

Groupe de travail Réseau
Request for Comments : 4518
Catégorie : Norme

K. Zeilenga
OpenLDAP Foundation
juin 2006

Protocole léger d'accès à un répertoire (LDAP) : Préparation de chaîne internationalisée

Statut de ce mémo

Le présent document spécifie un protocole de normalisation Internet pour la communauté de l'Internet, qui appelle à la discussion et à des suggestions pour son amélioration. Prière de se reporter à l'édition en cours des "Normes de protocole officielles de l'Internet" (STD 1) sur l'état de la normalisation et le statut de ce protocole. La distribution du présent mémo n'est soumise à aucune restriction.

Notice de Copyright

Copyright (C) The Internet Society (2006).

Résumé

Les précédentes spécifications techniques du protocole léger d'accès à un répertoire (LDAP) ne définissaient pas de façon précise comment on doit effectuer la confrontation des chaînes de caractères. Ceci a conduit à un certain nombre de problèmes d'utilisation et d'interopérabilité. Le présent document définit des algorithmes de préparation de chaînes pour les règles de confrontation fondées sur le caractère définies pour être utilisées dans LDAP.

1 Introduction

1.1 Fondements

Une règle de correspondance [RFC4517] du protocole léger d'accès à un répertoire (LDAP) [RFC4510] définit un algorithme pour déterminer si une valeur présentée correspond à une valeur d'attribut conformément au critère défini pour la règle. La proposition peut être évaluée comme True (*vrai*), False (*faux*), ou Undefined (*indéfini*).

True - l'attribut contient une valeur correspondante,

False - l'attribut ne contient pas de valeur correspondante,

Undefined - il ne peut pas être déterminé si l'attribut contient une valeur correspondante.

Par exemple, la règle caseIgnoreMatch peut être utilisée pour comparer si l'attribut commonName contient une valeur particulière sans considération de la casse et des espaces non significatifs.

1.2 Règles de correspondance de chaîne X.500

"X.520: Types d'attributs choisis" [X.520] fournit (entre autres choses) des syntaxes de valeur et des règles de correspondance pour comparer les valeurs couramment utilisées dans l'annuaire [X.500]. Ces spécifications sont inadéquates pour les chaînes composées de caractères Unicode [Unicode].

La règle de correspondance caseIgnoreMatch [X.520], par exemple, est simplement définie comme une comparaison insensible à la casse où les espaces non significatifs sont ignorés. Pour printableString, il y a seulement un caractère espace et la transposition de la casse est bijective, et donc, cette définition est suffisante. Cependant, pour les types de chaînes Unicode, telles que universalString, ceci n'est pas suffisant. Par exemple, une mise en œuvre de correspondance insensible à la casse qui ramènerait des caractères minuscules en majuscules donnerait des résultats différents d'une mise en œuvre qui ramène les majuscules en minuscules. Ou bien une mise en œuvre peut voir les espaces comme se référant au seul SPACE (U+0020), et une seconde mise en œuvre verra tout caractère avec la propriété de séparateur d'espace (Zs) comme un espace, et une autre mise en œuvre peut voir tout caractère avec la catégorie espace blanc (WS) comme un espace.

L'absence de spécification précise pour la correspondance de chaînes de caractères a conduit à des problèmes significatifs d'interopérabilité. Lorsque c'est utilisé pour la validation de chaînes de certificats, il peut en résulter un affaiblissement de la sécurité. Pour résoudre ces problèmes, le présent document définit des algorithmes précis pour la préparation des chaînes de caractères à la mise en correspondance.

1.3 Relations avec "stringprep"

Les algorithmes de préparation de chaînes de caractères décrits dans le présent document se fondent sur l'approche "stringprep" [RFC3454]. Dans "stringprep", les valeurs présentées et mémorisées sont d'abord préparées pour la comparaison de sorte qu'une comparaison caractère par caractère donne le résultat "correct".

L'approche utilisée ici est une amélioration de l'approche "stringprep" [RFC3454]. Chaque algorithme implique deux étapes de préparation supplémentaires.

