

# Introduction to Internationalized Domain Names (IDN)

IP Symposium for CEE, CIS and Baltic States  
Moscow, Russia  
16-19 September 2003

Robert Shaw

<robert.shaw@itu.int>

ITU Internet Strategy and Policy Advisor

# Agenda

- Background on ITU involvement
- Definition of DNS and IDN
- IDN History
- IETF Technical Solution
- Administration and Policy Directions
- National Experiences
- Tools & ITU Resources
- Some Reflections
- Future ITU Activities

# Background on ITU Involvement

- ITU Member States adopted two resolutions in 2002 related to Internet names and addresses guiding ITU's activities in this area:
  - Resolution 102: “Management of Internet Domain Names and Addresses”
  - Resolution 133: “Role of administrations of Member States in the management of internationalized domain names”
    - “to promote effectively the role of Member States in the internationalization of domain names and address of their respective languages”
  - Resolutions give instructions to Secretary-General, TSB Director, the BDT Director and the ITU Council, as well as inviting Member States to contribute to certain activities
  - See [www.itu.int/osg/spu/mina/](http://www.itu.int/osg/spu/mina/)

# What we hope to accomplish from our related activities...

- Bring together experts so that they can share experiences for the benefit of others
- Document and build knowledge base of materials available to ITU Member States
- Give snapshot of some of the ongoing national activities and their implementation experiences
- Discuss role of national administrations of ITU Member States and possible policy role they may wish to consider
- Discuss further cooperative measures at both regional and international levels, particularly with regard to assisting developing countries in their consideration of these new technologies

# The DNS Is...

- The “Domain Name System”
- What people use to refer to computers by name on the Internet
- The mechanism by which Internet software translates names to addresses and vice versa
- A globally distributed, loosely coherent, scalable, reliable, dynamic database
- The only database system that has been successfully deployed Internet-wide
- Which also makes it hard to change...

# DNS History

- Created in 1983 by Paul Mockapetris to address maintenance problems with the Internet hosts database, fondly remembered as HOSTS.TXT.
- Originally defined in IETF RFCs 1034 and 1035, then extended by numerous subsequent RFCs.
  - RFC stands for Request for Comments
  - Standards for Internet protocols are documented by RFCs
    - Not all Internet protocols have RFCs
    - Not all RFCs define standards
- Restricted for ~20 years to case-insensitive ASCII letters (a-z), digits (0-9) and hyphen (LDH)

# Names versus Addresses

- An address is how you get to (route) to a network node
  - Often hierarchical, which helps with scaling
    - Robert Shaw, ITU, Place des Nations, 1211 Geneva 20, Switzerland
    - 156.106.130.32
- A name is how a node is referenced
  - Hierarchical name structures can help scaling
    - recipes: chocolate: cookies
    - C:\WINDOWS\system32\drivers\disdn\
    - [www.itu.int](http://www.itu.int)
- Telephone numbers have aspects of both names and addresses

# DNS is a Database

- Keys to the database are “domain names”
  - [www.itu.int](http://www.itu.int), [www.aptsec.org](http://www.aptsec.org), 1.4.e164.arpa
- Over 100,000,000 domain names are now stored
- Each domain name contains one or more attributes, known as “resource records”
- Each attribute is individually retrievable



# Global Distribution

- Data is maintained locally, but retrievable globally
  - No single computer has all DNS data
- DNS lookups can be performed by any Internet-connected device
- Remote DNS data is locally cached to improve performance

# Demand for Multilingualism

- For example, largest percentage of Internet users are now in the Asia-Pacific region
- Consequence of the Internet “globalization” is growing number of users not familiar with ASCII
- Domain names in ASCII characters poses linguistic barriers
- Native speakers of Arabic, Chinese, Japanese, Korean, Russian, Tamil, Thai and others who use non-ASCII scripts at disadvantage
- Requirement for “internationalization” of the Internet’s Domain Name System

# IDN is...

- Abbreviation for “Internationalized domain name”
- Refers to a domain name where one or more characters not in historical subset of Latin LDH set (a-z), digits (0-9) and hyphen (LDH) used in the DNS
- Associated with Unicode (ISO 10646)-based labels
- Major transition from 38 characters to more than tens of thousands possible Unicode “code points”

# “Unicode” Examples

- Arabic (Arabic) یونیکوڈ
- Arabic (Persian) یونی کڈ
- Armenian Յւնիկոդ
- Bengali য়ুনিকোড
- Cyrillic (Russian) Юникод
- Devanagari (Hindi) यूनिकोड
- Georgian უნიკოდ
- Greek Γιουνικοντ
- Gujarati યૂનિકોડ
- Gurmukhi ਗੁਰਮੁਕੀ

