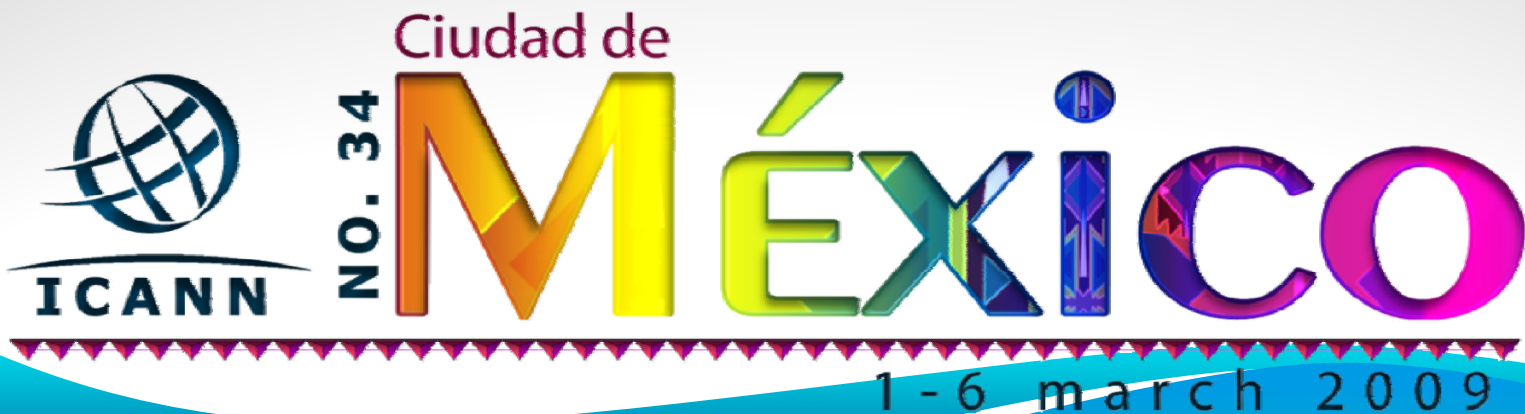


Internationalized Domain Names – an introduction

Tina Dam

Director, IDN Program

1 March 2009



Agenda

- Where are we and where are we headed
- IDN TLD Processes
- IDN Definitions
- How does IDNs work
 - including examples of applications and their various implementations
- IDN Confusability
 - Issues and solutions
- IDN wiki facility

What we have / what we need

- IDNs have existed as second level since 2003
 - under web protocol standards (under revision)
 - email protocol standards are underway (IETF)
- We also need IDN TLDs
 - 北京.中国; [xn--1lq9oi. xn--fiqs8s]

Domain Availability Today



ASCII domain names (a, b,...,z), (0,1,...,9), (-)	domainname.TLD icann.org
IDN second level	실레.TLD -under various existing TLDs
IDN TLDs	실레. 테스트

Future Addition



IDN TLD Processes

Implementation of Country-code IDN TLDs – Fast Track

- implementation of recommended policy from IDNC WG
- review of relationship and financial model for IDN TLD operators
- review of technical requirements for IDN TLD management
- review of variant management

Country-code IDN TLDs – Long Term

- Full policy that caters for all
- Follows the full ccNSO Policy Development Process

New Generic TLDs

- New ongoing policy for new gTLDs
- Includes internationalized domains
- Focus on non-ASCII squatting & confusingly similarity solutions

Characters in the DNS

- Search on “*US-ASCII character set*”
- The DNS can handle all US-ASCII characters
 - Examples:
 - (a...z), (o...9), (-)
 - () SPACE
 - (!) EXCLAMATION MARK
 - (") QUOTATION MARK
 - (#) NUMBER SIGN
 - (\$) DOLLAR SIGN
 - (%) PERCENT SIGN
 - (&) AMPERSAND
 - (') APOSTROPHE

Characters, DNS, and domain names...

- All TLD registries have implemented the LDH rule
 - Domain names can only contain:
 - (a,b,...z)
 - (0,1,...9)
 - (-)
- ...before internationalization....

