

Internationalized Domain Names - a primer

**APTLD Meeting
Bangkok, Thailand
13 November 2006**

**Tina Dam
IDN Program Director
ICANN**



Agenda

- What is an Internationalized Domain Name
- The Need for Internationalization
- Internationalization Overview
- IDNA – Protocol Functionality
- IDN Working Groups
- User Confusion and IDN Guidelines
- A range of other IDN relevant topics

What is an IDN?

- IDN stands for Internationalized Domain Name
 - Domain name labels that represent names containing non-host name characters.
 - Valid hostname characters are: a-z, 0-9, “-”
 - Only host name (or “LDH”) strings are entered into the DNS
 - Sometimes, IDN is used to refer to a fully-qualified domain name that contains at least one non-IDN label
 - IDN in general refers to both displayed form (Unicode) and stored form (punycode) of the domain name
- Example: rødgrød.tld → xn--rdgrd-vuad.tld
- IDN is also sometimes used to refer to other ways of internationalization or localization
 - “Keywords”,
 - Special searching or directory mechanisms, etc.

What is an IDN – User Perspective

- All you need is the name you want to register
- Registries will supply a list over available characters, usually in Unicode
- Registries will handle all encodings needed during registration process
- Example: `فرسالنهر.tld` → `xn--mgbtbg2evaoi.tld`
- May be useful to consider usability of the name, keyboards, business cards, and other practical limitations
- Encodings by for example:
 - <http://josefsson.org/idn.php>
 - VeriSign IDN conversion tool
 - <http://mct.verisign-grs.com/index.shtml>

The Need for IDNs and Internationalization

- Geographic expansion of the Internet
 - IDNs match needs of increased use by linguistic groups
 - IDNs used for identification of content reflecting linguistic diversity
- Localized system adapted to
 - Language
 - Writing system and character codes
 - Location
 - Interests
- Internationalization is
 - A means to localization
 - Necessary given the global nature of the Internet
- Global Interoperability
 - Network strength is to interoperate globally
 - Security and stability is primary focus
 - Avoid fragmentation of the Internet

Internationalization Overview

Domain Names Based on
ASCII / LDH Rule

- IDN Second level
- Internationalized top level

ASCII based browser/email
clients/...

- Web access in local
languages + IDN emails

Content have been available
in many languages for
some time

- Expected to continue to
expand

example.test → 실례.test and 실례.테스트

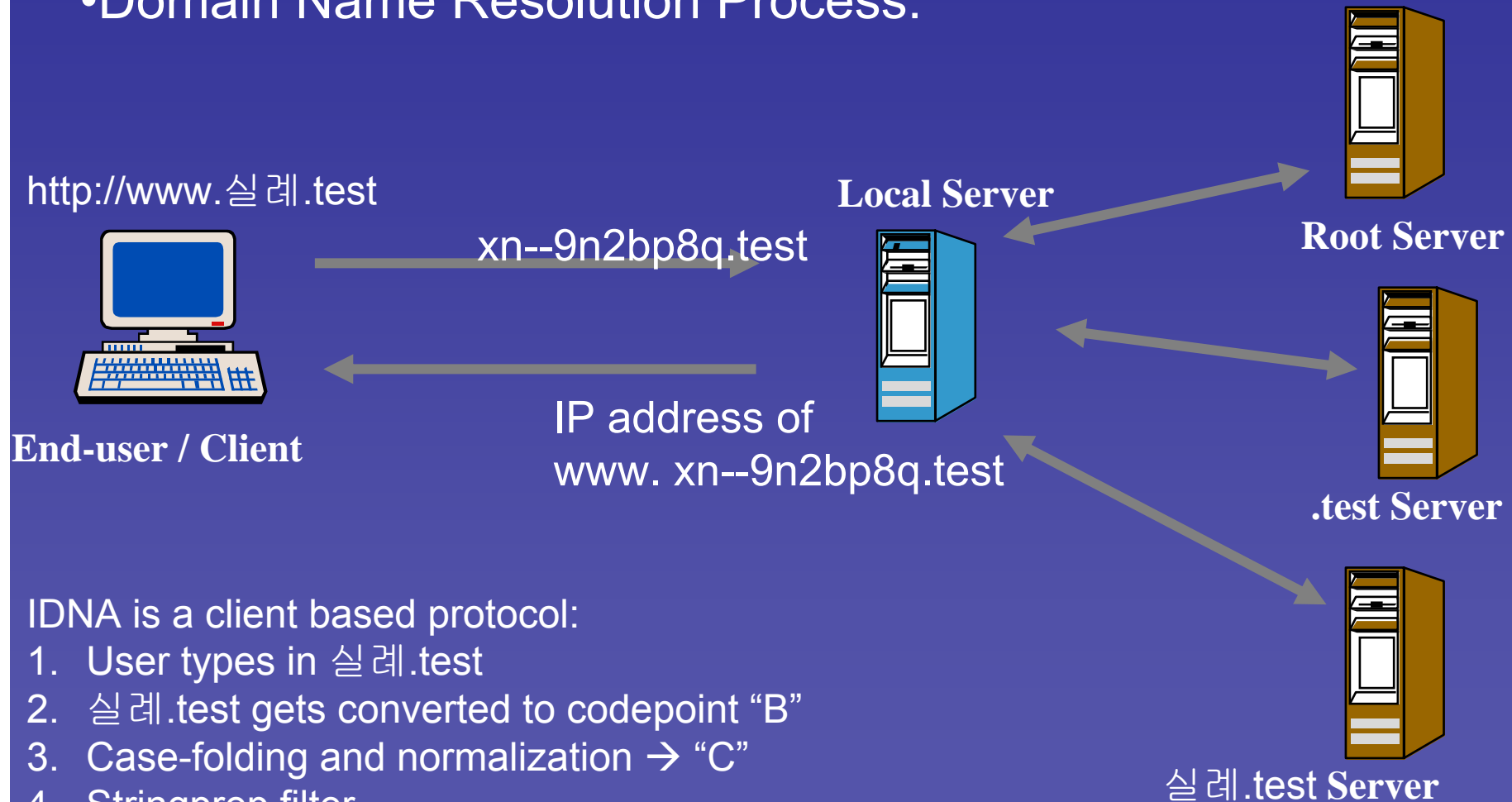
(stored form: example.test → xn--9n2bp8q.test and xn--9n2bp8q.xn--9t4b11yi5a)

