Internationalized Domain Names Operations (IDNOP)

Internationalized Domain Names Operations (IDNOP)

國際化域名維運機制

黃勝雄
Kenny Huang,
MIS, Chinese Institute of Technology, ROC.
Edmon Chung, Neteka, Toronto, Canada

Abstract

The introduction of internationalized domain names (IDN) is an important and critical feature of the Internet as a global medium and for domain names to continue to be relevant as a user-friendly naming system in the multinational context. While the IDN set of RFCs [3490-2] have outlined a protocol for the transportation and resolution of IDNs, the additional management for provisioning, zone preparations as well as transition was not prescribed.

The IDNOP describes a framework for the following outstanding issues on the management and operations of IDNs for a DNS Zone operator:
- Character Equivalency Preparations [Charprep]
- Zone Preparations for IDNs [Zoneprep]
- IDN Mapping for EPP [EPP-I]
- IDN Registry Implementation Guidelines [IDNREG]
- IDN Transitional Resolution [IDNX]

摘要

網際網路已成為全球性的通訊媒體，網域名稱(簡稱域名)也成為通訊上使用者所熟悉的命名識別機制，因而國際化域名(簡稱IDN)的服務就顯得格外重要。網際網路IDN標準之一RFC[3490-2]闡述IDN在通訊、域名解析時所使用的通訊協定，但在IDN 管理、部署、域名資料準備及相關過度方案並沒有提出相關標準。

國際化域名維運(簡稱IDNOP)建立一個機制，此機制將域名伺服器(DNS)管理及維運面的議題作整體性的整合。此整合機制涵蓋下列各項目
- 字元對映準備 [Charprep]
- 國際化域名資料檔準備 [Zoneprep]

- 198 -
1. Introduction

This document serves as an introduction to a set of papers surrounding the management and deployment of IDNs for zone operators, including domain registries, registrars and host managers.

Throughout the discussion of the IDN protocol at the IDN workgroup several operational issues such as the management of character equivalence issues inherent with the use of the Unicode repertoire of codepoints, as well as the transitional provisions for IDNs were dismissed as beyond the scope of the core protocol. The IDNOP set of papers will serve to recapture the concerns raised at the discussions as well as to describe a framework for the management and deployment of an orderly IDN zone operation.

1.1 Terminology

The key words "MUST", "SHALL", "REQUIRED", "SHOULD", "RECOMMENDED", and "MAY" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Character Equivalence Preparations (Charprep)

2.1 Introduction

Some discussions about Character equivalence preparations (Charprep) and Zone Preparations (Zoneprep) has been outlined in a previous Internet draft: draft-jseng-idn-admin-02.txt. While the IDN-Admin document provided a good overview of the problems as well as implications on certain policies, the IDNOP set of documents intend to describe a more technical framework for the actual management and implementation of the said policies. Also to provide a more generic structure for all types of IDNs and not only surrounding CJK (Chinese, Japanese and Korean) issues.

Charprep will discuss the need for character equivalence preparations to preserve the user friendliness of a domain name for a common user. Charprep will also provide a generic framework for the publishing of character equivalence
mapping tables for each language and how they would be used for Charprep mechanisms to create lists of reserved domains to be used in Zoneprep (which will then determine the framework for the publishing of the IDN and their Charprep-ed aliases into the DNS zonefiles).

Charprep do not intend to discuss or provide specific mapping tables, however will point to and suggest an archive for the mapping tables to be registered.

2.2 Nomenclature

As in the Unicode Standard [UNICODE], Unicode code points are denoted by "U+" followed by four to six hexadecimal digits.

The following terms will carry specific definitions within this document:

Zone Manager: A domain operator or service that manages sub-domain delegations. This would include domain registries such as TLD registries as well as domain operators of SLDs to issue third level domains.

Registration: Entry of a domain into the zone file of an authoritative name server. Resolution: Matching or lookup of domain names within the name server.

IDN: Internationalized Domain Names: domain names consisted of one or more characters out of the A-z 0-9 and "-" repertoire.

2.3 Importance of Charprep

The best way to illustrate the importance and need for Charprep is through the following simple example:

Suppose a person obtained a domain <alpha><beta>.example from the .example Zone Manager. The person now advertises his domain as <ALPHA><BETA>.example (Alpah & Beta in capital letters). A user seeing this perceives the domain as AB.example. The user now attempts to access the domain and fails.

