Building Indian IDNs

IDN Panel Discussion

APTLD Meeting
Dubai
3 June 2007
Ram Mohan
.IN Registry
rmohan@afilias.in
.IN Brief History

Opened in 1992, run by R&D institutions through 2004

6,600 names registered through 2004

From 2005:

- Managed by National Internet Exchange of India (NIXI), Afilias registry Technical Partner
- 2nd & 3rd level domains available on an unlimited basis
- Separation of Registry & Registrar functions

- 250,000 registrations through May 2007
- No IDN registrations so far
India’s Multi-lingual society

- Total Population 1.002 billion (2001 Census)
- 2 Classical Languages – Tamil (2nd century BC) & Sanskrit (5th century BC)
- 22 Official Languages
  - Devanāgarī (देवनागरी) script based (North Indian):
    - Hindi, Marathi, Sanskrit, Kashmiri, Sindhi, Nepali, Manipuri
  - Dravidian script based (South Indian):
    - Tamil, Telugu, Kannada, Malayalam, Konkani
  - Arabic Script Based: Urdu
  - Other script basis: Bengali, Oriya, Gujarati, Punjabi, Assamese

- Worldwide Audience:
  - Hindi - 400 Million Speakers
  - Bengali - 200 Million Speakers
  - Tamil - 70 Million Speakers
  - Telugu - 80 Million Speakers
  - 7 more languages with audience of 25-50 million speakers

- Schools teach in 58 different languages
- Newspapers publish in 87 languages
- Radio programmes broadcast in 71 languages
- Movies released in 15 languages
The .IN IDN Plan

- Introduce Hindi & Tamil IDNs for .IN first
- Develop Language Policies
- Consider Issues Brought About by Variants
- Ensure Standards-Compliance
- Effect Single-Root plan
- Enhance Dispute Resolution Policy to cover IDNs
Building Indian IDNs

Language & Architecture Challenges
- About 12 Indian languages are based on Devanagari script …
- Many-to-many problem set (many languages using more than one script for equivalent representation)

Other Linguistic & Script Representation Challenges
- Bi-directional text
- Multiple diacentric positioning
- Word Breaking

Current Unicode mapping considered insufficient by Indian government & linguists
- Unicode Consortium (UTC) & Indian Govt form South Asian Scripts subcommittee to evaluate & identify potential solutions
Launch Process Questions

**Registry Considerations:**
- Should an IDN Sunrise be imposed?
- Should there be a Landrush process?
- Should there be new reserved names?
- Should existing ASCII registrants get first shot at equivalent IDN registrations?

**Registrar Considerations:**
- New/modified client software to connect to registry
- Conformance to standards

**Standards:**
- Use a standards-based process
Next Steps To A Multi-Lingual Internet

- **Internationalize** … design technology to accommodate languages
  - *IDNs are only the start*
  - Build standard language tables/repertoire
- **Localize** … define character sets that define a language, and create local language strings
  - Convert to local languages and scripts
- **Standardize** … provide a uniform, cohesive and standard representation and protocol
  - Ensure we have a global standard – that works in all applications
    - Browsers
    - Email
    - Mobile Devices
Conclusions

- Over $3mm committed funding & >20 staff members dedicated to Indic script IDN efforts
- Since Indian languages are spoken worldwide, our work will likely impact other nations and their people
- Likely to be one of the most complex IDN implementations ever undertaken worldwide
Questions?

Contact: Ram Mohan
rmohan@afilias.in
www.registry.in