IDN Definitions

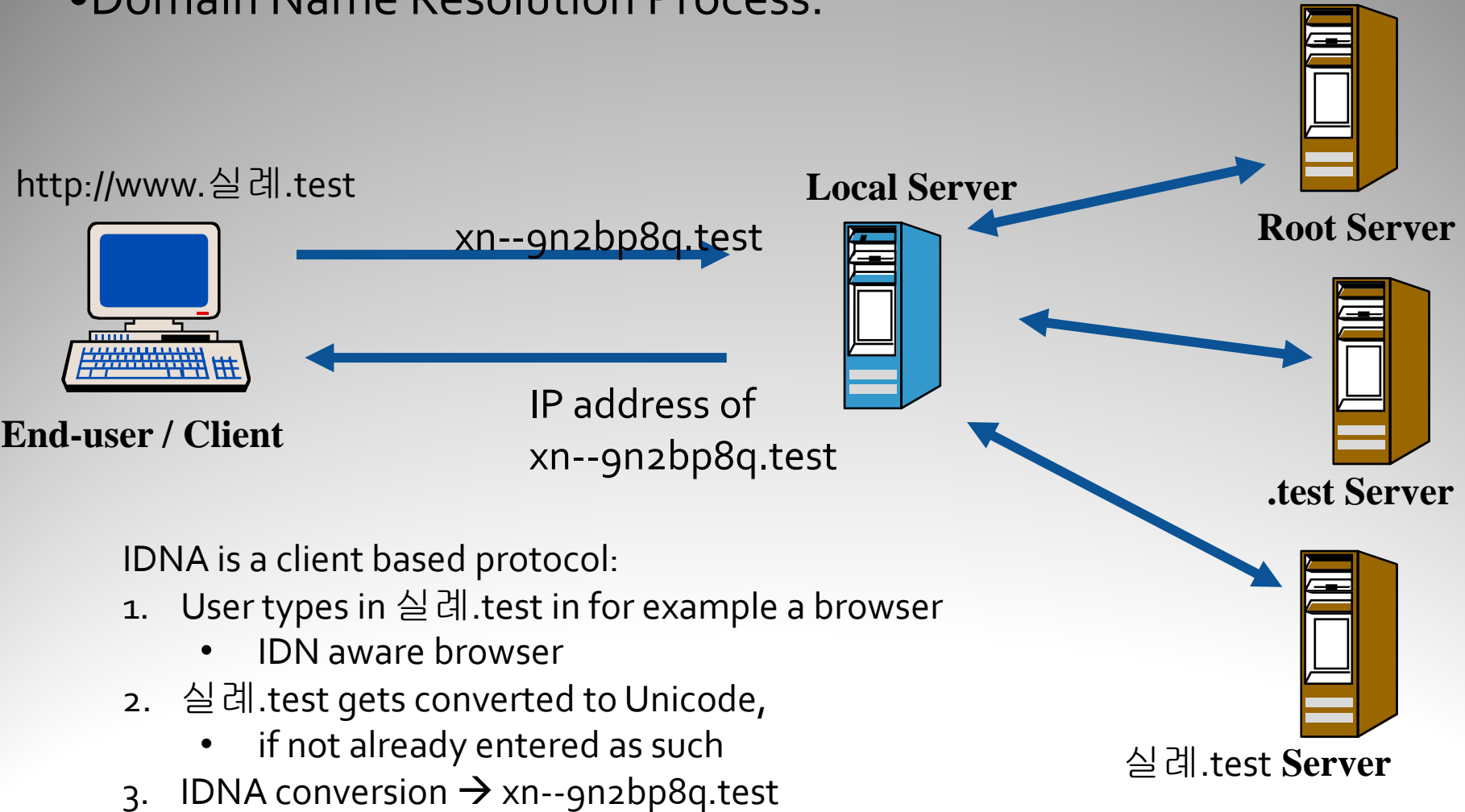
- Internationalized domain names are:
 - Names with characters other than the standard ASCII (a,b,...z), (0,1,...9), (-)
- Example:
ñandú.cl → xn--and-6ma2c.cl
- IDNs are about localized solutions
 - But need to be ‘internationalized’ due to the global nature of the Internet

Why Internationalization?