It is true that the characters <ALPHA> and <A> are not technically equivalent, but because of their perceived equivalence, it will cause confusion to the user and therefore defeats the purpose of having a human-friendly domain name system.

More importantly, it could create a security issue whereby a domain name is maliciously registered to confuse the end user. For example, suppose the AB.example site is an e-Commerce site, a malevolent registrant may register the domain <ALPHA><BETA>.example set up a link to it on a competing site. The end user will not be able to realize that s/he is being brought to a different site
because the display will always look like: B.example?

Charprep will provide a framework for the publishing of Charprep mapping tables that can be used by Zone Managers to create a set of variants from the original submitted domain (Primary Domain) that may cause user confusion. Further management of this set of variants with regards to zone file entries is discussed in Zoneprep.

2.4 Equivalence Mapping versus Prohibition

A common misconception is that Equivalence Mapping prohibits the use of mapped characters. This is NOT true. For example, even if <ALPHA> is mapped to <A>, and vice versa, it does not prohibit a Zone Manager to offer a domain name that contains <ALPHA>, or <A>, or both. To resolve possible conflicts, the first come first serve rule as employed by most zone managers today may naturally come into place.

Another common misconception is that character equivalence mapping requires word or phrase semantic equivalence. This is also NOT true. Charprep does not give much regard to the end phrase or word, but focus only on the character itself. Therefore, even though a mapped character may be semantically different, it should still be mapped as equivalent if it is visually identical (as in the case for <ALPHA> versus <A>). Or in another case, even though a mapped character may be visually different, it should still be mapped as equivalent if it is contextually identical (as in the case for Traditional versus Simplified Chinese characters).

2.5 Character Equivalency Preparations

Throughout the IDN discussions, character equivalency issues were repeatedly brought up. While it is appropriately dismissed as a core protocol concern, the importance of Charprep has never been discounted. Especially from zone operators who have started to deploy IDNs as well as from a policy point of view such as at the discussions at ICANN.

Charprep is important because characters that may be perceptually equivalent, whether visually or contextually, may occupy different codepoints (as specified in Unicode), and therefore make them technically distinct and unique "characters", yet in real-life they are perceived and considered to be the same.

There are two main types of perceptual equivalence that Charprep deals with: 1.
Visual Equivalence; and, 2. Contextual Equivalence. The CJK preparations consists mainly of contextual with some visual equivalence considerations, while the LCGA preparations are focused mainly on visual equivalencies with some case mapping considerations (e.g. the Turkish character "ı").

2.5.1 Visual Equivalency

One type of character equivalence issue is of a visual nature. For example, the Greek capital letter <ALPHA> is visually identical to the English capital letter <A>, yet they occupy two different codepoints in the Unicode scheme. The implication is that <ALPHA>.example and <A>.example are technically two distinct domain names even though, when displayed appear identical: "A.example", and "A.example", look exactly the same.

2.5.2 Contextual Equivalency

Another character equivalence issue is of a contextual nature. For example, within the Chinese language, one particular character may have a number of different visual representations, yet they are conceptually equivalent. The most noticeable case is the Traditional Chinese versus the Simplified Chinese representation of a character (e.g. . [U+767C("fa"-prosper)] and . [U+53B1("fa"-prosper | hair)]). To complicate matters these relationships may not be one-to-one, because within different context, a character may take on a semantically different meaning, therefore creating additional variances from the root character (e.g. . [U+53B1("fa"-prosper | hair)] and [U+9AEE("fa"-hair)]).

Furthermore, the Japanese and Korean languages share a subset of the Chinese character repertoire. Two characters that may be considered perceptually equivalent in the context of the Chinese language, however, may be considered distinct and unique in Japanese Kanji (e.g.[U+570B("guo"-country<cn>)("goku"-a name<jp>)] and . [U+56FD("guo"-country<cn>)("koku"-country<jp>)]).

It is therefore very important to preserve the perceptual expectations of the end user for multilingual domain names, to maintain the what-you-see-is-what-you-get? user-friendly spirit of domain names in order to allow it to continue to be a useful and human-friendly means of direct navigation and resource addressing on the Internet.

