Native Language Internet Address System (NLIAS)
draft-pjlee-nlias-00.txt

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on May, 2002.

Copyright Notice

Copyright (C) The Internet Society (2001). All Rights Reserved.

Abstract

This Draft is to introduce Native Language Internet Address Service that becomes popular as an alternative Internet Name Service. Especially, it describes the backgrounds, rationale and the specification of the Native Language Internet Address Service.

Generally in internet service when user types a word to connect to the specific website, a query typed in is so called as Keyword. However, keyword is a word used to describe the service, which is to type a word related to the information the user wants to get in the search engine. So the author, myself, would like to use Native Language Internet Address instead of Keyword.
1. Overview

As Internet service and the Internet infrastructure grow very fast, Internet Name Service that is a basis for the Internet service is also being developed rapidly. All the development is being carried out to give more convenience to the Internet users and these efforts are shown in many ways.

IP address and DNS (Domain Name System) are the general Internet addressing schemes from the start of the Internet era to nowadays. In case of DNS service, it is being extended to IDN for the convenience of end users. In addition to these traditional addressing schemes, a new approach, called Native Language Internet Address, is actively being discussed among in the Internet society.

This document surveys the development trend of Internet Addressing schemes and describes the rationale and the architecture of Native Language Internet Address Service as an alternative Internet Addressing schemes.

2. Introduction

2.1 Development of Internet Address

Internet Address, as of today, has been advanced to allow multilingual characters in the domain names using IDN. However, from the viewpoint of the end user convenience, IDN is not an ultimate destination of the Internet Address development, but it is just one of the experimental tries to go beyond new Internet Address. Among many Internet addresses, this document discusses the Native Language Internet Address Service that has been discussed in a few drafts.

The development stages of Internet Addresses are as follows:

1) IP Address
   (210.103.175.31)
2) Domain Name
   (netpia.com)
3) I18N Domain Name
   (한국어.netpia.com)
4) Full I18N Domain Name
   (한국어.한국어.한국어, 한국어.한국어)
5) Native Language Internet Address
   (한국어.netpia)
As the number of IP address which is a combination of numbers has increased, the server management with only host IP addresses and host names becomes more inconvenient. To resolve this problem, domain name has emerged. Domain name, however, also has problems of limited namespaces using LDH[1] only. As Internet has been spread to non-English speaking countries, the need for using their own languages as Internet Addresses has increased.

As a result, IDN has emerged but it neither provide the community with the full convenience, nor is fully serviced as well. Now, more convenient addresses, known as, Native Language Internet Address has emerged. From 1) to 4) of above, the technical advancement has been achieved through the need of community. Native Language Internet Address, 5), is, conceptually, a brand new Internet address requires legal support as well as the technical advancement and community’s need.

Native Language Internet Address is based on the assumption that it is better to recognize an Internet address without current Internet addressing hierarchies such as TLD and 2LD, and this is a more advanced Internet addressing schemes.

A legal background of Native Language Internet Address is as follows.

| +--------> Hangul character itself can express the ccTLD.    |
| | That is a character code corresponds to .kr.               |
| +--------> It identifies the characteristic of organizations  |
| | according to the traditional trademark principles.        |
| | Therefore, 2LD becomes unnecessary.                       |

The character sets or languages can be used as ccTLD or TLD by character set identification system. For example, Hangul character set itself becomes ccTLD. This means that the language itself can identify the country so .kr is not needed any more and can be omitted. Also, the traditional trademarks already imply the organizations. So the `.co`, which implies company or corporation, can be omitted.

For example, in "Å-À;ÅÁ»ÅÁ©.go.kr", the Native Language Internet Address "Å-À;ÅÁ»ÅÁ©" (City Council of Seoul), itself identifies the governmental organization. As another example, in "Å-À;ÅÁ’ÅÁÑÅÁ.edu", the Native Language Internet Address, since "Å-À;ÅÁ’ÅÁÑÅÁ" (The Seoul National University) itself implies the educational institution ,".edu" is not necessary.

That is, 2LD or TLD such as ".co", ".go", ".or" or ".edu", which identify the characteristic of organizations already according to the traditional trademark principles, so it can be omitted in the domain names.
In other words, ccTLD and TLD can be resolved by the character set identification system. By the traditional trademark principles, the trademark name itself identifies 2LD and the organization, for example, .co, .go, .or and etc. As technology advances, the system can identify the 2LD or TLD(ccTLD) from the name even if the user does not specifies it. It is a kind of more advanced system so that the users can use the internet more conveniently.

There are two more important advantages in this approach. First, from the user’s point, the availability of internet is one factor to consider. In the traditional domain system, the domain is the combination of English alphabets and numbers designed for universal use. However, this can be an obstacle to the Internet access for the non-English speaking people. But the Native Language Internet Address can identify the Internet Address by the very real name in the native language or notation, which make the Internet more available to the local, non-English speaking people. Second, the stability and the user friendliness of the system without using 2LD or TLD are another advantage of the system, which has been verified by the commercial service experience of the last 2 years.

In the traditional domain names, as the combination of English alphabets, in an abbreviated form, are used for the name, the organization identification is needed to reduce the conflict of the names. In Native Language Internet Address, a real trademark or a name is used in itself. The conflict can be minimized by using the real name, although the registration policy permits abbreviated name by warning the possible conflict. (legal issues)

Native Language Internet Address has emerged as a result of Internet Address development toward the convenience of the community and it made the Internet more accessible for the community by using the real names in its native language, which was not possible in the traditional Internet Addressing scheme. As mentioned above, Native Language Internet Address is an advanced method derived from the community needs and the technical and legal developments of traditional internet addresses.

