

# Internationalized Domain Names

**Doug Brent**  
**Chief Operating Officer**

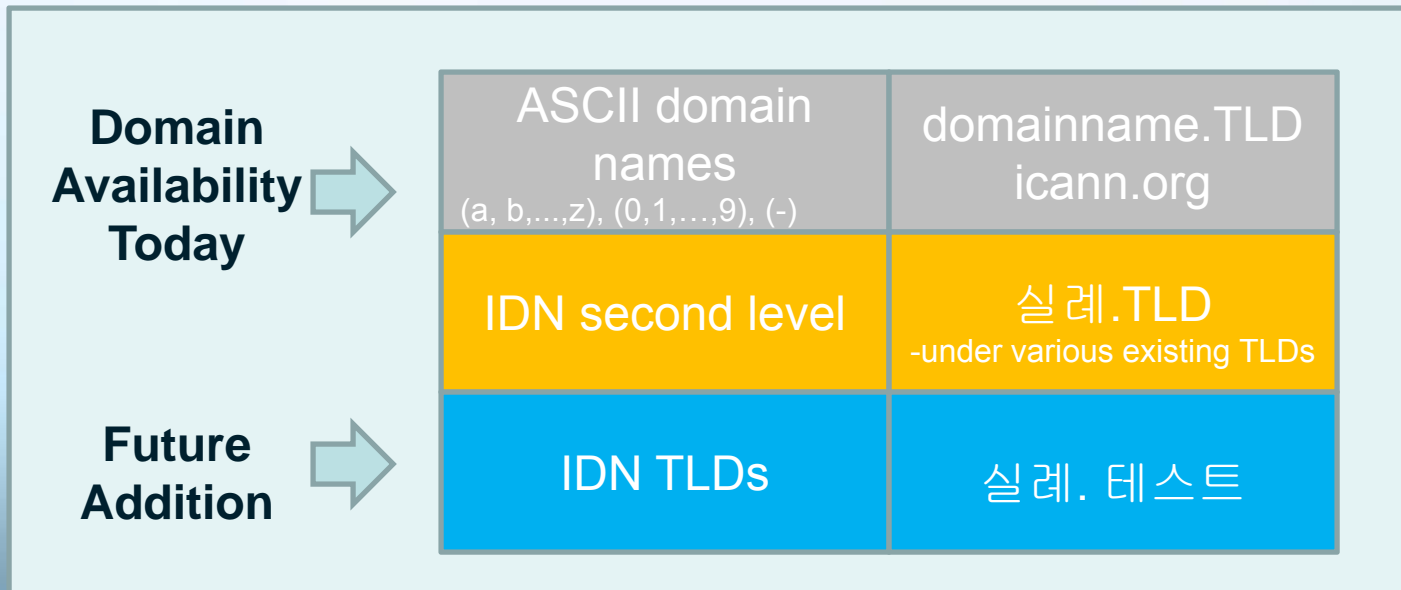
**Hong Kong**  
**24 July 2009**

# Agenda

- Internationalized Domain Names (IDNs) defined
- How IDNs work
- IDN confusability
  - Issues and solutions
- Remaining technically-related issues
- IDN processes at ICANN

# What we have / what we need

- IDNs have existed at the second level since 2003
  - under web protocol standards (under revision)
  - email protocol standards are underway (IETF)
- We also need IDN top level domains (TLDs)
  - 北京.中国; [xn--1lq90i. xn--fiqs8s]



# Characters in the DNS

- 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...