2.6 Charprep Mapping Tables and Mapping Profiles
Charprep deals with perceptual equivalency of characters. Characters are units of visual or graphical representation of the written form of languages. Scripts best define the collection of a set of characters. Charprep will utilize the ISO15924: Codes for the representation of names for scripts, as the guide for identifying scripts and managing Charprep mapping tables and mapping profile. Note clearly that this document does not intend to provide a full discussion on all of the scripts. Multiple scripts may share one Charprep mapping profile.

Each Charprep Mapping Profile MUST consist of two tables & a Report:
1. Reason For Equivalence Mapping Table
2. Variant Set Mapping Table
3. Accompanying Report

The Reason Table provides a clear rationale to the mapping process for reference, while the Variant Table is intended for technical processes to generate the set of Charprep variants for a given character.

A Charprep mapping profile MUST also be accompanied by a written report to provide background information on the derivation of the mapping tables.

3. Zone Preparations for IDNs (Zoneprep)

Following from Charprep, Zoneprep will describe a framework for the selection and inclusion of the equivalent domain names as identified by Charprep into the master zone files of the DNS.

More specifically, because the IDN protocol standards have not taken into consideration character equivalency issues, Charprep and Zoneprep intend to provide guidelines to zone operators to deal with potential confusion from end-users by either the contextual or visual identicalness of two domain labels.

For any given IDN, the primary domain will be identified as the first intended form, which is received by the zone operator. Subsequently, the Charprep mechanism should be employed to create a list of Reserved Variants (RV) from the primary domain. The RV set will be removed from the pool available for further provisioning by other users of the zone. From this RV set, additional strings will be selected via the Zoneprep mechanism to be included into or excluded from the master zone files for the DNS (Zone Variants ?ZV). Zoneprep is intended to provide a framework only, which will allow zone operators to implement different policies.
Internationalized Domain Names Operations (IDNOP)

based on the common expectations of their anticipated user base and demographics.

4. IDN Mapping for EPP (EPP-I)

In order for zone operators (registries) to announce the list of RVs and for clients (registrars) to further identify the ZVs to be included for each IDN using the Extensible Provisioning Protocol (EPP), a set of IDN Mapping will be introduced. In essence, Charprep and Zoneprep will provide the requirements for the IDN Mapping extensions required for EPP.

The IDN Mapping extensions will address EPP commands and responses for domain creation, domain update and domain info. More specifically to respond with a list of RVs upon domain creation, to allow for the selection and retiring of ZVs during domain update as well as to convey the existing allocations during domain info commands.

5. IDN Registry Implementation Guidelines (IDNREG)

There are two main areas of concern for domain registries (zone operators): Registration of IDNs and Transitional Resolution of IDNs. Because the IDNA standards currently do not discuss either of these issues, it presents a challenge to domain registries as they incorporate Charprep and Zoneprep to their registration policies.

There will be multiple mapping tables for Charprep as well as multiple policy profiles for Zoneprep, IDNREG will attempt to provide some guidelines for zone operators (registries) in the rolling out of IDNs to their users (registrants).

In addition, IDNREG will discuss transitional resolution issues. More specifically, what the zone operator (registry) could reasonably expect from DNS resolution requests as IDNs are being offered.

6. IDN Transition & Implementation (IDNX)

There is a common misconception that the IDN protocol standards once published could be implemented as is and that all users (registries, registrants, Internet users) can expect multilingual domain names to be universally functional immediately. In fact, the publishing of the IDN protocol RFCs is but the beginning of a long transition towards the universal incorporation of the standards.

The IDNX document will provide insights into what zone operators should
Internationalized Domain Names Operations (IDNOP)

expect as they implement IDNs as well as what types of transitional measures would be appropriate to smooth the transition and make it transparent for end users. It is critical for the success of IDN deployment to not confuse or frustrate the end users.

7. IANA & Security Considerations

This particular document does not require any IANA considerations, however the related documents for Charprep and Zoneprep will recommend for IANA to maintain registries for character equivalence mapping tables as well as for language policies respectively.

This document does not talk about DNS security issues, and it is believed that the proposal does not introduce additional security problems not already existent and/or anticipated by adding multilingual characters to DNS and/or using ACE.

Moreover, the considerations for Charprep and Zoneprep should help to improve the security and authenticity for the usage of IDNs.

References


Internationalized Domain Names Operations (IDNOP)


[RFC3492]  A. Costello, Univ. of California, Berkeley, Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)? March 2003