

Internationalized Domain Names

Business Access Meeting



Tina Dam
ICANN
Director, IDN Program
tina.dam@icann.org

Paris, France
23 June 2008

Agenda

- Definitions and basics
- How does IDNs work
- What work still needs to be done?
- Confusability Issues
- Summary

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)
 - (-)
- That was before internationalization....

IDN Definitions

- Internationalized domain names are:
 - Names with characters other than the standard ASCII (a,b,...z), (0,1,...9), (-)
- IDNs are about localized solutions
 - But need to be ‘internationalized’ due to the global nature of the Internet
- 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--1lq90i.xn--fiQs8S]

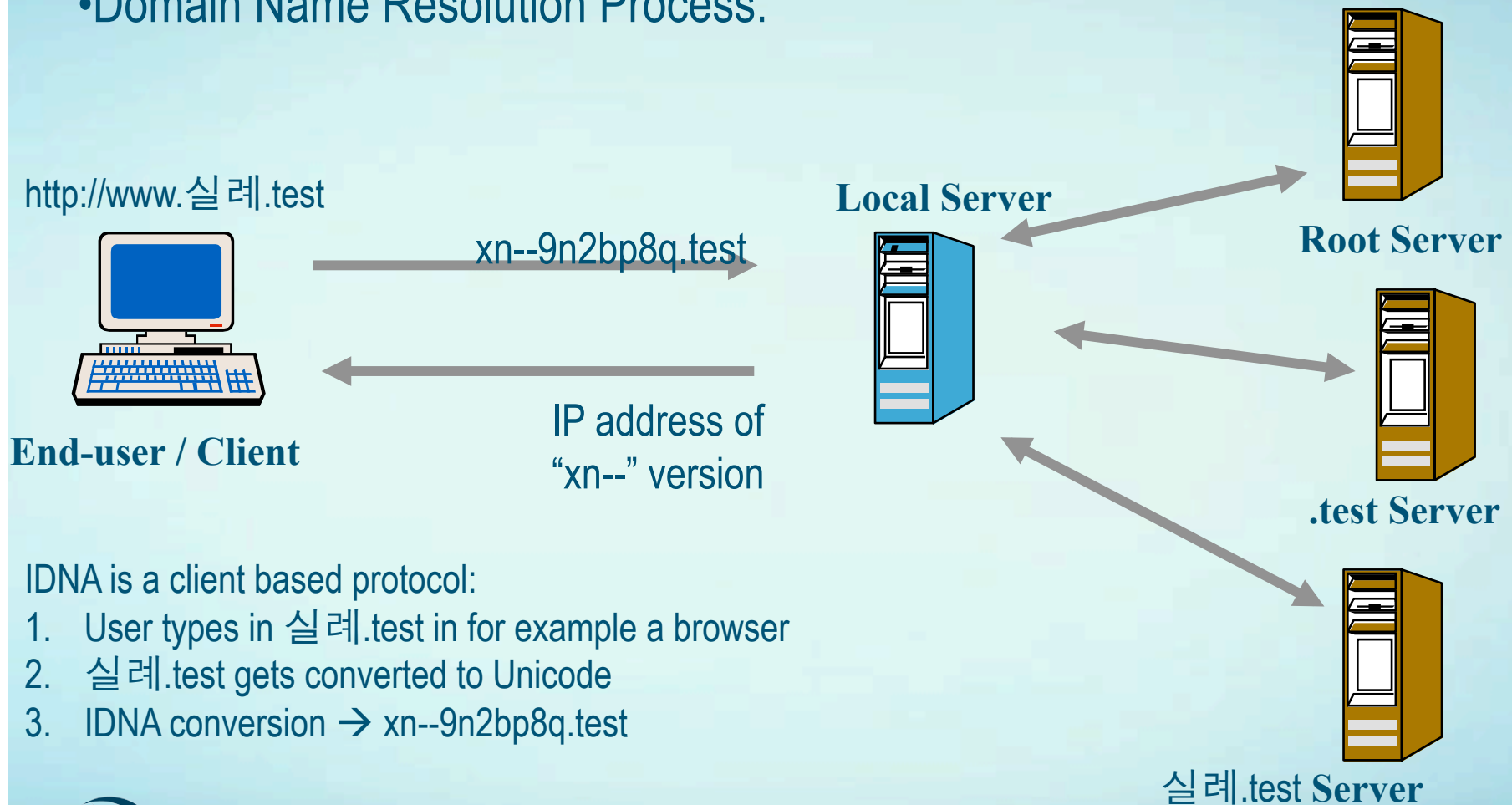
Why Internationalization?

- 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
 - SLD IDN registration available across many TLDs
 - 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
3. IDNA conversion → xn--9n2bp8q.test



IDNs only work if the application software works with IDNs

- standard implementation is important to ensure secure global user experience

- today browser developers have implemented IDNA differently

Internet Explorer - IDNwiki - Windows Internet Explorer

http://مثال.إختبار/%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

متال.إختبار/الصفحة_الرئيسية http://

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

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
- Usually the stored form usually gives no meaning
 - Example: `فارسالنهر.tld` → `xn--mgbtbg2evaoi.tld`
- However, there are exceptions:
 - `xn--gibberish` - decodes into the Arabic characters `بببب٨٧٩ب`
 - `xn--trademark` - with different versions of trademarks
 - This is coincidentally and hence not intentionally
- `xn--` prefix indicates to application software that the label needs to be decoded back into Unicode for proper display to the user



IDNA protocol – 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
- Try copy / paste between applications

Why are we not there yet?

