Internationalized Domain Names

21 June 2009

Tina Dam
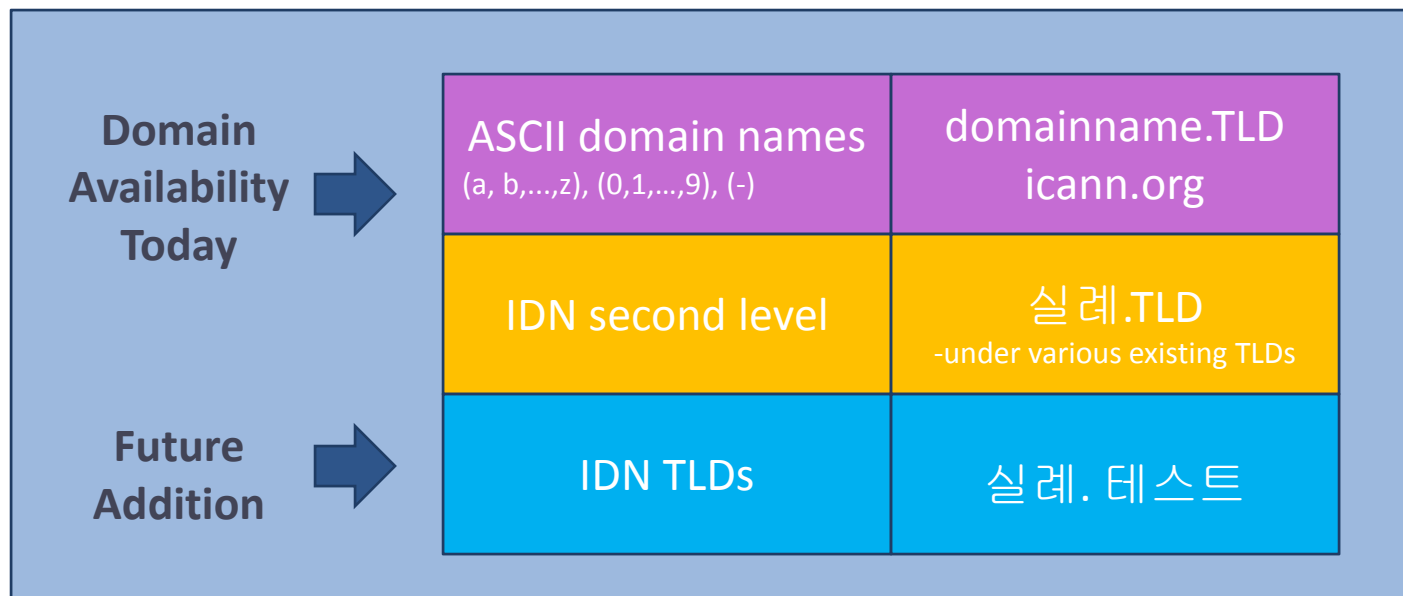
Sr. Director, IDNs

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
- IDN sessions this week in Sydney

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--1lq90i. xn--fiqs8s]

Domain Availability Today ➡	ASCII domain names (a, b,...,z), (0,1,...,9), (-)	domainname.TLD icann.org
	IDN second level	실례.TLD -under various existing TLDs
Future Addition ➡	IDN TLDs	실례. 테스트

IDN TLD Processes

Implementation: IDN ccTLDs Fast Track Process

- to introduce a limited number of IDN ccTLDs
- non-Latin scripts only, matching ISO3166 list
- must match country/territory names

