

ICANN 33



# INTERNATIONALIZED DOMAIN NAMES

## - A BASIC INTRODUCTION

2 November 2008

Tina Dam, Director, IDN Program

[tina.dam@icann.org](mailto:tina.dam@icann.org)

باهر عصمت، مدير علاقات إقليمية – الشرق الأوسط

[baher.esmat@icann.org](mailto:baher.esmat@icann.org)

# Agenda

- IDN definitions and basics
- How does IDNs work
  - including examples of applications and their various implementations
- IDN wiki facility
- IDN TLD introduction processes
- IDN TLD technical string requirements



# Characters in the DNS

- Search on "*US-ASCII character set*"
- The DNS can handle all US-ASCII characters
  - Examples:
    - (a...z), (o...g), (-)
    - ( ) 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)
    - (-)
- That was before internationalization...



# IDN Definitions

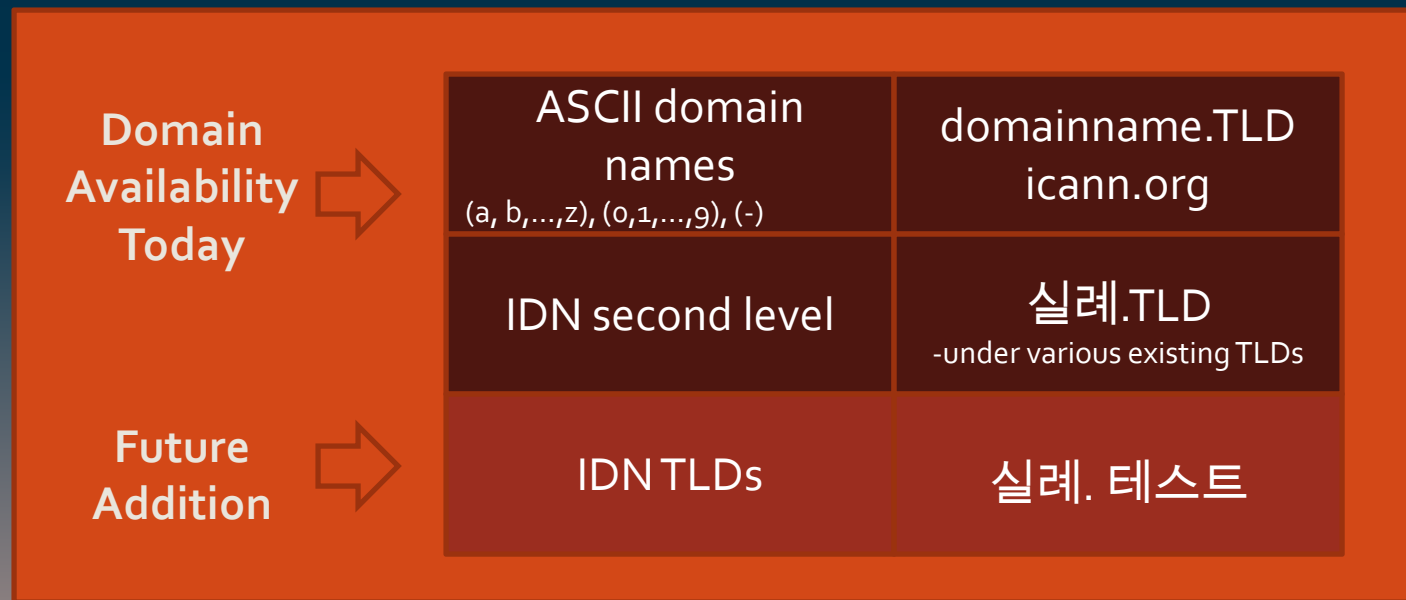
- Internationalized domain names are:
  - Names with characters other than the standard ASCII (a,b,...z), (0,1,...9), (-)
- Example:

xn--ngbrx4e.tld ← tld.عربي
- IDNs are about localized solutions
  - But need to be 'internationalized' due to the global nature of the Internet



# IDNs - What we have / what we need

- IDNs have existed as second level since 2003
  - under web protocol standards
  - email protocol standards are underway (IETF)
- We also need IDN TLDs
  - 北京.中国; [xn--1lq9goi.xn--fiQs8S]



