

Noms de domaines internationalisés (IDN)

Marc Blanchet

Viagénie

<http://www.viagenie.ca>

Présenté aux Journées Tunisiennes de l'Internet, Février 2008

Plan



- DNS
- Noms de domaines internationalisés
- Rôle et processus de l'ICANN
- Conclusion
- Références

DNS

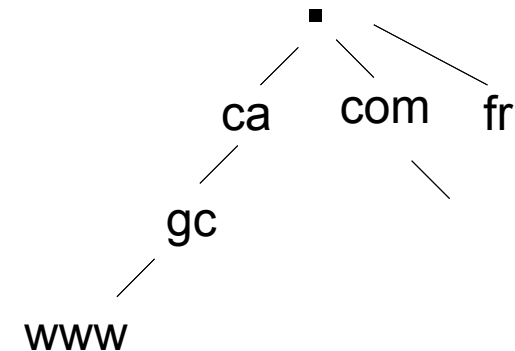


- DNS: Domain Name System
- Système de base de données distribuée
 - Globalement très fiable. Résiste très bien à des charges importantes (11 sept 2001), des attaques significatives, ...
- Permettant l'association entre un nom (de domaine) et une donnée de protocole:
 - Nom: ca ou test.fr ou xyz.exemple.com ou x.y.abc.org
 - Donnée de protocole:
 - Adresse IP (IPv4 ou IPv6), identifiant un ordinateur sur le réseau
 - Clés de sécurité
 - Localisation de services
 - Endroit où transférer un appel

DNS



- Pour ouvrir une page web: http://www.gc.ca/main_f.html
 - Votre navigateur extrait le nom: 'www.gc.ca'
 - Votre ordinateur fait une demande pour l'adresse IP de www.gc.ca, au serveur DNS de l'entreprise ou du fournisseur Internet.
 - En votre nom, le serveur DNS trouve la réponse en parcourant la hiérarchie du DNS:
 1. '.' (racine, root)
 2. 'ca'
 3. 'gc.ca'
 4. 'www.gc.ca'
 - Trouve la réponse: 198.103.238.30
 - Votre navigateur effectue une communication TCP/IP à l'ordinateur 198.103.238.30 avec le protocole HTTP.



Autorité



- À chaque niveau (séparé par des '.') correspond généralement un organisme et des serveurs DNS qui ont l'information pour ce niveau.
- Racine de la hierarchie du DNS: '.' (premier niveau)
 - Organisme: ICANN
 - 13 (en réalité ++) serveurs DNS, localisés partout dans le monde. Leurs adresses sont incluses avec tout logiciel serveur DNS.
 - Appelée 'root zone'.
 - Contient l'information des serveurs DNS pour le niveau plus bas:
 - Qui sont les serveurs DNS pour com., net., fr., ca., ...

Deuxième niveau



- Appelés: TLD (top-level domain)
- Plusieurs catégories:
 - Nom de pays: ca, fr, us, jp, ...
 - Génériques: com, net, org, info, ...
 - US: gov, mil, edu, ...
- Avant, la procédure pour avoir l'autorité sur un TLD était (trop) 'simple'. Procédure plus formelle maintenant.

Deuxième niveau: ca.



- .ca:
 - Organisme: CIRA/ACEI est responsable pour .ca.
 - Historiquement fait par John Demco de UBC.
 - Transféré à CIRA/ACEI 'par' le gouvernement canadien.
 - Zone .ca
 - Servie par des serveurs DNS sous l'autorité de CIRA/ACEI
 - Contient l'information pour les serveurs DNS pour le niveau plus bas:
 - Qui sont les serveurs DNS pour gc.ca, viagenie.ca, ...

Troisième niveau: gc.ca



- gc.ca est géré par les serveurs DNS du gouvernement canadien.
 - Contient l'information du niveau plus bas:
 - Quelle est l'adresse IP de ...?
 - www.gc.ca => 198.103.238.30

IDN



- IDN = Internationalized Domain Names
- Objectif: Supporter 'toutes' les langues (pour être plus précis: les écritures) dans les noms de domaine.
- Travaux d'ingénierie à l'IETF (Internet Engineering Task Force) de ~1999-2002.
- Solution:
 - Utilisation de presque la totalité des caractères identifiés dans le jeu Unicode/ISO-10646
 - Conversion par le logiciel client (ex: navigateur web, logiciel de courriel) entre les caractères et une représentation ascii, en ajoutant un préfixe identifiant un idn (xn--).