- DNS handling US-ASCII character set
 - a natural choice at the time
 - no expectation to current commercial value
 - Unicode was not available
- IDNs a natural expansion for global usability
 - allow users to use domain names in local scripts
 - no need to learn US-ASCII, user choice
 - some applications have implemented IDNA
 - still need internationalization of TLD

IDNA – Protocol Functionality

• Domain Name Resolution Process:



Internet Explorer - IDNwiki - Windows Internet Explorer

Address bar: http://مثال.إختبار/الصفحة_الرئيسية/%D8%A7%D9%84%D8%B5%D9%81%D8%AD%D8%A9_%D8%A7%D9%84%D8%B1%D8%A...

Opera - IDNwiki - Opera

File Edit View Bookmarks Widgets Tools Help

New tab الصفحة الرئيسية - IDNwiki

Address bar: http://مثال.إختبار/الصفحة_الرئيسية/%D8%A7%D9%84%D8%B5%D9%81%D8%AD%D8%A9_%D8%A7%D9%84%D8%B1%D8%A...

Mozilla Firefox - IDNwiki - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Address bar: http://xn--fsqu00a.xn--0zwm56d/%E9%A6%96%E9%A1%B5

Mozilla Firefox - IDNwiki - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Address bar: http://مثال.إختبار/%D8%A7%D9%84%D8%B5%D9%81%D8%AD%D8%A9_%D8%A7%D9%84%D8%B1%D8%A...

Displayed Form vs. Stored Form

- Historically the domain name you register is also the domain names stored and usable in the DNS
- This is changed with introduction of IDNs
- The stored form usually gives no meaning
 - Example: فرساننهر.tld → xn--mgbtbg2evaoi.tld
- However, there are exceptions:
 - xn--gibberish - decodes into the Arabic characters ب٨٧٩أ
 - xn--trademark - with different versions of trademarks
 - This is coincidentally and hence not intentionally, but implementations showing xn- to the user is a concern
- xn-- prefix indicates to application software that the label needs to be decoded back into Unicode for proper display to the user

IDN Confusability

- Not a new topic
 - “o” looks like “o” and “i” looks like “l”
- Increased problem with increased # of characters
 - ASCII: 37 characters used in domain names
 - IDNs -> approximately 100.000 characters available
- Well known examples:
 - “paypal” and “paypal”
 - “py” and “py”
- Some problems solved in the IDN Guidelines
 - Preventing mixing of scripts unless a linguistic need

Preventing confusability

- IDN Tables & variants

- IDN tables are developed by registry managers to
 - Inform users what characters are available
 - Eliminate confusability by listing variant characters
- IDN Tables are used both on second level and top level
 - Strongly urging collaboration across language communities when potential for confusion exists
 - Languages using the same script
 - Scripts looking alike (fx. Cyrillic, Greek, Latin)
 - IDN TLD variant strings
 - Proposed to be allocated or blocked in the IDN ccTLD Fast Track Process

IDNA Protocol Revision

- IDNA provides the technical requirement for IDN strings:
 - The label must be a valid internationalized domain name, as specified in technical standards <http://www.icann.org/en/topics/idn/rfcs.htm>.
- Protocol revision is ongoing in the IETF
- Main positive results are:
 - Unicode version independent
 - Fixing problems with right-to-left script strings

IDNA protocol and IDN wiki

- <http://idn.icann.org>
- <http://josefsson.org/idn.php>
 - IDNA ToASCII
 - IDNA ToUnicode
- If you can't type in an IDN then search for your favorite newspaper online and copy-paste it
 - or you can copy-paste text from the IDN wiki
- Try copy / paste between applications you normally use