- Initial registration availability resulted in
 - visual confusion issues
 - damaging uniqueness principle of the DNS
- Different words spelled identically
 - cap (cyrillic) is not homograph to cap (latin) but it is confusingly graphic similar
 - paypal.com (cyrillic a's) and paypal.com (latin a's)
 - They appear to be identical but are not



Why are we not there yet?

- display of **xn--mgbh0fb** instead of مثال
- display of **xn--mgb0dgl27d** instead of ايكوم
- display of **xn--1lqs71d** instead of 東京
- display of **xn--1lq90i** instead of 北京

→ Results in trademarks being displayed

where

the U-label version that was registered may be a different trademark

- more user confusion and fraud opportunity
 - Registration of microsoft.<tld> ?
- Protocol implementation experience and review showed other problems...



Towards IDN TLDs:

- What still needs to be done?

- IDN wiki – test facility
- IDN TLD processes at ICANN and IANA
- Main policy related question from users
- IDNA protocol revision at IETF
- IDN Guidelines

IDN wiki at <http://idn.icann.org>

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--jxalpdip
Hangul	Korean	실례.테스트	xn--9n2bp8q	xn--9t4b11yi5a
Hebrew	Yiddish	בבש.ה'פשיב	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

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
- 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

Status of the .test wiki

- Adding new “features”
 - New languages:
 - Amharic (4th level under idn.icann.org)
 - Hebrew (2nd level under existing TLD)
 - Thai (4th level)
 - Urdu (4th level)
 - process available for others languages
 - DNSSEC signing the zones
 - exploring IDN email addition
 - experimental status of technical standard



IDNwiki Access

- IDNwiki can be accessed at:
 - <http://idn.icann.org>
 - Thanks to users, moderators, and wiki staff,
 - includes useful information about IDN
 - in all available languages
 - Please:
 - Visit the site
 - Try IDNs
 - Report on results
 - Use information
 - Add information

Status of the IDNA revision

- Proposed revision at IETF
 - from extremely hard working participants
 - RFC4690 and associated internet drafts suggesting revisions and solutions to some problems
- Unicode version independent
 - Three categories by procedure not table
 - Protocol-valid (some w/ contextual rules)
 - Disallowed
 - Unassigned
- Attempting to plan for educational sessions on the difference between protocol versions

Status of IDNA revision

Basis in RFC4690, describing issues

- [draft-klensin-idnabis-issues-07.txt](#)
 - overall rationale and explanation
- [draft-klensin-idnabis-protocol-04.txt](#)
 - registration vs. resolution
- [draft-faltstrom-idnabis-tables-05.txt](#)
 - category operations procedure
 - not table, but holds Unicode-5.0 result for reference
- [draft-alvestrand-idna-bidi-04.txt](#)
 - to allow combining marks at end of string, by test



IANA management of IDN TLDs

- Process for insertion of IDN TLDs in root
 - exists for test domains only (IDN .test)
 - Developed w/ RSSAC & SSAC recommendations
 - need review, revision, and implementation
 - includes emergency removal procedure
 - for test IDN TLDs only
 - Initial review scheduled by IANA staff
 - Before 30 June 2008
 - Result to be published publicly

ICANN TLD Allocation Processes

Country-code IDN TLDs – Fast Track

- Deploy non-contentious ccTLD equivalents quickly
- Where demand/readiness exists
- Don't wait for full ccNSO PDP
- Tomorrow, half day session on policy issues and methodology

Country-code IDN TLDs – Long Term

- Full policy that caters for all
- Follows the full ccNSO Policy Development Process
- Issues paper scheduled for publication by ICANN Paris meeting (Jun08)

New Generic TLDs

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

If I have registered [idn.tld] then will I also be the registrant of [idn.idn-tld] ?

- Policy considerations
 - related to IP rights vs. competition options
 - difficult to do meaningful translation of existing strings
- GNSO Policy:
 - No precedence for existing registries
 - Objection rights exists for confusingly similarity
- ccTLD operators and GAC are considering needs for IDN ccTLD aliasing solution
- If requested then technical solution is needed



IDN Guidelines

- ICANN IDN Guidelines
 - Need revision to follow IDNA revision
 - Developed by ccTLD and gTLD registry operators
- Local community guidelines, for example:
 - Informal ‘Arabic script’ meetings in Dubai
 - language experts participate in IDNA review
 - Review of characters that are valid per the protocol
 - In addition to protocol, a need for:
 - local registration policies and variant tables

Summary

- Protocol – general validity of character and in some extend string validity
- IDN Guidelines – implementation of protocol, and global rules for scripts and languages
- Local Directions – local rules for scripts and languages

→→→→

- Registry testing and implementations
- Application testing and implementations
- Allocation and Delegation Procedures
- Apps Developer, Registry, Registrar, Reseller, Registrant, User:





**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>