IDN RFC



- Les RFC qui composent la spécification pour les idn:
 - [RFC3490] Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications (IDNA)", RFC 3490, March 2003.
 - [RFC3454] Hoffman, P. and M. Blanchet, "Preparation of Internationalized Strings ("stringprep")", RFC 3454, December 2002.
 - [RFC3491] Hoffman, P. and M. Blanchet, "Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)", RFC 3491, March 2003.
 - [RFC3492] Costello, A., "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)", RFC 3492, March 2003
- cet ensemble est aussi nommé IDNA2003 faisant référence à l'année de publication. Pour faire distinction à la nouvelle version en discussion.

IDN



- Exemple:
 - Dans le navigateur: <http://www.viagénie.ca> (notons le 'é')
 - Le navigateur convertit en: <http://www.xn--viagnie-eya.ca>
 - La requête DNS est pour: www.xn--viagnie-eya.ca
 - Même processus que précédemment. Même autorité.
- Pour supporter [viagénie.ca](http://www.viagénie.ca), il me faut ajouter (acheter) le nom idn (xn--viagnie-eya) dans .ca.
- Le codage est par niveau et indépendant:
 - 賓.viagénie.españa
-

Définitions et mots-clés



- Étiquette U (Unicode): viagénie
- Étiquette A (Ascii compatible encoding): xn—viagnie-
eya
- punycode: algorithme de codage unicode à ASCII
- IDNA: internationalized domain names in applications
- nameprep, stringprep: préparation des chaînes de caractères en vue de les coder en ASCII.
- LDH: letters, digits, hyphen: alphabet restreint utilisé pour les étiquettes des noms de domaines.

IDN

- Problématiques:

- Unicode/ISO-10646 a été conçu pour l'affichage de caractères à l'écran, sur papier. Le DNS a été conçu pour une correspondance unique. Que faire avec deux caractères identiques ou semblables ou ...?
 - Attaques possibles basées sur les ressemblances.
- Pour plusieurs écritures (exemple “Han”), il est important pour l'utilisateur de connaître la langue afin de comprendre le nom. Or la langue n'est pas identifiée dans les idn.
- Est-ce que tous les registres doivent supporter toutes les langues? (Est-ce que CIRA (.ca) doit offrir le chinois, le japonais, ...? Si oui, faudra avoir des personnes compétentes pour le support de ces langues. Les registraires aussi!
- Peut-on mélanger plusieurs langues dans un même nom?

IDN

- Problématiques:

- Les noms de domaines sont des identifiants. Mais ils ont une forte valeur de marque. Une même séquence de caractères dans certaines langues asiatiques peut avoir des significations forts différentes.
- Résolution de conflits sur les noms de domaines internationalisés!
- Explosion possible des TLD:
 - .es = .españa ? = .espagne?
 - .cn = .中華 ? = .中國 ? .zhōnghuá (pinyin)? .chunghua (Wade-Giles)
- Géopolitique: Taiwan ou République du peuple de Chine?
- Comment entrer les codes d'une langue X avec mon clavier Y?

Implémentations et registres

- Implémentations:
 - Mozilla/Firefox 1.4+, Netscape 7.1+, Opera 7.11+, Safari 1.2+, IE 7.
- Registres:
 - .ac, .ae, .at, .biz, .br, .cat, .com, .ch, .cl, .cn, .de, .dk, .fi, .gr, .hk, .hu, .info, .io, .is, .jp, .kr, .li, .lt, .lv, .museum, .net, .no, .nu, .org, .pl, .se, .sh, .tm, .tw, .vn
- Travail en cours:
 - Réviser les caractères à exclure.
 - Adresses de courriel internationalisées

Variations sur l'affichage



- certaines implémentations ont décidé de ne pas afficher l'étiquette U mais l'étiquette A selon certains critères spécifiques à l'implémentation.
- Firefox, par défaut, a une liste de registre de premier niveau où il montrera l'étiquette U. Les autres seront affichés par l'étiquette A. Par exemple, il ne montrera pas `http://www.viagénie.com` mais `http://www.xn--viagnie-eya.com`
 - cependant, on peut configurer pour que `.com` soit reconnu.
 - voir <http://marcblanchet.blogspot.com>
- Safari ne supporte pas certaines écritures. On peut changer la liste cependant.