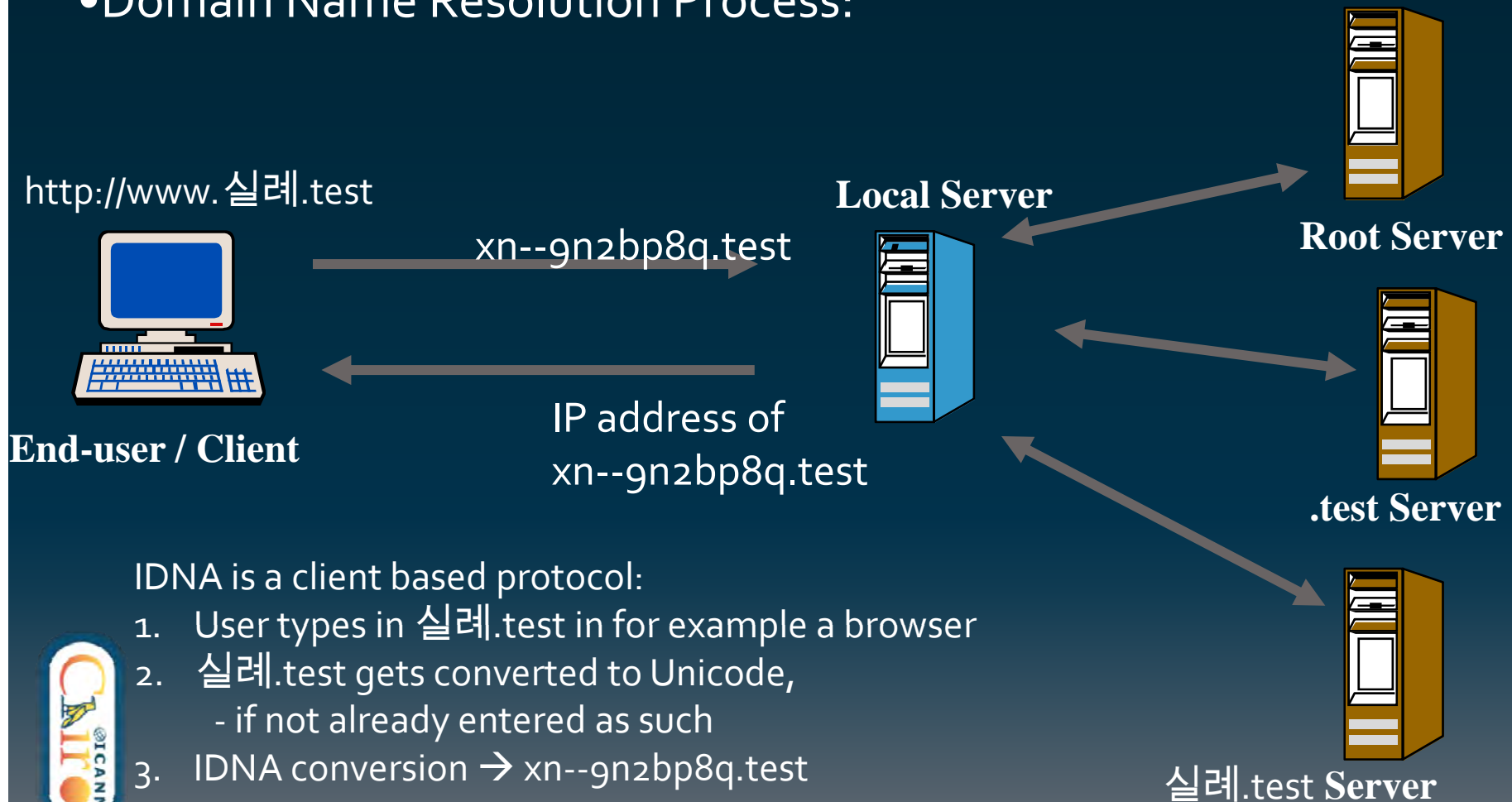
# Why Internationalization?

- 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:



IDNA is a client based protocol:

1. User types in 실례.test in for example a browser
2. 실례.test gets converted to Unicode,  
- if not already entered as such
3. IDNA conversion → xn--gn2bp8q.test





Internet Explorer - IDNwiki - Windows Internet Explorer

http://www.idnwiki.com/%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

http://www.idnwiki.com/%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

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

Mozilla Firefox - IDNwiki - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.idnwiki.com/%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: فرسالنهر.tld → xn--mgbtbg2evaoui.tld
- However, there are exceptions:
  - xn--gibberish - decodes into the Arabic characters ب٨٧٩فأ
  - 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



# IDNA protocol and IDN wiki – try it out

- <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 at <http://idn.icann.org>

The screenshot shows the IDNwiki website interface. At the top, there's a navigation bar with tabs for 'article', 'discussion', 'view source', and 'history'. The main heading is 'IDNwiki' with a sub-heading 'Welcome to the IDN TLD evaluation gateway!'. Below this, there's a section for 'Contents' with a list of numbered items: 1 Introduction, 2 Your participation is important!, 3 Limited evaluation period, 4 Things to test, 5 Further information about the IDNwiki, and 6 The example.test names. The left sidebar contains a 'navigation' section with links to the main page, email test, basic concepts, technical topics, and software. Below that is an 'interaction' section with links for recent changes, contact, and wiki help. The 'evaluation' section lists various languages in their native scripts: العربية, 简体中文, 繁體中文, Ελληνικά, हिन्दी, 日本語, 한국어, فارسی, Русский, தமிழ், and עברית. At the bottom, there's an 'additional' section.

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	בבש.לפישׂב	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--hlej6aya9esc7a

## 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 could not 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
- 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



# IDN TLD Introduction Processes

## Implementation of Country-code IDN TLDs – Fast Track

- implementation of recommended policy from IDNC WG
- review of contractual elements and relationship to IDN TLD operators
- review of technical requirements for IDN TLD 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

# IDN TLD Technical Requirements I

- The label must be a valid internationalized domain name, as specified in technical standards <http://www.icann.org/en/topics/idn/rfcs.htm>. This includes the following, non-exhaustive, list of limitations:
  - Must only contain Unicode code points that are defined as “Valid” in *The Unicode Codepoints and IDNA* (Internet Draft “draft-faltstrom-idnabis-tables”)
  - Must be fully Normalization Form C compliant, as described in *Unicode Standard Annex #15: Unicode Normalization Forms*.
  - Must not contain characters with a mixture of directionality properties.
  - Must not contain any leading or trailing digits (of any kind of digits).



# IDN TLD Technical Requirements II

- The label must meet the relevant criteria of the ICANN *Guidelines for the Implementation of Internationalized Domain Names*. This includes the following, non-exhaustive, list of limitations:
  - All code points in a single label will be taken from the same script as determined by the *Unicode Standard Annex #24: Unicode Script Property*.
  - Exceptions are permissible for languages with established orthographies and conventions that require the commingled use of multiple scripts. However, even in the case of this exception, visually confusable characters from different scripts will not be allowed to co-exist in a single set of permissible code points unless a corresponding policy and character table is clearly defined.





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



شكراً لكم

tina.dam@icann.org

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