- a) Avant d'appliquer les étapes de préparation de chaîne de Unicode exposées dans "stringprep", la chaîne est transcodée en Unicode.
- b) Après avoir appliqué les étapes Unicode de préparation de chaîne exposées dans "stringprep", la chaîne est modifiée pour traiter de façon appropriée les caractères non significatifs pour la règle de correspondance.

Et donc, la préparation des chaînes de caractères pour la correspondance de X.500 [X.500] [X.501] implique les étapes suivantes :

- 1) Transcoder
- 2) Transposer
- 3) Normaliser
- 4) Interdire
- 5) Chercher les caractères bidirectionnels
- 6) Traiter les caractères non significatifs

Ces étapes sont décrites à la Section 2.

Il faut noter qu'alors que divers tableaux de caractères Unicode inclus ou référencés par la présente spécification sont dérivés des données Unicode [Unicode], ces tableaux sont à considérer comme définitifs pour les besoins de la mise en œuvre de la présente spécification.

1.4 Relations avec la spécification technique LDAP

Le présent document fait partie intégrante de la spécification technique LDAP [RFC4510], qui rend obsolète dans sa totalité la spécification technique LDAP [RFC3377] précédemment définie.

Le présent document expose de nouveaux algorithmes LDAP de préparation de chaîne de caractères internationalisée utilisés par la [RFC4517] et d'autres spécifications techniques possibles définissant les syntaxes et/ou règles de correspondance LDAP.

1.5 Relations avec X.500

LDAP est défini [RFC4510] dans les termes de X.500 comme un mécanisme d'accès X.500. Comme tel, il y a un fort désir d'alignement entre la syntaxe et la sémantique de LDAP et de X.500. Les algorithmes de préparation de chaînes de caractères décrites dans le présent document se fondent sur la proposition du groupe d'étude conjoint UIT/ISO n° 2 "Règles de correspondance de chaîne internationalisée pour X.500" [XMATCH].

1.6 Conventions et termes

Les mots clé "DOIT", "NE DOIT PAS", "EXIGE", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMANDÉ", "PEUT", et "FACULTATIF" dans le présent document sont à interpréter comme décrit dans le BCP 14 [RFC2119].

Dans le présent document, les noms de caractères utilisent la notation de la norme Unicode [Unicode] pour les noms et les séquences codées. Par exemple, la lettre "a" peut être représentée soit par <U+0061> soit par <LATIN SMALL LETTER A>. Dans les listes des transpositions et des caractères interdits, le "U+" est laissé de côté pour rendre les

listes plus lisibles. Les commentaires pour les gammes de caractères sont donnés entre crochets (comme "[CARACTÈRES DE CONTRÔLE]") et ne font pas partie de la norme.

Note : un glossaire des termes utilisés dans Unicode figure dans [Glossary]. On peut trouver des informations sur le modèle de codage des caractères Unicode dans [CharModel].

Le terme "marque de couplage", tel qu'utilisé dans la présente spécification, se réfère à toute séquence codée Unicode [Unicode] qui a une propriété de marquage (Mn, Mc, Me). L'Appendice A donne une liste définitive des marques de couplage.

2 Préparation de chaîne

Le processus en six étapes suivant DOIT être appliqué à chaque valeur d'attribut présentée et en préparation pour l'évaluation de règle de correspondance de chaîne de caractères.

- 1) Transcodage
- 2) Transposition
- 3) Normalisation
- 4) Interdiction
- 5) Recherche de caractères bidirectionnels
- 6) Traitement des caractères non signifiants

L'échec d'une seule étape cause l'évaluation de l'assertion à Indéfini.

Le répertoire de caractères de ce processus est Unicode 3.2 [Unicode].

Noter que cette spécification de processus en six étapes est destinée à décrire un comportement de correspondance attendue. Les mises en œuvre ont toute liberté pour utiliser des processus de remplacement pour autant que le comportement d'évaluation de règle de correspondance soit cohérent avec le comportement décrit par la présente spécification.

