Internationalized Domain Names

RIPE NCC Dubai 7-9 Dec 2003

James Seng jseng@pobox.org.sg (former) co-chair, IDN Working Group

Internationalized Domain Names

华人.公司.cn 華人.商業.tw 高島屋.会社.jp 삼성.회사.kr 三星.회사.kr الاهرامم viagénie.qc.ca ישראל.קום ทีเอชนิค.พาณิชย์.ไทย

現代.com ヤフー.com

IETF IDN Working Group

- IETF is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet.
- It is open to any interested individual.
- From the IDN WG charter "The goal of the group is to specify the requirements for internationalized access to domain names and to specify a standards track protocol based on the requirements."

A brief history of time...

- 04/98 IRDU initiated iDNS pilot project
- 07/98 APNG initiated iDNS Working Group
- 02/99 APNG iDNS Testbed
- 06/99 INET'99 progress report
- 09/99 Japan IDN WG formed in Aizu
- 11/99 IDN BoF at IETF#46
- 01/00 IDN WG formed to work on requirements
- 03/00 IETF#47 First requirements draft
- 06/00 IETF#48 Re-chartered to include standard protocol, ~13 Internet Drafts
- 07/00 (JET formed in Yokohama)
- 08/00 Three Design Teams formed, handling
- 11/00 ICANN IDN Workshop

A brief history of time...

- 12/00 IETF#49 ~26 Internet Drafts
- 03/01 IETF#50 ~30 Internet Drafts
- 08/01 IETF#51 ~13 Internet Drafts
- 12/01 IETF#52 ~9 Internet Drafts
- 01/02 WG Last Call for 4 Internet Drafts
- 02/02 submit the I-Ds to IESG for RFC publication
- 06/02 IESG initiated IETF-wide last call
- 07/02 Another WG Last Call on 2 Technical change
- 10/02 IESG approves the publication of the
 - 4 Internet Drafts
- 10/02 ICANN IDN Workshop
- 03/03 IDN RFCs published by IETF
- 06/03 ICANN announced IDN Guidelines

What's so difficult?

Where is IDN?



Snapshot of network layers to provide some Internet Service

Stability and Compatibility

- From the charter "A fundamental requirement in this work is to not disturb the current use and operation of the domain name system, and for the DNS to continue to allow any system anywhere to resolve any domain name."
- IDN is changing a fundamental service
 - IAB RFC2825 "Issues of I18N, Domain Names, and the other Internet protocols"
 - IAB RFC2826 "Unique Root"
- A lot of existing Internet Protocols assumed domain names and assumed that it is A-Z, 0-9 and "-" only

Confusion

Language and Script

- A language is a way that human interact
- A script is the written form of a language
- Many written languages share the same script
- Some written languages use more than one script

 Example: If "ø.se" is registered in Sweden, should that imply "ö.se" is also registered? Or blocked?

Confusion

Name and Identifier

- Name is a word or phrase that constitutes the distinctive designation of a person or thing
- Identifier is a string of characters that uniquely identify a person or thing
- o Example
 - "James Seng" is a Name
 - "jseng" is an Identifier
 - is jseng.com a Name or Identifer?

Domain Name is an Identifier not a Name

Confusion

- Internationalization, Localization and Multilingualism
 - Internationalizing make the protocol able to handle more scripts
 - Localization involves tailoring interaction with users in the languages they know
 - Multilingualisam makes the protocol able to handle multiple languages

• "I" in IDN is Internationalization

Problems

Encoding

- Using one universal character set or multiples caracters set?
- What encoding to use? UTF-8? UTF-16?
- Matching
 - yahoo.com = YAHOO.com
 - ø.se = ö.se

Local issues

Language specific consideration

IETF RFCs

Punycode RFC 3492

- Designed for use with Internationalized Domain Names
- It uniquely and reversibly transforms a Unicode string into an LDH string (Letter, Digit, Hyphen) (aka ASCII Compatible Encoding – ACE)
- Some compression to produce shorter string
- "xn--" prefix has been selected by IANA
- e.g. fältström => xn--fltstrm-5wa1o

Stringprep/Nameprep RFC 3454 & RFC 3491

- Prepare internationalized domain name labels in order to increase the likelihood that name input and name comparison work in ways that make sense for typical users throughout the world.
- Based on UTR#15 (Normalization) & UTR#22 (Case Mapping)
- Stringprep is the generic processing
- Nameprep is a profile of stringprep for Internationalized Domain Names
- **Based on Unicode 3.2** (not 4.0)



 IDNA only upgrade in applications to handle IDN

- Consideration of legacy encoding and interopability
- Enforce Nameprep in applications

Uses Nameprep-ACE'ed IDN over the wire

IDNA-Nameprep-Punycode



Encoding

Input from user
 Fa["]Itström

O 2. Apply Nameprep
fältström

3. Apply Punycode
 xn--fltstrm-5wa1o

IDNA-Nameprep-Punycode



Implications

 Two different strings in Unicode might be "equal" according to the rules

• Example:

- Fältström and fa["]ltström
- xn--fltstrm-5wa1o
- Today Faltstrom and faltstrom are equal

• IDNA does not change DNS rules

Implications

- Two strings "looking" the same might be different Unicode strings and different strings according to the rules
- Example:
 - CYRILLIC SMALL LETTER IE (U+0435) : e
 - LATIN SMALL LETTER E (U+0065) : e
- Spot any two differences between the two "e"s?

