Introduction to an IDN Revision

John Klensin
History

- IDNA adopted in 2003
- Significant controversy started 2+ years ago
  - Some unsolvable issues
    - Domain names don’t map well onto languages
    - Unicode is best we have, but not ideal for this
    - Characters look alike
  - Some others
- RFC 4690
Evolution

• IAB effort and RFC 4690
  – Extensive discussion
  – Extensive review

• Current effort
  – Design team starting over two years ago
  – Previous BOF
  – Open mailing list, thousands of messages, many revisions and adjustments

• How many have read docs ??
Some Key Issues I

- Some important characters and scripts excluded
  - To a user of a script, it is important
  - To a user of a language, the entire orthography is important
    - Most long-lasting arguments have been about one character at a time
    - DNS is about mnemonics, not writing novels
    - IETF has no real way to resolve those arguments
  - Globally-dangerous characters

- Some characters *not* excluded – back to LDH-extended?
Some Key Issues II

• Some scripts perceived as badly handled for this purpose
  – Very passionate arguments
  – Threats about alternate roots or worse
  – IETF has no good way to resolve these either, but we need to listen carefully and try

• Confusing terminology
Some Key Issues III

• Structure of IDNA2003 definition
  – Unicode version dependency
    • Applications don’t know, regardless of what standard says
    • Looking up code points that aren’t defined
  – Poor ability to understand
    • What is permitted
    • What happens to it.
  – Poor extensibility and ability to check
    • In some ways, any change at all is incompatible
    • List of label separators
Some Key Goals

• Get a standard that
  – Unicode version agnostic
  – Easier to understand
  – More predictable wrt what happens
  – More adaptable to local conditions – realistic interoperability
  – Deals with issues discussed earlier to extent possible
Current Structure

• Several documents
  – Organization is somewhat opportunistic
    • Reflects division of labor rather than fundamental, necessary, logic
  – Rationale and explanation
    • No real correspondence to IDNA2003 documents
    • People need to understand IDNA or they will make their own adjustments – some experience with that already
  – Protocol
  – Tables
  – Bidi
DNS Internationalization

• Not just IDNA
• Many areas of responsibility to make things work well
  – Standard Protocol
  – Registries and Registry Restrictions
  – Registrants – Increasing or decreasing confusion
  – Lookup implementors and Applications
  – Common sense by users and others
• Still mnemonics, not literature
• Can’t solve confusion, but can provide better tools
Issues and Rationale
Issues and Rationale

- draft-hoffman-idna200x-topics is an attempt at a summary
- Not a substitute
Address Primary Issues

• Unicode relationships
• Registration-Lookup
  – Mostly about explanation
  – Some substantive changes, but subtle
• Characters
  – Internationalize LDH, not “everything possible, somehow”
  – Make to/from ACE more symmetric
  – Cannot get elegance of ASCII model without server-side matching… still making compromises
New Terminology

• Much confusion in IDNA2003
• Standard about labels, not FQDNs…
  – Still some Right-to-left (bidi) questions
  – Hard issues with label separators
• U-label, A-label, LDH label
  – And all kinds of invalid stuff
The Front End

• What does mapping mean/do
• A minimal-variation on-the-wire form
• Compatibility and transition
Summary

• Trying to move ahead to make IDNs far more useful and predictable as mnemonics
• No big-jump revisions in future
• Still can’t write literature