

Nouveau programme gTLD : Variantes IDN

Février 2010

Introduction

Les communautés de langage qui songent à utiliser des IDN et des caractères variants sont concernées par la gestion et l'implémentation de variants au sein des nouveaux TLDs. Cela touche à l'implémentation des gTLD et des ccTLD. L'objectif est de déployer un panel de procédures consistant concernant le développement et l'utilisation des tables IDN, et pour l'allocation et la gestion de TLD variants.

Définitions :

1. Une table IDN fournit la liste des caractères disponibles pour l'enregistrement de noms de domaines en accord avec la politique du registre, et contient tous les caractères variants.
2. Les caractères variants sont des caractères qui peuvent être représentés de plusieurs façons (qui risquent de semer la confusion par leurs similitudes).
3. Les TLD variants sont identiques entre eux à l'exception du fait que des caractères variants sont inter changés au sein de la chaîne de caractères.

La gestion des tables IDN et des caractères variants pour les chaînes de premier niveau a été débattu au sein de la communauté dans de multiples contextes. Un groupe de travail indépendant concentré sur l'implémentation a été créé pour examiner ces questions après les discussions ayant eu lieu lors des réunions de l'ICANN à Mexico et à Sydney en 2009. Une partie des responsabilités du groupe de travail reposait sur l'étude de la gestion des variants au sein des chaînes de caractères TLD et la proposition d'une possible solution. L'équipe intégrait des experts techniques et linguistiques venus de différentes communautés de langage, et était codirigé par deux membres du comité exécutif de l'ICANN spécialisés dans les questions d'IDN et de DNS.

La recommandation du groupe de travail fut de ne pas associer de variants aux TLD pour le moment, mais de tester certains mécanismes pour permettre l'utilisation future de TLD variants. Il était spécifiquement recommandé que le mécanisme d'association de variants DNAME soit systématiquement testé pour arriver à la formulation d'une solution appropriée et assurer ainsi une certaine consistance lors du déploiement de variants au sein des DNS. La conclusion de l'équipe statuait que, si les variants seront bien attribués dans le futur, certaines conditions doivent être remplies, selon un arrangement (i.e. un accord) entre l'ICANN et le postulant. Par ailleurs, une proposition d'enregistrement d'une nouvelle ressource (appelée BNAME) a été réalisée à travers des ébauches Internet au sein de l'IETF, en tant que solution au problème de gestion des TLD variants. L'ICANN testera DNAME et BNAME, ainsi que toute autre solution viable, et comparera les résultats.

Les recommandations suggéraient également qu'un postulant à un TLD doit prouver que toute chaîne de caractères variants pour les TLD est identifiable au sein de la table IDN correspondante, et si le postulant souhaite utiliser ces variants, ils lui seront attribués de manière temporaire jusqu'à ce qu'un mécanisme d'application soit défini. Il était également recommandé que le postulant doive présenter les variants "indésirables" devant être bloqués de manière préventive afin d'éviter qu'ils ne lui soient attribués pour ces chaînes.

Le rapport du groupe de travail sur l'implémentation est disponible à <http://www.icann.org/en/announcements/announcement-2-03dec09-en.htm>.

Pour voir davantage d'information sur ce sujet, aller sur <http://www.icann.org/en/topics/idn/fast-track/proposed-implementation-details-idn-tables-revision-1-clean-29may09-en.pdf>

Conditions préalables au débat

A ce jour, les ébauches du Manuel du postulant gTLD n'évoquent pas la possibilité de chaînes de caractères IDN variants de premier niveau. Le texte proposé permet la future possible attribution de TLD variants en cas d'achèvement d'un mécanisme et d'un processus encore à développer.

Un éventuel texte pour un manuel basé sur les recommandations du groupe de travail est inclus plus bas pour aider à nourrir la discussion. La section appropriée apparaît dans le module 1 du manuel; voir le module entier sur <http://icann.org/en/topics/new-gtlds/intro-04oct09-en.pdf> Le premier module fournit une introduction à l'application et au processus d'évaluation des gTLD, incluant les renseignements sur les conditions spécifiques nécessaires aux postulants aux gTLD IDN.