2.1 Transcodage

Chaque valeur de chaîne non Unicode est transcodée en Unicode.

Les valeurs de PrintableString [X.680] sont transcodées directement en Unicode.

Les valeurs UniversalString, UTF8String, et bmpString [X.680] n'ont pas besoin d'être transcodées car elles sont des chaînes fondées sur Unicode (dans le cas de bmpString, un sous-ensemble de Unicode).

Les valeurs TeletexString [X.680] sont transcodées en Unicode. Comme il n'y a pas de norme pour la transposition des valeurs TeletexString en Unicode, la transposition est laissée aux organes locaux.

Pour ces raisons et quelques autres, l'utilisation de TeletexString est NON RECOMMANDÉE.

Le résultat est la chaîne transcodée.

2.2 Transposition

Les séquences codées SOFT HYPHEN (U+00AD) et MONGOLIAN TODO SOFT HYPHEN (U+1806) ne sont pas transposées. Les séquences codées JOINER (U+034F) et VARIATION SELECTOR (U+180B-180D, FF00-FE0F) ne sont pas transposées non plus. Le caractère OBJECT REPLACEMENT CHARACTER (U+FFFC) n'est pas transposé.

Les caractères CHARACTER TABULATION (U+0009), LINE FEED (LF) (U+000A), LINE TABULATION (U+000B), FORM FEED (FF) (U+000C), CARRIAGE RETURN (CR) (U+000D), et NEXT LINE (NEL) (U+0085) sont transposés en SPACE (U+0020).

Toutes les autres séquences codées de contrôle (par exemple, Cc) ou séquences codées avec une fonction de contrôle (par exemple, Cf) ne sont pas transposées. Ci-après figure une liste complète de ces séquences codées : U+0000-0008,

000E-001F, 007F-0084, 0086-009F, 06DD, 070F, 180E, 200C-200F, 202A-202E, 2060-2063, 206A-206F, FEFF, FFF9-FFFB, 1D173-1D17A, E0001, E0020-E007F.

ZERO WIDTH SPACE (U+200B) n'est pas transposé. Toutes les autres séquences codées avec une propriété de séparateur (espace, ligne, ou paragraphe) (par exemple, Zs, Zl, ou Zp) sont transposées en SPACE (U+0020). Ci-après figure une liste complète de ces séquences codées : U+0020, 00A0, 1680, 2000-200A, 2028-2029, 202F, 205F, 3000.

Pour les règles de correspondance de chaîne case ignore (*ignorer la casse*), numeric (*numérique*), et stored prefix (*préfixe mémorisé*), les caractères sont changés de casse conformément au paragraphe B.2 de la [RFC3454].

Le résultat est la chaîne transposée.

2.3 Normalisation

La chaîne d'entrée est à normaliser en forme Unicode KC (couplage de compatibilité) comme décrit dans [UAX15]. Le résultat est la chaîne normalisée.

2.4 Interdiction

Toutes les séquences codées non allouées sont interdites. La liste des séquences codées non allouées figure dans le Tableau A.1 de la [RFC3454].

Les caractères qui, selon le paragraphe 5.8 de la [RFC3454], changent les propriétés d'affichage ou sont déconseillés sont interdits. La liste de ces caractères figure au Tableau C.8 de la [RFC3454].

Les séquences codées à usage privé (*Private Use*) sont interdites. La liste de ces caractères figure au Tableau C.3 de la [RFC3454].

Toutes les séquences codées de non caractère sont interdites. La liste de ces séquences codées figure au Tableau C.4 de la [RFC3454].

Les codes de substitution sont interdits. La liste de ces caractères figure au Tableau C.5 de la [RFC3454].

La séquence codée REPLACEMENT CHARACTER (U+FFFD) (*caractère de remplacement*) est interdite.

Cette étape échoue si la chaîne d'entrée contient une séquence codée interdite. Autrement, la sortie est la chaîne d'entrée.

2.5 Recherche de caractères bidirectionnels

Les caractères bidirectionnels sont ignorés.

