RFC3743 and IDN TLD tests

YAO Jiankang
姚明康
CNNIC
What is RFC 3743?

- RFC3743 JET Guidelines for IDN Registration and Administration for Chinese, Japanese, and Korean
- Chinese Japanese and Korean characters have variants for the same conceptual character
- **It means:** One conceptual character can be identified with several different Code Points in character sets for computer use.
- The principle of RFC3743 is to assign a bundle of names with character variants to the same registrant.
CJK K character variants issues

- characters with different code points to have the same meaning – example: the term “International”

- Chinese
  - Simplified Chinese: 国际 (56FD 9645)
  - Traditional Chinese: 國際 (570B 969B)
  - Other variants combination: 囯際 (56EF 9645)

- Japanese
  - Modern Japanese: 国際
  - But there is: 國際

- Other Possible Combination:
  - 國際
For example

- Simplified-Chinese (SC) and Traditional-Chinese (TC) characters equivalence for example in English, ‘A’ = ‘a’; ‘B’ = ‘b’
  Similarly in Chinese, ‘国’ = ‘國’; ‘华’ = ‘華’

- Chinese character variants, some Chinese characters has many variants but they are same in meaning for example {国,國,国,圀}
Why RFC 3743?

- http://www.icann.org/topics/idn/rfc3743.html
- BANK.cn equals bank.cn?
  Upper case BANK equals lower case bank
- 銀行.cn equals 銀行.cn?
  Simplified Chinese characters 銀行 equals Traditional Chinese characters 銀行
CNNIC policy for registration of CDN

- Each registrant registering Chinese Domain Name will be assigned the three domain names: (1) One original domain name (2) One traditional domain name for free (3) One simplified domain name for free

For example:
- 清 -> {清，清，清}
- 华 -> {华，華，華}
- 学 -> {学，學，孝}

Input Original Form: 清華大学.cn
Free TC-only Form: 清華大學.cn
Free SC-only Form: 清华大学.cn
A registrant registering ‘清华大学.cn’ in CNNIC will have:

**CDN package with a bundle of names:**

{ 清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；…}

**CDN used in DNS resolution**

- Input Original Form: 清华大学.cn
- Free TC-only Form: 清华大学.cn
- Free SC-only Form: 清华大学.cn

**Other CDNs which will not be used in DNS resolution and be reserved only for this registrant**

{ 清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；清华大学.cn；…}

- It protects the interests of the registrants of CDN and eliminate the possible “phising” problems

Note: The text is presented in Chinese characters, which are translated into English for better understanding.
‘例子.测试’ = ‘例子.測試’？

- For Chinese, Simplified Chinese characters 银行 equals Traditional Chinese characters 銀行 and their meaning is same although their scripts are different.
- If Simplified Chinese 银行.cn and Traditional Chinese銀行.cn does not belong to the same registrant, there will be a possible ‘phising’ problem.
  银行.cn xn--jn2A72X.cn  192.168.1.1
  銀行.cn xn--jn2Ax2s.cn  192.168.254.254
- Luckily, we have RFC3743.
- ‘例子.测试’ equals ‘例子.測試’.
  测 -> {测 測}
  试 -> {试 試}
  测试 = 測試
a new ‘phising’ opportunity?

- ‘bank.测试’ = ‘bank.測試’?
- The principle of RFC 3743 is to protect the interests of the registrant.
- If ‘bank.测试’ ≠ ‘bank.測試’ and ‘bank.测试’ and ‘bank.測試’ belong to the different registrants, a new ‘phising’ opportunity?
A suggestion to IDN TLD test

- Simple Chinese 例子.测试 and traditional chinese例子.測試 for ‘example.test‘ in the IDN wiki should link to the same webpage on which there are 2 versions, Simple Chinese 例子.测试 version and traditional chinese 例子.測試 version

- If that happens, it told users that 例子.测试 equals 例子.測試 and it eliminates the possible ‘phising’ problem
Internationalized Email Address

- 中文.cn
- 姚健康@中国互联网络信息中心.cn
IETF EAI WG progress

- Overview and Framework for Internationalized Email (RFC4952 July 2007)
- SMTP extension for internationalized email address (Finish WG Last Call)
- Internationalized Email Headers (Finish WG Last Call)
- International Delivery and Disposition Notifications (Finish WG Last Call)
Q&A

YAO Jiankang
yaojk@cnnic.cn
姚明@互联网中心.cn