A travers l'approche proposée ici, l'ICANN collecte les listes des chaînes de caractères TLD variants pour tous les nouveaux postulants gTLD. Pour le processus de postulation aux nouveau gTLD (comme pour le Fast Track ccTLD IDN), les tables IDN doivent être fournies avec le dossier de postulation pour la chaîne de caractère associée. Les chaînes de caractères variants sont générées par la table IDN du postulant, pour la langue adéquate.

Au moment de postuler, le postulant doit indiquer quels sont les variants désirés (i.e. les chaînes de caractères que le postulant souhaite s'approprier en tant que TLD une fois qu'un processus sera développé), et quels variants ne sont pas désirés (i.e. les chaînes de caractères que le postulant ne souhaite pas user, mais qui devraient être bloquées). Cependant, aucun variant n'est attribué dans la zone racine par le biais d'une postulation. Les variants ne peuvent être attribués que lorsqu'un processus développé et testé existe, et lorsque le postulant a démontré qu'il pouvait développer ces mécanismes selon les normes en vigueur. Ces mesures sont similaires à celles prises concernant le Fast Track ccTLD IDN. Pour le Fast Track, les demandes doivent indiquer les chaînes de caractères variants extraits de la table IDN correspondante et indiquer les chaînes désirées. Comme expliqué dans le Manuel des demandeurs pour les participants Fast Track ccTLD IDN, ceci ne signifie pas que le TLD variant désiré soit attribué à la zone racine des DNS.

L'approche d'implémentation discutée ici rassemble les renseignements des variants à travers les postulants et permet que des détails supplémentaires soient éclaircis le long du processus dans ce domaine.

L'ICANN encourage les commentaires sur les formules fournies dans ce document. Ce lexique n'est là que pour le débat, et n'a pas pour le moment été incorporé dans le Manuel du postulant. Les commentaires seront pris en compte pour la version numéro 4 de l'ébauche finale du guide du postulant, prévu pour être publié en juin 2010.

1.3 Information pour les postulants aux noms de domaines internationalisés

Certaines chaînes de caractères gTLD ayant fait l'objet d'une demande devraient devenir des Noms de domaine internationalisés (IDN) requièrent l'insertion d'un label-A encodé IDN dans la zone racine des DNS. Les IDN sont des noms de domaines incluant des caractères usés dans la représentation locale des langues non-écrites avec l'alphabet latin classique (a - z), les chiffres Européo-Arabes (0-9) et le tiret (-).

1.3.1 Conditions spécifiques aux IDN

Un postulant à une chaîne IDN doit fournir certaines informations montrant la conformité avec le protocole IDNA et d'autres conditions. Le protocole IDNA peut être trouvé sur <http://icann.org/en/topics/idn/rfcs.htm>

Les postulants doivent fournir les chaînes gTLD demandées sous la forme d'un **label-U** et un **label-A**.

Un label-A est la forme ASCII d'un label IDN. Chaque label-A commence par le préfixe IDNA ACE, "xn--", suivi d'une chaîne qui est le produit d'un algorithme Punycode, d'une longueur maximale de 59 caractères ASCII. Le préfixe et la chaîne réunis doivent être conformes à toutes les conditions nécessaires à un label qui pourrait se retrouver dans le DNS, ceci incluant la conformité avec la règle du LDH (nom d'hôte) décrite dans le RFC 1034, RFC 1123, et ailleurs.

Le label-U est la forme Unicode d'un label IDN, celle qu'un usager s'attend à voir s'afficher.

En prenant comme exemple la chaîne de test IDN actuelle en script cyrillique, le label-U est <испытание> et le label-A est <xn--80akhbyknj4f>. Un label-A doit pouvoir être produit par la conversion depuis un label-U, tout comme un label-U doit pouvoir être produit par un label-A.

Les postulants aux gTLD IDN devront également fournir les informations suivantes au moment de leur demande:

1. Version courte de la chaîne (en Anglais). Le postulant devra fournir une courte description de ce que la chaîne représenterait ou signifierait en anglais.
2. Langue du label (ISO 639-1). Le postulant spécifiera la langue de la chaîne TLD demandée, en respectant les codes ISO de la représentation des noms de langue et en Anglais.