2.6 Traitement des caractères non significatifs

Dans cette étape, la chaîne est modifiée pour garantir un traitement approprié des caractères non significatifs dans la règle de correspondance. Cette modification diffère selon la règle de correspondance.

Le paragraphe 2.6.1 s'applique à la correspondance de chaîne case ignore et exact.

Le paragraphe 2.6.2 s'applique à la correspondance numericString.

Le paragraphe 2.6.3 s'applique à la correspondance telephoneNumber.

2.6.1 Traitement de l'espace non signifiant

Pour les besoins de ce paragraphe, un espace est défini comme étant la séquence codée SPACE (U+0020) non suivie de signe de couplage.

NOTE – L'étape précédente assure que la chaîne ne peut pas contenir de séquence codée autres que SPACE (U+0020) dans la classe des séparateurs.

Pour les chaînes d'entrée qui sont des valeurs d'attribut ou des valeurs d'assertion qui ne sont pas des sous-chaînes : si la chaîne d'entrée ne contient pas de caractère non espace, la sortie est exactement deux SPACE. Autrement (la chaîne d'entrée contient au moins un caractère non espace), la chaîne est modifiée de telle sorte que la chaîne commence par exactement un caractère espace, se termine par exactement un caractère SPACE, et toute séquence interne (non vide) de caractères espace est remplacée par exactement deux caractères SPACE. Par exemple, la chaîne d'entrée "foo<SPACE>bar<SPACE><SPACE>", donne le résultat "<SPACE>foo<SPACE><SPACE>bar<SPACE>".

Pour les chaînes d'entrée qui sont des valeurs d'assertion de sous-chaîne : si la chaîne en cours de préparation ne contient pas de caractère non espace, la chaîne de sortie est exactement un SPACE. Autrement, on suivra les étapes ci-après :

- Si la chaîne d'entrée est une sous-chaîne initiale, elle est modifiée pour commencer par exactement un caractère SPACE ;
- Si la chaîne d'entrée est une sous-chaîne initiale ou n'importe quelle sous-chaîne qui se termine par un ou plusieurs caractères espace, elle est modifiée pour se terminer par exactement un caractère SPACE ;
- Si la chaîne d'entrée est une sous-chaîne quelconque ou finale qui commence par un ou plusieurs caractères espace, elle est modifiée pour commencer par exactement un caractère SPACE ; et
- Si la chaîne d'entrée est une sous-chaîne finale, elle est modifiée pour se terminer par exactement un caractère SPACE.

Par exemple, pour la chaîne d'entrée "foo<SPACE>bar<SPACE><SPACE>" comme sous-chaîne initiale, la sortie devrait être "<SPACE>foo<SPACE><SPACE>bar<SPACE>". Comme sous-chaîne quelconque ou finale, la même entrée aurait pour résultat "foo<SPACE>bar<SPACE>".

L'Appendice B expose les raisons de ce comportement.

2.6.2 Traitement de caractère non signifiant numericString

Dans ce paragraphe, un espace est défini comme étant une séquence codée SPACE (U+0020) suivie d'aucun signe de couplage.

Tous les espaces sont considérés comme non signifiants et sont à retirer.

Par exemple, retrait des espaces de la chaîne de forme KC :

"<SPACE><SPACE>123<SPACE><SPACE>456<SPACE><SPACE>" résulterait en la chaîne de sortie : "123456" et la chaîne de forme KC : "<SPACE><SPACE><SPACE>" donnerait la chaîne de sortie : "" (une chaîne vide).

2.6.3 Traitement de caractère non signifiant telephoneNumber

Dans ce paragraphe, un trait d'union est défini comme étant une séquence codée HYPHEN-MINUS (U+002D), ARMENIAN HYPHEN (U+058A), HYPHEN (U+2010), NON-BREAKING HYPHEN (U+2011), MINUS SIGN (U+2212), SMALL HYPHEN-MINUS (U+FE63), ou FULLWIDTH HYPHEN-MINUS (U+FF0D) suivi d'aucun signe de couplage, et un espace est défini comme une séquence codée SPACE (U+0020) suivi d'aucun signe de couplage.