- Han (Chinese) 统一码  
統一碼  
万国码  
萬國碼
- Hangeul 유니코드
- Hebrew תַּיִקוּד
- Hiragana ゆにこおど
- Khmer អ៊ូនីកូដ
- Malayalam യൂനികോഡ്
- Syriac ܝܘܢܝܩܘܕܝܬܐ
- Tamil யூனிகோட்
- Thai ยูนิโคต

# Some IDN History

- Late 1990s
  - Multilingual domain names first developed at National University of Singapore
- July 1998
  - Asia Pacific Networking Group
  - iDNS Working group : development of the experimental implementation of an Internationalized multilingual multiscrypt Domain Names Service
  - iDomain Working Group : creation of an iDNS testbed in Asia Pacific countries: China, Japan, Korea, Hong Kong, Singapore, Taiwan, Thailand...

# IDN History cont'd

- 1998-1999
  - Prototypes demonstrated in international conferences
  - BoFs held in international conferences (APRICOT, INET)
  - Singapore, China, Hong Kong, Korea, Japan expressed interests in implementation
- November 1999
  - BoF in IETF
  - IETF Mailing list discussion
- January 2000
  - IETF IDN Working Group formed

# IDN History cont'd

- End 1999
  - Testbeds emerge and companies began commercialization
- July 2000
  - Multilingual Domain Names Consortium (MINC) and Country/regional organizations formed (e.g. AINC, CDNC, INFITT, JDNA)
- March 2001
  - IDN Working Group formed in ICANN Board
- September 2001
  - Creation of ICANN IDN Committee

# IDN History cont'd

- December 2001
  - ITU/WIPO/MINC Symposium ([www.itu.int/mdns](http://www.itu.int/mdns))
- December 2002
  - ITU Plenipotentiary Resolution 133
- March 2003
  - Technical solution in RFCs 3490, 3491, 3492 published
- June 2003
  - ICANN publishes guidelines
- 2003....
  - National and regional administration frameworks under development



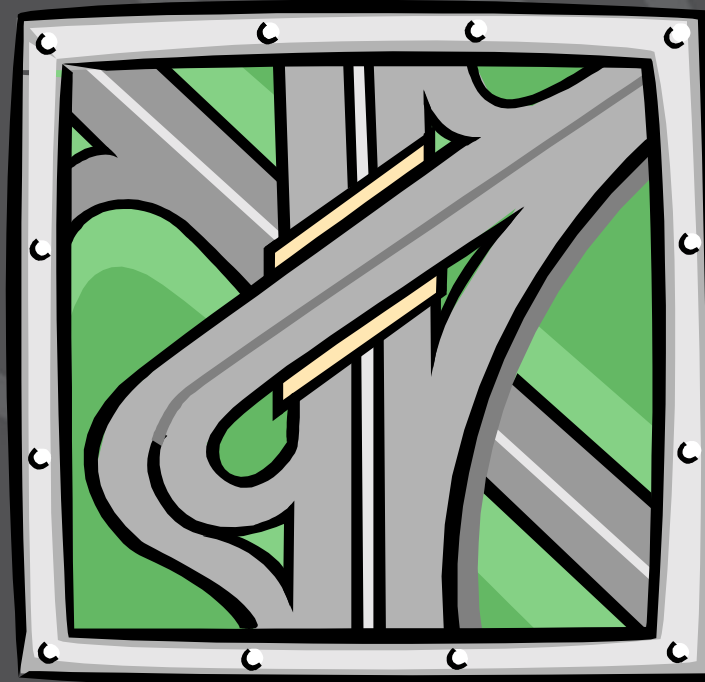
# The IETF Technical Solution

- Internationalizing Domain Names in Applications (IDNA)
- Based on code points in Unicode (ISO 10646)
- Case folding and normalization process to encode Unicode representation into ASCII Compatible Encoding (ACE)
- Keep ASCII “on the wire” for compatibility with existing software and DNS infrastructure
- Domain labels start with “xn--” represent ACE encoded “internationalized” label

# Relevant Technical RFCs

- 3490: Internationalizing Domain Names in Applications (IDNA)
  - <http://www.rfc-editor.org/rfc/rfc3490.txt>
- 3491: Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)
  - <http://www.rfc-editor.org/rfc/rfc3491.txt>
- 3492: Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)
  - <http://www.rfc-editor.org/rfc/rfc3492.txt>