3. Script du label (ISO 15924). Le postulant spécifiera le script de la chaîne gTLD demandée, en respectant les codes ISO de la représentation des noms de scripte et en Anglais.
4. Points de code Unicode. Le postulant fera une liste des tous les points de code contenus dans le label-U, selon sa forme Unicode.
5. Tables IDN. Une table IDN fournit la liste des caractères éligibles pour l'enregistrement d'un nom de domaine selon la politique de registre en vigueur. Elle contiendra tous les caractères multiples pouvant être considérés "égaux" dans le but des enregistrements de second niveau ("caractères variants"). Une fois utilisées par un registre TLD actif, les tables seront logées dans le reposoir des pratiques IDN. Pour plus d'information, voir les tables existantes sur <http://iana.org/domains/idn-tables/>, et le guide des demandes sur <http://iana.org/procedures/idn-repository.html>.

Les tables IDN doivent être présentées, au minimum, la langue ou le script de la chaîne de caractères gTLD demandée. Les tables IDN doivent être présentées pour chaque langue ou script dans lesquels le postulant pense proposer des enregistrements IDN de second niveau.

Il est demandé aux postulants de considérer les questions linguistiques et de système d'écriture dans leur travail de définition des caractères variants, et de coopérer avec d'autres opérateurs TLD offrant des enregistrements de noms de domaine avec les mêmes caractères ou des caractères visuellement similaires. L'ICANN peut aussi comparer la table IDN soumise par le postulant avec d'autres tables ayant été soumises pour les mêmes langues, et enquêter sur les inconsistances.

6. Les postulants devront également démontrer qu'ils ont fait les efforts nécessaires pour s'assurer de ce que la chaîne IDN encodée ne cause pas de problèmes opérationnels ou d'interprétation. Par exemple, des problèmes ont été identifiés dans des chaînes contenant des caractères de direction différentes (droite-à-gauche et gauche-à-droite) lorsque des caractères numériques sont adjacents au caractère séparateur (i.e., un point). Si un postulant fait une demande de chaîne connue pour ses problèmes récurrents, il devra documenter les mesures prises pour éviter ou minimiser ces problèmes dans les applications. Même s'il n'est pas possible d'assurer que tous les problèmes d'interprétation seront évités, il est important que le plus grand nombre d'entre eux soit identifié au plus tôt et que le potentiel opérateur du registre soit au courant de ces problèmes. Les postulants peuvent se familiariser avec ces problèmes en comprenant le protocole IDNA et en particulier la nouvelle version proposée du protocole IDNA (voir <http://www.icann.org/en/topics/idn/rfcs.htm>), et en

participant activement au wiki des IDN (voir <http://idn.icann.org/>) ou plusieurs problèmes d'interprétation sont expliqués.

7. **[Optionnel]** - Représentation du label en alphabet phonétique. Le postulant peut choisir de fournir la chaîne de caractères gTLD demandée annotée selon l'alphabet phonétique international (<http://www.langsci.ucl.ac.uk/ipa/>). Veuillez noter que cette information ne sera pas évaluée, ni notée. L'information, si fournie, sera utilisée comme guide par l'ICANN en cas d'enquête ou pour parler de la candidature lors de présentations publiques.

1.3.2 *TLD à IDN variants*

Chaque candidature contient une chaîne gTLD que le postulant tente de se voir attribuer. Pour la chaîne en question, le postulant peut également spécifier toute chaîne variante existante. Une chaîne variante résulte du remplacement d'un ou de plusieurs caractères dans la chaîne demandée, par des caractères variants. Chaque chaîne variante listée doit aussi être conforme aux conditions énoncées dans la section 2.1.1.3.2.

Pour chaque liste variante listée, le postulant doit spécifier s'il s'agit de:

- Une variante désirée. Cela signifie que le postulant souhaite faire usage de la chaîne comme TLD dans le futur.

ou

- Une variante non-désirée. Ceci signifie que le postulant a l'intention de ne jamais faire usage de la chaîne.

Veuillez noter que deux caractères variants peuvent se ressembler ou être différents visuellement. De plus, il n'y a aucune obligation à choisir des caractères variants se ressemblant. Cela dépend du script en question.

