

Looking at email Internationalization

John Klensin

IETF 58

Disclaimer

- Prior reading of both proposals is assumed
- Mailing list “discussion” has been mostly a waste of time, showing the worst of the IETF process.
- This presentation is a review of issues, not a replacement for proposal-reading

An Architectural Problem

- Some interesting questions – where do we want to end up and what is important?
 - Multilingual Internet with English as one language *or* English Internet with some other scripts grafted on ?
 - Fully internationalized email *or* a kludge forever ?
 - Smooth working in an internationalized environment *or* ability to transmit encoded non-ASCII information to non-updated environments.
- Familiar, high-quality localization versus global interoperability

Global Interoperability versus Good Localization

- Can we, and should we, keep addressing in ASCII on the theory that it is all about protocol elements?
- Pro
 - Cannot get rid of “@” (and maybe “,” and “:”)
 - Better for global interoperability
 - May need ASCII alternatives forever for global communication.
- Con
 - People should be able to use their own languages, especially in communications with other local/native speakers.
 - Most email communications are fairly local to country or language

Bottom line

- As we have seen with URLs, email addresses will be internationalized.
- The only real question is whether it will be done in a standard way or locally and incompatibly
- Users who hate transliteration will not see incomprehensible encodings as a step forward.

The Concept

- Important to
 - Concentrate on making internationalized environments work well.
 - Less important to be sure that non-i18n environments can send/transmit/receive i18n mail
- Use transport capability negotiation to be sure recipient systems can handle new options
- Do not alter fundamental interoperability or flexibility properties of email by making local-part non-opaque or restricting the flexibility to ASCII.

Transport Address

- UTF-8-string@UTF-8-domain
- All canonicalization/ normalization problems are UTC problems, not IETF problems
- Provision of ASCII fallback in parameter ?
- Requirement for 8BITMIME ?
- The probably unthinkable: “@” is an ASCII character.

The Bad (?) News – This Changes Everything

- Changes to Received
- Changes to header From/To/Cc/etc. Or new fields?
- What to do about threading, minor, or user-defined fields?

The Error Cases

- Bouncing actually is acceptable
 - These new addresses will be most useful within communities who share language/ script/ degree of upgrading (localization)
 - Best address for unknown person halfway around the world is likely to be ASCII for a *long* time (global interoperability)
- Fallback address (traditional) in envelope. ?

Rethinking Mail Reading/ Delivery

- POP3 and IMAP need upgrades
- Could use IMAA-like approach for local transition, delivery MTA to MSA or MSA to MUA: keep kludges local so they can be removed when no longer needed, rather than turning into permanent parts of mail backbone.

Why is this Plausible?

- Email's local parts are more important than domain names – domains can usually be hidden.
- Domain names are multiprotocol; we are only talking about email here.
- MUAs in many non-English countries internationalized already. There is large demand which will likely accelerate as IDNs come into use
- Hopes for really fast transitions of most end user application software to handle IDNA encodings have been dashed:
 - There are no quick fixes
 - Any approach will take some deployment time
 - Let's get it right

An Internationalized Internet,
Not an English Internet with
Tolerance for Some Other Scripts