Registre IDN de l'IANA



- Registres de niveau 1 ont identifié les sous-ensembles de Unicode qu'ils supportent.
- Ces sous-ensembles sont documentés dans un registre de l'IANA:
 - <http://www.iana.org/assignments/idn/index.html>

Nouvelle révision du protocole



- Ensemble de caractères:
 - présentement: inclure tous les caractères Unicode et exclure un sous-ensemble
 - nouvelle proposition: inclure seulement un sous-ensemble déterminé.
- Certains problèmes découverts avec l'expérience de la première version.
- En cours à l'IETF
 - draft-klensin-idnabis-*
 - draft-falstrom-idnabis-tables*
 - draft-alvestrand-idna-bidi*
 - Discussions sur la formation d'un groupe de travail

ICANN: test idn tld



- Besoin de vérifier si les serveurs DNS de la racine (root servers) seraient affectés par l'introduction de TLD internationalisés.
- Un analyse a été faite avec une expérimentation pratique. La conclusion est qu'il n'y aucun effet.

ICANN: example.test



- Pour valider le support des IDN (incluant les TLD) dans les différentes applications Internet,
- ICANN a mis en place un ensemble de 11 noms de domaines représentant example.test en 10 langues en 11 écritures.

ICANN: example.test



- | | | | |
|-----------------------------|---------------------------|---------------------------|----------|
| • http:// مثال.إختبار | http:// مثال.إختبار | Arabic | Arabic |
| • http:// 例子 . 测试 | http:// 例子 . 测试 | Simplified Chinese | Chinese |
| • http:// 例子 . 測試 | http:// 例子 . 測試 | Traditional Chinese | Chinese |
| • http:// παράδειγμα.δοκιμή | http:// παράδειγμα.δοκιμή | Greek | Greek |
| • http:// उदाहरण.परीक्षा | http:// उदाहरण.परीक्षा | Devanagari | Hindi |
| • http:// 例え . テスト | http:// 例え . テスト | Kanji, Hiragana, Katakana | Japanese |
| • http:// 실례 . 테스트 | http:// 실례 . 테스트 | Hangul | Korean |
| • http:// مثال.آزمایشی | http:// مثال.آزمایشی | Perso-Arabic | Persian |
| • http:// пример.испытание | http:// пример.испытание | Cyrillic | Russian |
| • http:// உதாரணம்.பரிட்சை | http:// உதாரணம்.பரிட்சை | Tamil | Tamil |
| • http:// בישפיל.טעסט | http:// בישפיל.טעסט | Hebrew | Yiddish |

ICANN: idn TLD



- Processus en cours pour déployer les TLD internationalisés.
 - “fast track” pour certains ccTLD bien circonscrits.
 - développement d'une politique

Prochaines étapes



- IETF:
 - finalisation de la nouvelle version du protocole
- ICANN:
 - tests et évaluations
 - évaluation des impacts sur la sécurité du DNS
 - politique et règles sur les IDN
 - approbation
- Registres TLD et registraires
 - déploiement

Adresses de courriel



- partie gauche de l'adresse de courriel
- groupe de travail à l'IETF: eai

Conclusion



- Les noms de domaines internationalisés sont une technologie importante pour le futur de l'Internet.
- Ils génèrent aussi un ensemble de problématiques qui sont en train d'être étudiées.
- Il est souhaitable que tous les pays non-anglophones participent activement aux travaux à l'IETF et ICANN.

Questions?



Contact:

Marc.Blanchet@viagenie.ca

Présentation disponible à <http://www.viagenie.ca/publications/>

Références

- Hoffman, P. and M. Blanchet, "Preparation of Internationalized Strings ("stringprep")", RFC 3454, December 2002.
- Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications (IDNA)", RFC 3490, March 2003.
- Hoffman, P. and M. Blanchet, "Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)", RFC 3491, March 2003.
- Costello, A., "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)", RFC 3492, March 2003.
- Klensin J., Faltstrom P. and C. Karp, "Review and Recommendations for Internationalized Domain Names (IDNs)", RFC 4690, September 2006.
- Internet drafts nouvelle version de idna: draft-*idna**
- ICANN IDN: <http://www.icann.org/topics/idn/>
- IDN, Wikipedia, http://en.wikipedia.org/wiki/Internationalized_domain_name
-