Les traductions ou translittérations de chaînes TLD ne sont pas considérées comme des variants. Le GNSO a recommandé dans son étude sur le développement des politiques gTLD que les droits de priorité pour les nouvelles chaînes de premier niveau ne dérivent pas de chaînes existantes.

L'ICANN enregistrera différentes chaînes de caractères variants listées par le postulant pour information, mais n'effectuera pas d'évaluation pour le moment.

Si la candidature est couronnée de succès, seule la chaîne gTLD demandée sera attribuée comme gTLD. Les chaînes de

caractères variants ne pourront être attribuées que lorsqu'un processus de gestion des TLD variants aura été achevé et testé par l'ICANN. A ce moment, les postulants se verront demander de fournir les informations complémentaires exigées par le processus de gestion des TLD variants et de participer au processus d'évaluation conséquent, ce qui provoquera d'éventuels frais supplémentaires, en fonction des étapes du processus qui restent à déterminer.

1.3 Information pour les postulants aux Noms de domaines internationalisés (souligné en rouge pour montrer la différence avec la version 3 du manuel)

Certaines chaînes de caractères gTLD ayant fait l'objet d'une demande devraient devenir des Noms de domaine internationalisés (IDN) requérant l'insertion d'un label-A encodé IDN dans la zone racine des DNS. Les IDN sont des noms de domaines incluant des caractères usés dans la représentation locale des langues non-écrites avec l'alphabet latin classique (a - z), les chiffres Européo-Arabes (0-9) et le tiret (-).

1.3.1 Conditions spécifiques aux IDN

Un postulant à une chaîne IDN doit fournir certaines informations montrant la conformité avec le protocole IDNA et d'autres conditions. Le protocole IDNA est en cours de révision et la documentation le concernant est disponible sur <http://icann.org/en/topics/idn/rfcs.htm> .

Les postulants doivent fournir les chaînes gTLD demandées sous la forme d'un **label-U** et un **label-A**.

Un label-A est la forme ASCII d'un label IDN. Chaque label-A commence par le préfixe IDNA ACE, "xn--", suivi d'une chaîne qui est le produit d'un algorithme Punycode, d'une longueur maximale de 59 caractères ASCII. Le préfixe et la chaîne réunis doivent être conformes à toutes les conditions nécessaires à un label qui pourrait se retrouver dans le DNS, ceci incluant la conformité avec la règle du LDH (nom d'hôte) décrite dans le RFC 1034, RFC 1123, et ailleurs.

Le label-U est la forme Unicode d'un label IDN, celle qu'un usager s'attend à voir s'afficher.

En prenant comme exemple la chaîne de test IDN actuelle en script cyrillique, le label-U est <**испытание**> et le label-A est <**xn--80akhbyknj4f**>. Un label-A doit pouvoir être produit par la conversion depuis un label-U, tout comme un label-U doit pouvoir être produit par un label-A.

Les postulants aux gTLD IDN devront également fournir les informations suivantes au moment de leur demande:

1. Version courte de la chaîne (en Anglais). Le postulant devra fournir une courte description de ce que la chaîne représenterait ou signifierait en anglais.
2. Langue du label (ISO 639-1). Le postulant spécifiera la langue de la chaîne TLD demandée, en respectant les codes ISO de la représentation des noms de langue et en Anglais.

3. Script du label (ISO 15924). Le postulant spécifiera le script de la chaîne gTLD demandée, en respectant les codes ISO de la représentation des noms de scripte et en Anglais.
4. Points de code Unicode. Le postulant fera une liste des tous les points de code contenus dans le label-U, selon sa forme Unicode.
5. Tables IDN. Une table IDN fournit la liste des caractères éligibles pour l'enregistrement d'un nom de domaine selon la politique de registre en vigueur. Elle contiendra tous les caractères multiples pouvant être considérés "égaux" dans le but des enregistrements de second niveau ("caractères variants"). Une fois utilisées par un registre TLD actif, les tables seront logées dans le repository des pratiques IDN. Pour plus d'information, voir les tables existantes sur <http://iana.org/domains/idn-tables/>, et le guide des demandes sur <http://iana.org/procedures/idn-repository.html>.