Tous les traits d'union et les espaces sont considérés comme non signifiants et sont à retirer.

Par exemple, retrait des traits d'union et des espaces de la chaîne de forme KC :

"<SPACE><HYPHEN>123<SPACE><SPACE>456<SPACE><HYPHEN>" donnerait la chaîne de sortie : "123456" et la chaîne de forme KC : "<HYPHEN><HYPHEN><HYPHEN>" donnerait la chaîne de sortie (vide) : "".

3 Considérations sur la sécurité

Les considérations sur la sécurité de "Préparation des chaînes internationalisée ("stringprep")" [RFC3454] s'appliquent de façon générale aux algorithmes décrits ici.

4 Remerciements

L'approche utilisée dans le présent document se fonde sur des principes et algorithmes de conception décrits dans "Préparation des chaînes internationalisée ("stringprep")" [RFC3454] par Paul Hoffman et Marc Blanchet. Certaines lignes directrices supplémentaires ont été tirées des normes techniques Unicode, des Rapports techniques, et des Notes.

Le présent document a été produit par le groupe de travail Révision LDAP (LDAPBIS) de l'IETF.

5 Références

5.1 Références normatives

[RFC2119] Bradner, S., "Mots clé à utiliser dans les RFC pour indiquer les niveaux d'exigence", BCP 14, RFC 2119, mars 1997.

[RFC3454] Hoffman, P. et. Blanchet, "Préparation des chaînes internationalisées ("stringprep")", RFC 3454, décembre 2002.

[RFC4510] Zeilenga, K., Ed., "Protocole léger d'accès à un répertoire (LDAP) : Descriptif des spécifications techniques", RFC 4510, juin 2006.

[RFC4517] Legg, S., Ed., "Protocole léger d'accès à un répertoire (LDAP): Syntaxes et règles de correspondance", RFC 4517, juin 2006.

[Unicode] The Unicode Consortium, "Norme Unicode, Version 3.2.0" est définie par "Norme Unicode, Version 3.0" (Reading, MA, Addison-Wesley, 2000. ISBN 0-201-61633-5), telle qu'amendée par "Norme Unicode, Annexe n°27: Unicode 3.1" (<http://www.unicode.org/reports/tr27/>) et par "Norme Unicode, Annexe n°28: Unicode 3.2" (<http://www.unicode.org/reports/tr28/>).

[UAX15] Davis, M. et M. Duerst, "Norme Unicode, Annexe n°15: Formes de normalisation Unicode, Version 3.2.0". <<http://www.unicode.org/unicode/reports/tr15/tr15-22.html>>, mars 2002.

[X.680] Union Internationale des Télécommunications – Secteur de la normalisation des Télécommunications, "Notation de syntaxe abstraite n° 1 (ASN.1) - Spécification de la notation de base", X.680(2002) (aussi ISO/CEI 8824-1:2002).

5.2 Références informatives

[X.500] Union Internationale des Télécommunications – Secteur de la normalisation des Télécommunications, "L'annuaire – Généralités sur les concepts, modèles et services," X.500 (1993) (aussi ISO/CEI 9594-1:1994).

[X.501] Union Internationale des Télécommunications – Secteur de la normalisation des Télécommunications, "L'annuaire -- Modèles," X.501 (1993) (aussi ISO/CEI 9594- 2:1994).

[X.520] Union Internationale des Télécommunications – Secteur de la normalisation des Télécommunications, "L'annuaire : Types d'attributs choisis", X.520 (1993) (aussi ISO/CEI 9594-6:1994).

[Glossary] The Unicode Consortium, "Glossaire Unicode", <<http://www.unicode.org/glossary/>>.

[CharModel] Whistler, K. et M. Davis, "Rapport technique Unicode n° 17, Modèle de codage de caractères", UTR17, <<http://www.unicode.org/unicode/reports/tr17/>>, août 2000.

[RFC3377] Hodges, J. et R. Morgan, "Protocole léger d'accès à un répertoire (v3) : Spécification technique", RFC 3377, septembre 2002.