More Implications...

 What is "domain name" and what is in zone file are two different things

What registrant wanted to register Fältström.se



What's in the zonefile xn--fltstrm-5wa10.se

What one get which decode the domain name fältström.se



Questions and Answers?



What we have so far



CJK Administration Guideline draft-jseng-idn-admin-05.txt Policy \circ JET effort (started in Feb 2002) delegation to re**tuce pote**r The goal dispute by reservat variants **RRP/EPP** registration • Provide an algorithm for the variants but allow zoneeadminto define further rules how to bridge echnical res Other Protocols ection. to IESG for pu

Internationalized Email Address

- Internationalized Mail Address in Applications (IMAA)
 - draft-hoffman-imaa
- Internationalization of Email Address
 Address
 - draft-klensin-emailaddr-i18n



Internationalized Resource Identifier draft-duerst-iri Policy erioa S \$ registratio Define a new protocol element resolution ??? Other Protocols

Some Suggestions...

Policy

Understand the technical limitation

- Script vs. Language
- Name vs. Identifier
- Internationalization vs. Localization
- Per label basis

Understand what users wants

- Script vs. Language
- Name vs. Identifier
- Internationalization vs. Localization
- Per "FQDN" (whole domain name) basis

Some Stogtges Softsvare ...



Software Libraries

 GNU IDN Library - implementation of Stringprep, Punycode and IDNA specifications released as LGPL.

http://www.gnu.org/software/libidn/

 International Components for Unicode (ICU) Libraries.

http://oss.software.ibm.com/icu/

 JPNIC idnkit - IDN processing library and patches/utilities released under a BSD-like license.

http://www.nic.ad.jp/ja/idn/idnkit/download/index.html



Software Libraries

• Paul Hoffman's Perl Libraries.

http://www.imc.org/idna/

 Verisign IDN Software Development Kit (C & Java).

http://www.verisign-grs.com/idn/sdk_download.html

 Python - version 2.3 has IDNA support built into the language. http://www.python.org/doc/current/lib/module-

encodings.idna.html



IDNA Clients

- Mozilla browser IDN support is released to the public since version 1.4. (Netscape 7.0) http://devedge.netscape.com/viewsource/2003/idn/
- Konqueror browser uses GNU IDN Library to support IDN. May be available in version 3.2?

http://devedge.netscape.com/viewsource/2003/idn/

 Verisign i-Nav - Internet Explorer plug-in that provides IDN support.

http://www.idnnow.com/index.jsp



IDN Open Source (IDN-OSS)

- Conceptualized during ICANN Montreal Jun 03; Launch Nov 2003
- IDN Open Source Software (IDN-OSS) project to develop quality RFC-compliant software to support IDN capabilities in different applications (web browsers, email and instant messaging clients, etc) running on all major operating systems (Windows, Mac OS X, Unix etc).
- We believe that contributing towards a common IDN-OSS platform which can be used by all registries and all users will be a cheaper and more effective alternative, compared to in-house development or licensing.

ISC and Advisory Council

- Internet Software Consortium (ISC) has agreed to host the IDN-OSS project.
- The Advisory Council will be set up as the body to advise the Project Convener on all technical, legal, development, and administrative issues with the IDN-OSS project. The initial Advisory Council includes:
 - Vint Cerf
 - Mark Davis
 - Martin Dürst
 - John Klensin
 - Paul Hoffman

What can you do help?

- If you are an organization that stands to benefit from IDN-OSS, we encourage you donate some money to the project.
- If you are a programmer, we need you! We need programmers to contribute codes towards the project.
- Even if you are just a user, you can help in translating the software to your own language, you can help to distribute the software and to file bugs reports.
- In short: **We need people!**

Contact Information

News on IDN: <u>http://idn.isc.org/</u>
IDN Wiki: <u>http://idn.isc.org/wiki/wiki.cgi</u>

Contact

Funding & ISC matters
 Paul Vixie <paul@vix.com>

Technical

- William Tan <wil@dready.org>
- All other matters
 - o James Seng <jseng@pobox.org.sg>





