0. Cette référence imprimée a maintenant été mise à jour en ligne pour refléter des codets supplémentaires. Pour les codets, la référence au moment de la publication de ce document est à Unicode 5.2.

Appendice A. Résumé des changements majeurs par rapport à IDNA2003

1. Mise à jour du jeu de caractères de base de Unicode 3.2 en une référence indépendante de la version Unicode.
2. Séparation des définitions des activités "d'enregistrement" et de "recherche".

3. Interdiction des caractères de symbole et de ponctuation sauf lorsque des exceptions particulières sont nécessaires.
4. Suppression des étapes de transposition et de normalisation du protocole pour les faire réaliser par les applications elles-mêmes, éventuellement de façon locale, avant d'invoquer le protocole.
5. Changement de la façon dont le protocole spécifie quels caractères sont permis dans les étiquettes depuis les "humains qui décident ce que contient le tableau des codets" jusqu'à la "décision sur les codets se fonde sur les propriétés d'Unicode plus une petite liste d'exclusions créée par des personnes".
6. Introduire le nouveau concept de caractères qui peuvent être utilisés seulement dans des contextes spécifiques.
7. Permet d'exprimer les mots et noms typiques dans des langues comme le Dhivehi et le Yiddish.
8. Rend l'affichage des noms de domaine bidirectionnels (chaînes délimitées d'étiquettes, pas seulement des étiquettes autonomes) moins surprenant, qu'ils apparaissent dans des contextes de noms de domaine évidents ou au titre du texte qui court dans des paragraphes.
9. Suppression du point séparateur de la partie obligatoire du protocole.
10. Rendues invalides certaines étiquettes actuellement valides qui ne sont en fait pas des étiquettes IDNA.

Adresse de l'auteur

John C Klensin
1770 Massachusetts Ave, Ste 322
Cambridge, MA 02140
USA
téléphone : +1 617 245 1457
mél : john+ietf@jck.com