# Some IDN administration & policy “works in progress”



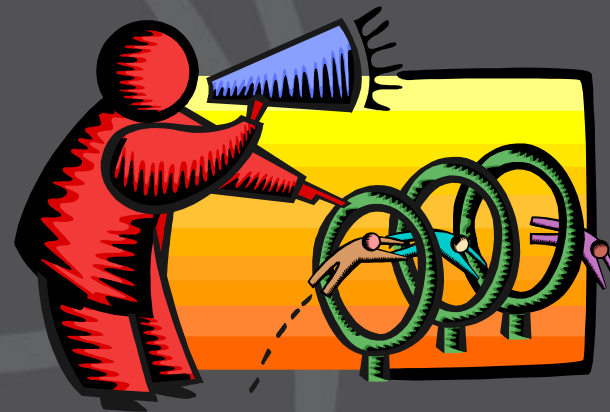
# Administration & Policy Directions

- Much work to do - will take many years
- Difficulties include:
  - identifying responsible “entity” to coordinate activities
  - dealing with complex administrative and policy arrangements, intellectual property, dispute resolution, sensitivities related to cultural and social issues
- MINC’s prior work in defining matrix of languages/scripts/country describes complexity
  - See “Who is the Language Authority for Multilingual Domain Names?” in ITU briefing paper at [www.itu.int/mdns/](http://www.itu.int/mdns/)



# Administration & Policy Directions

- No single “generic framework” possible but pieces of puzzle starting to emerge...
- Many hoops to jump through!



- Part of WSIS draft “action plan” (before 2010)

# Some problems

- Mitigating user confusion
  - Principle of least astonishment
  - Example: shouldn't megève.com and megeve.com lead to same end-user experience?
  - Many opportunities for mischief
- Some language scripts are much more complex
- Definition of valid UNICODE code points for language scripts
- What language scripts to support?

# Issue of Character Variants

- Simple example: should geneve.ch be equivalent to genève.ch if “e” is considered variant of “è”?
  - registration policies in Switzerland suggest it will be...
- Very complex examples in Chinese, Japanese and Korean (CJK) scripts : alignment between simplified and traditional Chinese, selection of “preferred variants”
- Leads to concept of multiple registration of “domain name packages” or “bundles”

# Snapshot of some ideas for frameworks

caveat: incomplete...



# A Method for Registering Internationalized Domain Names

- Paul Hoffman: draft model registration framework for internationalized domain names
  - <http://www.ietf.org/internet-drafts/draft-hoffman-idn-reg-01.txt>
- Provides simple generic model for administrative arrangements

## Hoffmann draft: suggested practices

- Before accepting registrations of domain names into a zone, valid codepoints in the Unicode character should be defined
- Decide whether particular characters in a registered domain name should cause registration of multiple equivalent domain names
  - these domain names can be added to zone or blocked from registration
- How to handle character variants in registering IDNs, and how to publish tables that list the character variants

# Guidelines for registration policies for Internationalized Domain Name Registration and Administration

- Guidelines for administration of domain names that contain characters drawn from Chinese, Japanese, and Korean (CJK) scripts.
  - <http://www.ietf.org/internet-drafts/draft-jseng-idn-admin-04.txt>
  - <http://www.ietf.org/internet-drafts/draft-jseng-idn-admin-04.pdf> (displays CJK characters)
- From Joint Engineering Team (JET), group composed of members of CNNIC, TWNIC, KRNIC, and JPNIC as well as other individual experts.
- Very complex to deal with complex CJK environment

# Set of Drafts by Edmung Chung, Neteka

- Charprep: <http://www.ietf.org/internet-drafts/draft-chung-idnop-charprep-00.txt>
- Zoneprep: <http://www.ietf.org/internet-drafts/draft-chung-idnop-zoneprep-00.txt>
- EPP: <http://www.ietf.org/internet-drafts/draft-chung-idnop-epp-idn-00.txt>
- IDN Registry Implementation & Operations Guidelines  
<http://www.ietf.org/internet-drafts/draft-chung-idnop-guide-00.txt>

# ICANN IDN Guidelines

- Must comply with RFCs 3490, 3491, and 3492
- Must identify permissible Unicode code points and block non-compliant registrations
- Must associate registration with one or more languages and employ language specific registration rules (e.g. reservation of domain names associated with character variants)
- Registries and registrars should provide informational resources and services in all languages for which they offer IDN registrations

# Some implementation experiences

# Korea

- Announcement from Korean Ministry of Information and Communication (MIC) and Korea Network Information Center (KRNIC) that registrations in the Hangeul script (with the .kr extension) would start on August 19, 2003:
  - [http://www.mic.go.kr/eng/pol/pol\\_inf\\_view.jsp?idx=105](http://www.mic.go.kr/eng/pol/pol_inf_view.jsp?idx=105)
- To minimize domain name disputes, registrations are implemented in a phased approach
  - <http://domain.nic.or.kr/menu/hanrequest1-3.html>

# France

- France
  - See Stephane Bortzmeyer, AFNIC: IDN Deployment in France (PDF)
  - [http://www.eurocio.org/domainname/documents/2003/presentations/presentations\\_at4/stephane\\_bortzmeyer\\_ppt\\_ang.pdf](http://www.eurocio.org/domainname/documents/2003/presentations/presentations_at4/stephane_bortzmeyer_ppt_ang.pdf)
  - Availability end 2003?