[RFC4515] Smith, M., Ed. et T. Howes, "Protocole léger d'accès à un répertoire (LDAP) : Représentation de chaîne des filtres de recherche", RFC 4515, juin 2006.

[XMATCH] Zeilenga, K., "Règles de correspondance de chaîne internationalisée pour X.500", Travail en cours.

Appendice A Signes de couplage

Le présent appendice est normatif.

Le présent tableau est tiré des fichiers de données Unicode [Unicode] ; il fait la liste de toutes les séquences codées qui ont la propriété Mn, Mc, ou Me. Ce tableau est à considérer comme définitif pour les besoins de la mise en œuvre de la présente spécification.

```
0300-034F 0360-036F 0483-0486 0488-0489 0591-05A1
05A3-05B9 05BB-05BC 05BF 05C1-05C2 05C4 064B-0655 0670
06D6-06DC 06DE-06E4 06E7-06E8 06EA-06ED 0711 0730-074A
07A6-07B0 0901-0903 093C 093E-094F 0951-0954 0962-0963
0981-0983 09BC 09BE-09C4 09C7-09C8 09CB-09CD 09D7
09E2-09E3 0A02 0A3C 0A3E-0A42 0A47-0A48 0A4B-0A4D
0A70-0A71 0A81-0A83 0ABC 0ABE-0AC5 0AC7-0AC9 0ACB-0ACD
0B01-0B03 0B3C 0B3E-0B43 0B47-0B48 0B4B-0B4D 0B56-0B57
0B82 0BBE-0BC2 0BC6-0BC8 0BCA-0BCD 0BD7 0C01-0C03
0C3E-0C44 0C46-0C48 0C4A-0C4D 0C55-0C56 0C82-0C83
0CBE-0CC4 0CC6-0CC8 0CCA-0CCD 0CD5-0CD6 0D02-0D03
0D3E-0D43 0D46-0D48 0D4A-0D4D 0D57 0D82-0D83 0DCA
0DCF-0DD4 0DD6 0DD8-0DDF 0DF2-0DF3 0E31 0E34-0E3A
0E47-0E4E 0EB1 0EB4-0EB9 0EBB-0EBC 0EC8-0ECD 0F18-0F19
0F35 0F37 0F39 0F3E-0F3F 0F71-0F84 0F86-0F87 0F90-0F97
0F99-0FBC 0FC6 102C-1032 1036-1039 1056-1059 1712-1714
1732-1734 1752-1753 1772-1773 17B4-17D3 180B-180D 18A9
20D0-20EA 302A-302F 3099-309A FB1E FE00-FE0F FE20-FE23
1D165-1D169 1D16D-1D172 1D17B-1D182 1D185-1D18B
1D1AA-1D1AD
```

Appendice B Correspondances de sous-chaînes

Le présent appendice n'est pas normatif.

En l'absence de correspondance de sous-chaîne, le traitement de l'espace non significatif pourrait être simplifié pour la correspondance case ignore/exact. En particulier, le traitement pourrait être d'exiger que toutes les séquences d'un ou plusieurs espaces soient remplacées par un espace et, si la chaîne contient des caractères non espace, le retrait de tous les espaces en tête et en fin.

En présence de correspondance de sous-chaînes, ce traitement simplifié des espaces conduirait à un comportement de correspondance inattendu et indésirable. Par exemple :

- 1) (CN=foo\20*\20bar) correspondrait à la valeur CN "foobar" ;
- 2) (CN=*\20foobar\20*) correspondrait à "foobar", mais pas (CN=*\20*foobar*\20*).

Note pour les lecteurs qui ne sont pas familiarisés avec la correspondance de sous-chaînes LDAP : L'assertion de filtre LDAP [RFC4515] (CN=A*B*C) dit "faire correspondre chaque valeur (de l'attribut CN) qui commence par A, contient B après A, se termine par C lorsque C est aussi après B."