Policy Development: IDN ccTLDs – Long Term

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

Implementation: New gTLDs

- Includes internationalized domains
- IDN technical requirements same as Fast Track
- 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), (0...9), (-)
    - ( ) SPACE
    - (!) EXCLAMATION MARK
    - (") QUOTATION MARK
    - (#) NUMBER SIGN
    - ($) DOLLAR SIGN
    - (%) PERCENT SIGN
    - (&) AMPERSAND
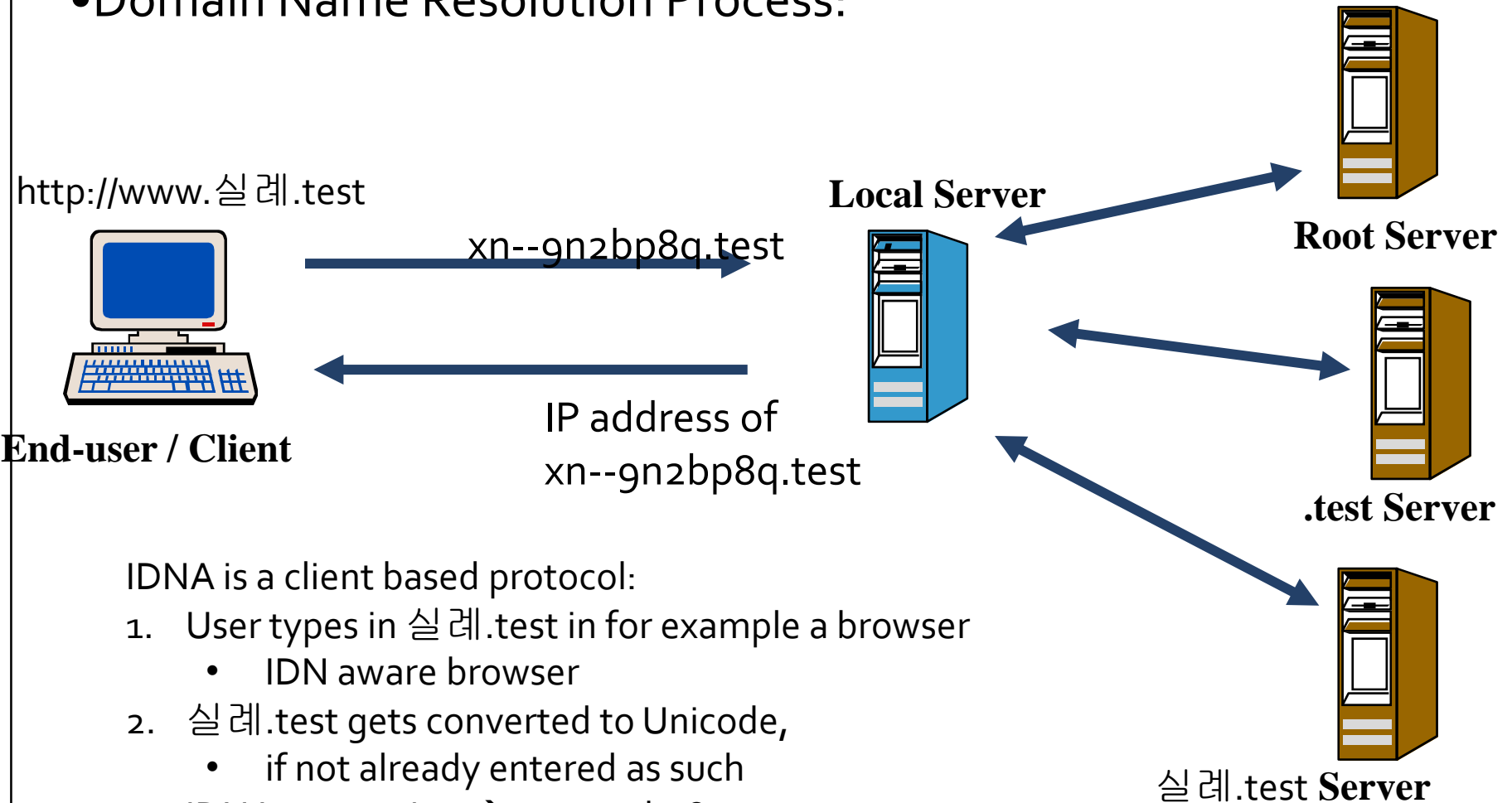   - (') APOSTROPHE

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

IDNA – Protocol Functionality

- Domain Name Resolution Process:

http://www.실례.test

Local Server

Root Server

xn--9n2bp8q.test

IP address of
xn--9n2bp8q.test

End-user / Client

.test Server

IDNA is a client based protocol:
1. User types in 실례.test in for example a browser
   - IDN aware browser
2. 실례.test gets converted to Unicode,
   - if not already entered as such
3. IDNA conversion → xn--9n2bp8q.test

실례.test **Server**

□□□□□□ □□□□□□□□ - IDNwiki - Windows Internet Explorer

http://مثالخ.اختبار/%D8%A7%D9%84%D8%B5%D9%81%D8%AD%D8%A9_%D8%A7%D9%84%D8%B1%D8%A

□□□□□□ □□□□□□□□ - IDNwiki - Opera

File   Edit   View   Bookmarks   Widgets   Tools   Help

New tab   الصفحة الرئيسية - IDNwiki

http://الصفحة_الرئيسية/اختبار.مثال

□□ - IDNwiki - Mozilla Firefox

File   Edit   View   History   Bookmarks   Tools   Help

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

□□□□□□ □□□□□□□□ - IDNwiki - Mozilla Firefox

File   Edit   View   History   Bookmarks   Tools   Help

http://اختبار.مثال/%D8%A7%D9%84%D8%B5%D9%81%D8%AD%D8%A9_%D8%A7%D9%84

IDN Confusability

- Not a new topic
  - "0" looks like "o" and "1" 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-past 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--hxajbheg2az3al	xn--jxalpdlp
Hangul	Korean	실례.테스트	xn--9n2bp8q	xn--9t4b11yi5a
Hebrew	Yiddish	בײַשפּיל.טעסט	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

IDNwiki

article | discussion | view source | history

Welcome to the IDN TLD evaluation gateway!

There are new articles on Additional languages and Technical

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

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
- ייִדיש

additional

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?

IDN TLDs – *technical related issues ahead*

- Variant TLD reservation or blocking
  - No technical solution to alias TLDs in the root, but community need
    - for example: Pakistan: پاکستان and پاکستان
- Number of characters in a label
  - ccTLDs 2char labels → 2+char
  - gTLDs 3+char →1char, 2char, 3char, and 3+char ?
  - Usability and acceptability across applications?
- IDNA technical standard revision
  - Next IETF meeting, Stockholm, July 2009
- Security and stability impact of number of TLDs
  - Scaling study initiated by SSAC and RSSAC

IDN ccTLD Fast Track Process

- Revised Draft Implementation Plan
  - Public comment period open until 15 July 2009
  - Operational Elements and Application Form Details to be finalized by Q4-2009:
    - online application form, IDNA implementation, technical string requirements, TAS/gTLD interface, formation of stability panel, finalize linguistic process and contacts
- Revised DoR paper:
  - Content and form seems agreeable
  - Dispute/Enforcement mechanism needed
- Detailed Financial Papers w/Supporting analysis
  - Pre-arranged processing fee (25-50K$)
  - Annual Contribution fee (1-3% of revenue)

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

Thank you

tina.dam@icann.org
http://icann.org/topics/idn

IDN Sessions this week in Sydney:

- Monday, 11-12.30: New gTLDs Overview
- Monday 17-18.30: IDN ccTLD Fast Track Session
- Tuesday 9-10.30: GAC IDN Technical Requirements
- Tuesday: Various Constituency meetings
- Wednesday 9-10.45: Fast Track & IDN PDP