# Poland

- The Polish Research and Academic Computer Network (NASK) (administrator of .pl)
  - <http://www.nask.pl>
- Internet draft documents accepted Unicode codepoints for internationalized domain name (IDN) registrations under .pl
  - <http://www.ietf.org/internet-drafts/draft-bartosiewicz-idn-pltld-00.txt>
- Adds support for ą ć ę ł ń ś ó ź ż characters
- No variants/bundles
- Estimated available September 2003

# Interoperability Event

- IDNConnect: “virtual” interoperability event for testing implementations of internationalized domain names to be held September 23-27, 2003
  - <http://idnconnect.jdna.jp/>
- Sponsored by the Japanese Domain Names Association with Paul Hoffmann, IMC & VPNC, co-chair IETF IDN WG

# Some IDN Software Tools

- GNU IDN Library
  - <http://www.gnu.org/software/libidn/>
- International Components for Unicode (ICU) Libraries
  - <http://oss.software.ibm.com/icu/> with IDNA demo at <http://oss.software.ibm.com/cgi-bin/icu/idnademo>
- JPNIC Toolkit
  - <http://www.nic.ad.jp/ja/idn/idnkit/download/>
- Paul Hoffman's Perl Libraries
  - <http://www.imc.org/idna/>

# IDN Software Tools cont'd

- Verisign IDN Software Development Kit (C & Java)
  - [http://www.verisign-grs.com/idn/sdk\\_download.html](http://www.verisign-grs.com/idn/sdk_download.html)
- UTF Converter (no punycode)
  - <http://www.macchiato.com/unicode/convert.html>
- Netscape 7.1 first browser to natively support internationalized domain names (IDN), see implementation description:
  - <http://devedge.netscape.com/viewsource/2003/idn/>

# ITU Newslog Channel on IDN

- News related to IDN from ITU
  - <http://www.itu.int/osg/spu/newslog/categories/internationalizedDomainNames/>
  - RSS XML feed  
<http://www.itu.int/osg/spu/newslog/categories/internationalizedDomainNames/rss.xml>
- Generic news on Internet Names and Addresses from ITU
  - <http://www.itu.int/osg/spu/newslog/categories/internetNamesAndAddresses>
  - RSS XML feed  
<http://www.itu.int/osg/spu/newslog/categories/internetNamesAndAddresses/rss.xml>

# Some Reflections

- We're at very early stage of IDN implementation
- The “ICT Superpowers” are advancing well but developing countries will need much assistance
- Complexity means “one size fits all” policy approach will not work
- Current administrative drafts don't address some complex issues (e.g., character position sensitive variants in Greek and Hebrew)
- IDN complexity exposes many weakness in DNS administration models

# Some Reflections cont'd

- Unclear alignment of ccTLD and gTLD policies
- What about internationalized top level domains?
  - Can linguistic variant bundles be created for TLDs?
  - Confusion? For example, .ru in Cyrillic script is .py which could be confused with the ccTLD for Paraguay (.py)
- In some cases, governments need to act as facilitators, particularly when there is no clear “language authority” or other initiatives not seen as “legitimate”...

# Some Reflections cont'd

- Liked John Klensin's IETF drafts on 'longer-term solutions'
- By the time fully implemented, will DNS still be viable solution?
  - Don't we need a better model for a future world of billions of named objects?
- Is future path in "federated" (Latin for trust) naming structures?
  - Key to XML-based web services security (SAML)
  - NB: History teaches that technical hierarchical federations usually not successful (examples: PKI, "The Directory")
  - Lessons from E.164 naming/numbering plan that has no single technical root?



# Future ITU Activities

- IDN implementation experiences discussions in number of ITU forums (future IDN workshops (e.g., pan-Arab region and CIS activities, IP policy manuals)
- Bring together experts so that they can share experiences for the benefit of others
- Build knowledge base of materials and implementations available to ITU Member States
- Discuss role of national administrations of ITU Member States and possible policy role they may wish to consider
- Discuss further cooperative measures at both regional and international levels, particularly with regard to assisting developing countries in their consideration of these new technologies?
- Ideas?

Thank you

International  
Telecommunication  
Union

Helping the world communicate