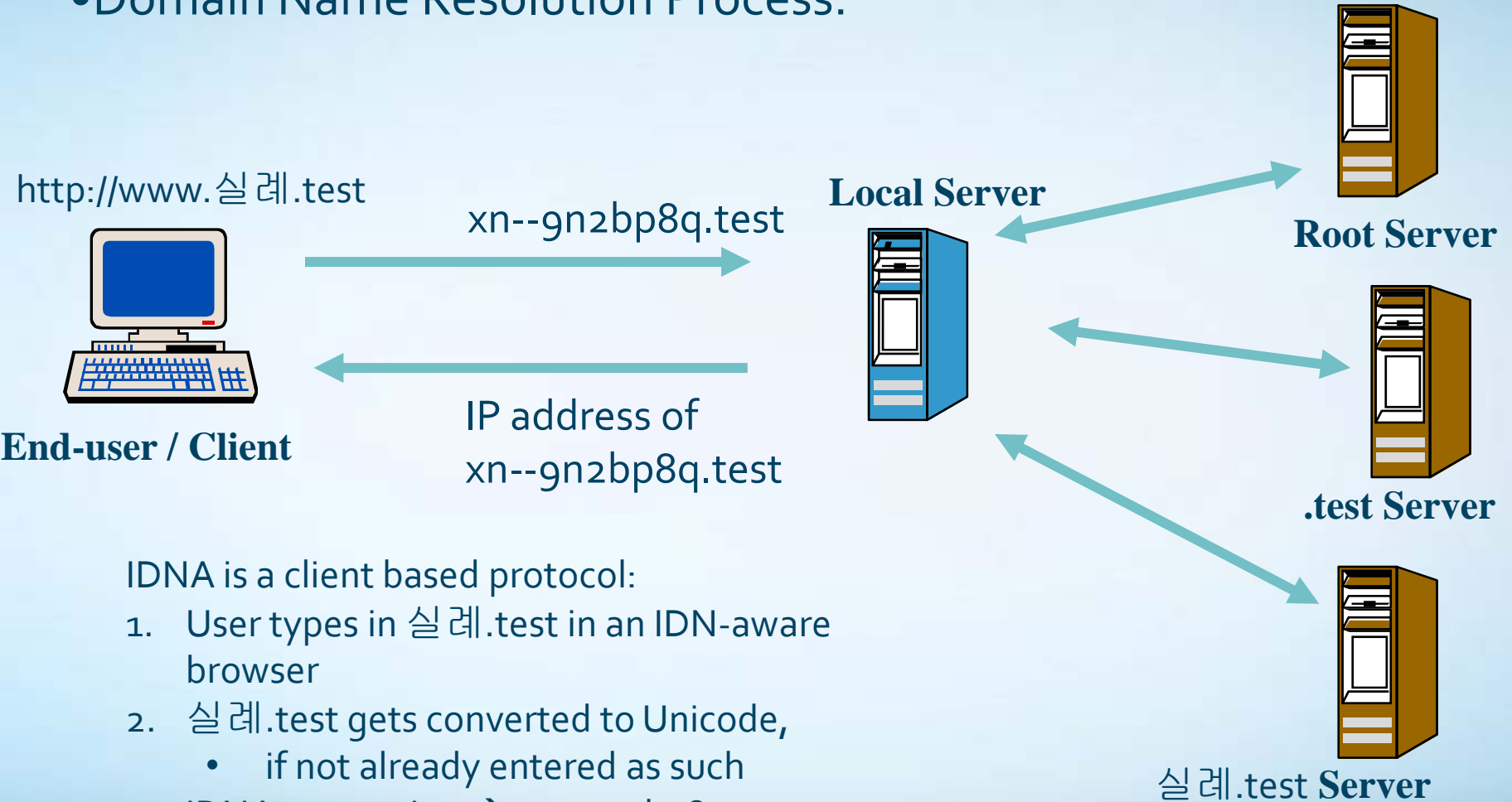
- All TLD registries have implemented the LDH (letter digit hyphen) rule
  - Domain names can only contain:
    - (a,b,...z)
    - (0,1,...9)
    - (-)
- ....before internationalization....

# 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 offer localized solutions...
  - ...but need to be ‘internationalized’ due to the global nature of the Internet

# IDNA – Protocol Functionality

## •Domain Name Resolution Process:



IDNA is a client based protocol:

1. User types in 실례.test in an IDN-aware browser
2. 실례.test gets converted to Unicode,
  - if not already entered as such
3. IDNA conversion → xn--gn2bp8q.test
4. DNS resolution proceeds in the same way as an ordinary ASCII string

# Client Software Still Maturing

- Not all browsers handle IDN strings in the same way
- Email protocol for IDNs still in definition
- Beyond desktop software, need to consider behavior of mobile and embedded clients
- Key is early testing and awareness
  - Many regional Asian efforts in this regard
  - ICANN has contributed through ICANN wiki “.test” effort





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.פּערשױב	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

Windows Internet Explorer - IDNwiki - IDNwiki

http://idn.icann.org/

File Edit View Favorites Tools Help

IDNwiki - IDNwiki

article discussion view source history

## IDNwiki

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

### 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?

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

# IDN Confusability

- Not a new topic
  - “0” looks like “o” and “1” looks like “l”
- Increased problem with increased number of characters in use, and length of strings
  - 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

- IDN tables are developed by registry managers to
  - Inform users which 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 that have similar appearance (e.g. Cyrillic, Greek, Latin)

# IDNA Protocol Revision

- IDNA provides the technical specification 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
  - Progress expected at the IETF meeting next week

# IDN TLDs – issues remain

- Request from many that multiple “variant” top level domain names behave as one
  - A variant is where two strings use different characters, but are the same words
    - for example: Pakistan: پاکستان and پاکستان
    - no existing technical solution to this at root level
  - May be addressed through name reservation or blocking; local technical solutions
- Number of characters in a label
  - ccTLDs 2 char labels → 2+char
  - gTLDs 3+char → 1char, 2char, 3char, and 3+char ?
  - Usability and acceptability across applications

# IDN Processes at ICANN

*IDNs essential to two aspects of ICANN's work*

- IDN ccTLDs (IDN country code top level domains)
  - Country code registries (e.g. .cn, .in, .eg) may want to offer IDN versions
  - Particularly pressing in parts of the world that don't use Latin characters
- New gTLDs
  - New IDN names are envisioned to be available at the launch of the new gTLD process

# Current IDN Processes at ICANN

## Implementation: IDN ccTLDs Fast Track Process

- to introduce a limited number of IDN ccTLDs
- non-Latin scripts only, matching ISO3166 list
- will 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

# IDN ccTLD Fast Track: Key Remaining Issues



- Technical concerns mentioned previously:
  - IDNA protocol finalization
  - Handling of variant strings
  - Issues associated with the number of characters in a name
- Broad agreement by all on need to adhere to technical standards in this still-evolving area; form of agreement not yet decided
- Suitable, fair mechanism for cost recovery



# IDN ccTLD Fast Track Progress

- Active project for past 18 months at ICANN
  - Active engagement with many potential requestors, several drafts of plan, community consultation
- Operational elements and processing details to be finalized by Q4 2009
  - Goal is to launch Fast Track process by year end
  - Like new gTLD process, will require community support and Board approval

# IDNs in New gTLDs

- Shares same technical issues as the IDN ccTLD Fast Track program
- Will launch along with overall new gTLD program

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

ICANN is working to support that objective