Les tables IDN doivent être présentées, au minimum, la langue ou le script de la chaîne de caractères gTLD demandée. Les tables IDN doivent être présentées pour chaque langue ou script dans lesquels le postulant pense proposer des enregistrements IDN de second niveau.

Il est demandé aux postulants de considérer les questions linguistiques et de système d'écriture dans leur travail de définition des caractères variants, et de coopérer avec d'autres opérateurs TLD offrant des enregistrements de noms de domaine avec les mêmes caractères ou des caractères visuellement similaires. L'ICANN peut aussi comparer la table IDN soumise par le postulant avec d'autres tables ayant été soumises pour les mêmes langues, et enquêter sur les inconsistences.

6. Les postulants devront également démontrer qu'ils ont fait les efforts nécessaires pour s'assurer de ce que la chaîne IDN encodée ne cause pas de problèmes opérationnels ou d'interprétation. Par exemple, des problèmes ont été identifiés dans des chaînes contenant des caractères de directionnalités différentes (droite-à-gauche et gauche-à-droite) lorsque des caractères numériques sont adjacents au caractère séparateur (i.e., un point). Si un postulant fait une demande de chaîne connue pour ses problèmes récurrents, il devra documenter les mesures prises pour éviter ou minimiser ces problèmes dans les applications. Même s'il n'est pas possible d'assurer que tous les problèmes d'interprétation seront évités, il est important que le plus grand nombre d'entre eux soit identifiés au plus tôt et que le potentiel opérateur du registre soit au courant de ces problèmes. Les postulants peuvent se familiariser avec ces problèmes en comprenant le protocole IDNA et en particulier la nouvelle version proposée du protocole IDNA (voir <http://www.icann.org/en/topics/idn/rfcs.htm>), et en

participant activement au wiki des IDN (voir <http://idn.icann.org/>) ou plusieurs problèmes d'interprétation sont expliqués.

7. **[Optionnel]** - Représentation du label en alphabet phonétique. Le postulant peut choisir de fournir la chaîne de caractères gTLD demandée anotée selon l'alphabet phonétique international (<http://www.langsci.ucl.ac.uk/ipa/>). Veuillez noter que cette information ne sera pas évaluée, ni notée. L'information, si fournie, sera utilisée comme guide par l'ICANN en cas d'enquête ou pour parler de la candidature lors de présentations publiques.

1.3.2 *TLD à IDN variants*

Chaque candidature contient une chaîne gTLD que le postulant tent de se voir attribuer. Pour la chaîne en question, le postulant peut également spécifier toute chaîne variante existante. Une chaîne variante résulte du remplacement d'un ou de plusieurs caractères dans la chaîne demandée, par des caractères variants. Chaque chaîne variante listée doit aussi être conformes aux conditions énoncées dans la section 2.1.1.3.2.

Pour chaque liste variante listée, le postulant doit spécifier s'il s'agit de:

- Une variante désirée. Cela signifie que le postulant souhaite faire usage de la chaîne comme TLD dans le futur.

ou

- Une variante non-désirée. Ceci signifie que le postulant a l'intention de ne jamais faire usage de la chaîne.

Veillez noter que deux caractères variants peuvent se ressembler ou être différents visuellement. De plus, il n'y a aucun obligation à choisir des caractères variants se ressemblant. Cela dépend du script en question.

Les traductions ou translittérations de chaînes TLD ne sont pas considérées comme des variants. Le GNSO a recommandé dans son étude sur le développement des politiques gTLD que les droits de priorité pour les nouvelles chaînes de premier niveau ne dérivent pas de chaînes existantes.

L'ICANN enregistrera différentes chaînes de caractères variants listées par le postulant pour information, mais n'effectuera pas d'évaluation pour le moment.

Si la candidature est couronnée de succès, seule la chaîne gTLD demandée sera attribuée comme gTLD. Les chaînes de

caractères variants ne pourront être attribuées que lorsqu'un processus de gestion des TLD variants aura été achevé et testé par l'ICANN. A ce moment, les postulants se verront demander de fournir les informations complémentaires exigées par le processus de gestion des TLD variants et de participer au processus d'évaluation conséquent, ce qui provoquera d'éventuels frais supplémentaires, en fonction des étapes du processus qui restent à déterminer.