IDN wiki

Script	Language	SLD.TLD U-labels	SLD A-label	TLD A-label
Arabic	Arabic	مثال.إختبار	xn--mgbh0fb	xn--kgbechtv
Arabic	Persian	مثال.آزمایشی	xn--mgbh0fb	xn--hgbk6aj7f53bba
Chinese, simplified	Chinese	例子.测试	xn--fsqu00a	xn--0zwm56d
Chinese, traditional	Chinese	例子.測試	xn--fsqu00a	xn--g6w251d
Cyrillic	Russian	пример.испытание	xn--e1afmkfd	xn--80akhbyknj4f
Devanagari	Hindi	उदाहरण.परीक्षा	xn--p1b6ci4b4b3a	xn--11b5bs3a9aj6g
Greek	Greek	παράδειγμα.δοκιμή	xn--hxajbhcg2az3al	xn--jxalpdlp
Hangul	Korean	실례.테스트	xn--9n2bp8q	xn--9t4b11yi5a
Hebrew	Yiddish	טעסט.דאָקומענט	xn--fdbk5d8ap9b8a8d	xn--deba0ad
Kanji Hirigana, and Katakana	Japanese	例え.テスト	xn--r8jz45g	xn--zckzah
Tamil	Tamil	உதாரணம்.பரிட்சை	xn--zkc6cc5bi7f6e	xn--hlcj6aya9esc7a

IDNwiki - IDNwiki - Windows Internet Explorer
http://idn.icann.org/
File Edit View Favorites Tools Help
IDNwiki - IDNwiki


ICANN
article discussion view source history
IDNwiki
Welcome to the IDN TLD evaluation gateway!
There are new articles on [Additional languages](#) and [Technical](#)

- navigation
- IDNwiki main page
 - E-mail test
 - Basic concepts
 - Technical topics
 - Software
- interaction
- Recent changes
 - Contact
 - Wiki help
- evaluation
- العربية
 - 简体中文
 - 繁體中文
 - Ελληνικά
 - हिन्दी
 - 日本語
 - 한국어
 - فارسی
 - Русский
 - தமிழ்
 - עברית
- additional

Contents [hide]

- 1 Introduction
- 2 Your participation is important!
- 3 Limited evaluation period
- 4 Things to test
- 5 Further information about the IDNwiki
- 6 The example.test names

Introduction

This page provides an introduction to a test of IDN top-level domain names that ICANN is coordinating. The test is based on eleven new internationalized domains representing the name **example.test** entirely in scripts other than the familiar Latin characters that appear in current top-level labels. The languages initially selected for illustrating this are listed in the table below, and the rationale behind their choice is discussed in the sidebar article on [basic concepts](#). These TLDs can be accessed by clicking on the links in the first column in the table. However, as with any other IDNs, if they are typed or copied and pasted directly into the address line of a browser, they will only work if that browser has full support for IDN. The names in the second column are intended to be used in that manner and, if they don't initially perform as intended, some software reconfiguration may help. Additional articles discuss [local configuration](#) and individual [software applications](#).

Your participation is important!

Public participation in the evaluation of these domains is one of the most important parts of the project. Joining this initiative requires nothing more than for you to click through one or more of the links in the table and report about the experience on the "discussion page" indicated with a tab at the top of this and the other IDNwiki articles.

- Were the results what you expected?
- Were there any problems that you couldn't solve?

Status of the .test wiki

- Purpose of the IDNwiki:
 - Introduce users to IDN TLDs
 - Applications test environment for usability
 - Registry information about user problems
 - General information about use of IDNs (fx. Fonts)
- Conduct an experiment with IDN TLDs
 - not a pre-requisite for production in root zone
 - no registrations are available
- Functions as a “normal wiki”, user access

**Internationalization of the internet means
that the internet is equally accessible from
all languages and scripts**

Gracias

tina.dam@icann.org

<http://icann.org/topics/idn>

IDN Sessions this week in Mexico are mixed within:

- **Constituency meetings**
- **GAC meeting**
- **ccNSO meeting**
- **gTLD sessions**
- **Monday, 11-12.30: IDN issues ahead**