Aim: An internationalized Internet



IDNA – Protocol Functionality

•Domain Name Resolution Process:



IDNA is a client based protocol:

1. User types in 실례.test
2. 실례.test gets converted to codepoint “B”
3. Case-folding and normalization → “C”
4. Stringprep filter
5. Punycode conversion → xn--9n2bp8q.test
6. (Registry prohibition list)

IDN Working Groups and Activities

- ICANN President's Advisory Committee for IDNs
 - Formed on 23 November 2005
 - Initially tasked with IDN TLD Technical Challenges
- IDN TLD Registries
 - ccNSO and gTLD reps focused on IDN Guidelines
- IETF and IAB
- Browser Community
- GNSO, ccNSO, GAC, ALAC
- Regional Model Initiatives

User Confusion and Spoofing Issues

- IDNs Expanding Risk of Known Problems
- Many characters can be confused with others
 - Problem exists in ASCII as well
 - Digit “1” and lower-case “l”
 - Digit “0” and upper-case “O”
 - IDNs increasing the character collection
 - From 64 in ASCII (LDH)
 - To tens of thousands in Unicode (nameprep)
- This kind of confusion
 - create opportunities for user mistakes
 - and fraud

Examples Visual Confusion

- Well-known example: paypal.com
 - Second character is U+0430, Cyrillic small a
 - Looks like Roman/ASCII “a”
 - Would have been prevented by “one label, one script” rule
- Other example:
 - Russian ccTLD is .ru
 - Cyrillic “r” and “u” is: py
 - .py (in latin) is ccTLD for Paraguay
 - Note: Russia did not ask for .py, this is just an example
 - Process needed to determine labels matching ccTLDs

Project: Revision of IDN Guidelines

- IDN TLD Working Group
 - ccNSO and gTLD Constituency membership
 - Monthly meetings through to Morocco, bi-weekly toward Sao Paulo
 - Created version 2.0, 2.1, and pending further revisions
 - Provide regular updates to the Presidents IDN Advisory Committee
- Version 1.0
 - Commitment by .cn, .jp, .tw, .info, and .org
 - Further authorized and implemented by .info, .museum, .biz, .org, .name, .com and .net
- Version 2.0
 - Opened for script based implementation in addition to language based
 - No mixing of scripts within labels
- Version 2.1
 - Included the provision in relation to the “UseSTD3ASCIIRules” flag
- Future Version:
 - Focus on IDN TLD principles & BCP practices to avoid further confusion between TLD implementations

Some user confusion is being solved by

- protocol adjustments
- IDN guidelines

Remaining user confusion need to be solved by

- implementation of adequate registry policies
- education of community

IDN Policy Development

- Policy development is run by ccNSO, GNSO, GAC
- Multiple meetings during and in-between ICANN meetings
- Latest face-to-face GNSO meeting in Amsterdam:
 - IDN policy issues combined with introduction of new gTLDs
 - IDN issues report and ToR are available online for comments
 - Proposes joint task force between GNSO, ccNSO, GAC to prioritize issues and propose structure for further work
 - Joint meeting planned for Sao Paulo
- Introduction of new gTLDs are focused on:
 - Whether to introduce new gTLDs
 - Allocation methods
 - Selection criteria
 - Contractual Conditions

Internationalized TLD Principles

- Global uniqueness and interoperability of the DNS
 - unique and unambiguous domain names
 - Same functionality regardless of geographic placement of access
 - URLs and emails connect as expected regardless of geographic placement of access
- Promote “Future-Proof” solutions
 - Define Unicode characters to be allowed
 - Provides ability for adding new languages, new characters far in the future
- Avoid or diminish as much as possible user confusion
 - Technical limitations
 - Implementation requirements
 - Registry restricted list and policies
 - User education
- Promote multi-stakeholder involvement

Range of Other Relevant IDN Topics:

- IDN Technical testing – (later today)
- IDNA Protocol revision – (later today)
- Impact of reviewed changes
- DNAME as resource record
- Any new kind of resource record
- IANA repository for IDN tables
- Language-script Specific Issues
- Dispute issues and rules
- Statistics on # of registrations at second level
- Much more....

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