Le premier cas illustre que ce traitement d'espace simplifié amènerait à considérer comme non significatifs les espaces en tête et en fin dans les sous-chaînes de la chaîne. Cependant, seuls les espaces en tête et en fin (ainsi que les espaces multiples consécutifs) de la chaîne (comme un tout) sont non significatifs.

Le second cas illustre que ce traitement simplifié des espaces pourrait causer des échecs de sous-partitionnement. C'est-à-dire, si une sous-chaîne préparée quelconque correspond à une partition de la valeur de l'attribut, une assertion construite en subdivisant cette sous-chaîne en plusieurs sous-chaînes devrait aussi correspondre.

Pour concevoir une approche appropriée du traitement des espaces pour la correspondance de sous-chaîne, on doit étudier les aspects clés de la correspondance case exact/ignore de X.500. X.520 [X.520] dit :

La règle [substrings] retourne VRAI s'il y a une partition de la valeur d'attribut (en portions) telle que :

- les sous-chaînes spécifiées (initiale, quelconque, finale) corresponde aux différentes portions de la valeur dans l'ordre de séquence de la chaîne ;
- l'initiale, s'il en est, corresponde à la première portion de la valeur ;
- la finale, s'il en est, corresponde à la dernière portion de la valeur ;
- un quelconque, s'il en est, corresponde à toute portion arbitraire de la valeur.

C'est à dire que l'assertion (CN=foo\20*\20bar) de sous-chaîne corresponde à la valeur d'attribut "foo<SPACE><SPACE>bar" lorsque la valeur peut être partagée en portions "foo<SPACE>" et "<SPACE>bar" qui satisfont aux exigences ci-dessus.

X.520 dit aussi :

[L]es espaces suivants sont considérés comme non significatifs :

- les espaces en tête (c'est-à-dire, ceux qui précèdent le premier caractère qui n'est pas un espace) ;
- les espaces (c'est-à-dire, ceux qui suivent le dernier caractère qui n'est pas un espace) ;
- plusieurs espaces consécutifs (qui sont tenus pour équivalents à un seul caractère espace).

Cette déclaration s'applique aux valeurs d'assertion et aux valeurs d'attribut comme chaînes entières, et non individuellement aux sous-chaînes d'une valeur d'assertion. En particulier, ces déclarations devraient être comprises comme signifiant que si une valeur d'assertion et une valeur d'attribut correspondent sans aucune considération de caractères non significatifs, la valeur d'assertion devrait alors aussi correspondre à toute valeur d'attribut qui diffère seulement par inclusion ou retrait de caractères non significatifs.

Et donc l'assertion (CN=foo\20*\20bar) correspond à "foo<SPACE><SPACE><SPACE>bar" et "foo<SPACE>bar" car ces valeurs ne diffèrent de "foo<SPACE><SPACE>bar" que par l'inclusion ou le retrait d'espaces non significatifs.

Le lecteur astucieux du présent document aura aussi noté qu'il y a des cas particuliers dans lesquels le traitement d'espace spécifié n'ignore pas des espaces qui pourraient être considéré comme insignifiants. Par exemple, l'assertion (CN=\20*\20*\20) ne correspond pas à "<SPACE><SPACE><SPACE>" (espaces non significatifs présents dans la valeur) ou " " (espaces non significatifs non présents dans la valeur). Cependant, comme ces cas n'ont pas d'application pratique, ils ne peuvent pas être satisfaits par des assertions simples, par exemple, (cn=\20), et cette anomalie mineure qui peut seulement être réglée par un algorithme de préparation à utiliser conjointement avec un partitionnement et une correspondance caractère par caractère, est considérée comme acceptable.

Adresse de l'auteur

Kurt D. Zeilenga
OpenLDAP Foundation
EMail: Kurt@OpenLDAP.org

Déclaration de copyright

Copyright (C) The Internet Society (2006).

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

Le présent document et les informations y contenues sont fournies 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 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 CI-ENCLOSES NE VIOLENT 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 pourrait 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 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.

RemerciementLe financement de la fonction d'édition des RFC est fourni par la Administrative Support Activity (IASA) de l'IETF.