2.2 Overview of Native Language Internet Address Service

Native Language Internet Address Service connects the traditional domain name to the unique information such as organizational name, trademark, service name, person’s name, telephone number, HAM or pager number, barcode and so on. Native Language Internet Address should be serviced in a regional legal boundary. Also, Native Language Internet Address is provided by user’s locale information to map that information with the address of the cyber world.
2.3 Characteristics

Since the characteristics of Native Language Internet Address can express all the unique aspects of a given name, it should include all unique identifiers that a user can understand. Because Native Language Internet Address Service respects the registered names and can extend the implied name space easily, the name space shortage problem can be relatively alleviated.

Although it fails to satisfy all the needs of Internet Address as any other method, it enhances the accessibility and convenience of the end user, especially in non-English speaking regions.

3. Current Native Language Internet Address Service

Until now, four different attempts have been tried to provide the Native Language Internet Address Service.

The following requirements should be satisfied for the future of Internet and its community.

1) user friendliness
2) consistency
3) extensibility
4) stability
5) recognizability
6) universal validity

Service methods are as follows:

1) by application
2) by OS support
3) by network device
4) by N/S extension

One of the service methods is to provide Native Language Internet Address Service by every application. This method is simple and easy method to do service, but each implementation may be responded with different answers. This causes a lot of confusions to the community who uses Native Language Internet Address as Internet Address, that is, it lacks the uniformity. This makes the Internet as a closed private service like most of the local communication service provides, which is not the goal of Internet service.

Another try is to enhance the OS resolver or to use special network devices to provide the Native Language Internet Address Service. But these methods are still in the experimental stage and lots of time and efforts are needed to apply these methods to the community. For now, these lack the extensibility and the universal validity.
Yet another method that uses the NS query is being tried in many ways. This method is based on the fact that, in many applications, the Native Language Internet Address typed in (though sometimes this is not the case) by the end user is transferred to the NS. In fact, there are various attempts to extend the name server. These alternative NS or extended NS can actively respond to the queries from various applications.

This method has advantages that it can provide the same response from different applications, which means that this can provide the uniqueness of the Internet Address.

Even though DNS was not made to provide the Native Language Internet Address Service, it gives a hint on what should be done to provide the Native Language Internet Address Service. Naming Service should provide an unique service for the various applications. Native Language Internet Address Service falls on these category. That means that it can provide consistency.

Future Native Language Internet Address Service must support not only specific application program, but also the general naming scheme to be usable as an Internet Address. It should be compatible with the existing service and easily extensible to the future Internet applications. And it also shouldn’t affect the existing services.

4. Native Language Internet Address Service

4.1 Overview

There exist many identification methods for our daily lives a personal home page, e-mail, ICQ number, Telephone number, fax number and snail mail address are examples to name a few. These identifiers are used in a real life without serious problems and it is being extended to the Internet service. In fact, I send fax and telephone calls using the desktop PC. Also, there is some service providers which allows users to have an effect of sending a snail mails just by sending an e-mail in Korea.

Native Language Internet Address will provide a framework to interconnect these identifiers easily. For example, I have to search the address book many times to send some faxes to my friends. If there is a fax program that can send just by typing the name of receiver, the service would be much convenient.

This problem is not limited to the fax number. We have many problems to memorize many identifiers and sometimes we even fail to find the specific identifier we need. It would be more convenient to use Native Language Internet Address Service not only for the fax program but also for many other application programs.
4.2 NLIAS (Native Language Internet Address System)

As mentioned above, Native Language Internet Address should evolve as a form of Naming System which can resolve the queries generated from various applications. Differentiating this service from DNS, we call it as NLIAS (Native Language Internet Address System).

| +--- Name Server +--- |
| |     +---+     |   |
| | |     DNS     |   |
| | +---+     |   |
| | |     NLIAS  |     |
| | +---+     +---+ |

4.3 Client Server Model

The Native Language Internet Address System should be a C/S model like Domain Name System. The server should respond to the queries without difficulty from various users. For this, a C/S model is adequate because it is simple and it reduces overhead. Especially, the new protocol should not increase the network loads.

5. Requirements

For the technical requirements described below, we define a "service" for those related to something provided to the end users and we define "protocol" for those related to implementation.

5.1 Requirements for Compatibility and Interoperability

[1] Service should be a separate system from DNS. In these days DNS is so important that it can be referred as Internet itself. It should be a separated Naming System from DNS. After the verification of stability, the attempts to integrate into DNS should be pursued.

[2] Native Language Internet Address Service can provide the basic IP Address in no time. Otherwise, the DNS should be queried.

[3] The response from the same Native Language Internet Address should be same regardless of the server type whether it is a cache server or a root server.


[6] Unique Native Language Internet Address for each language (character set) should be serviced.
5.2 Requirement for the Internationalization

[1] The character set used in the service should be extended from the Unicode.
[2] Protocol should do the Name Preparation for the Native Language Internet Address before the service.
[3] Conflicts on Native Language Internet Address should be reduced by Name Preparation.
[4] For the case mapping, the upper case letters should be converted to the lower case ones since most user use lower case letters.
[5] Service should be based on the user’s language.

5.3 Requirement for the service administration

[1] Service should be restricted to 1:1 match.
[2] Native Language Internet Address Table should be easily modifiable
[3] Native Language Internet Address System should not give any influence on the traffic of the DNS system.
[4] Native Language Internet Address should be maintained according to the character set. No two different character sets can share the same table. The character sets means the extension of Unicode.
[5] Native Language Internet Address has n:1 correspondence with the Internet Resource. Internet Resource means the information table for the Native Language Internet Address.
[6] In a given Native Language Internet Address table, the Native Language Internet Address should be unique. The same Native Language Internet Address can be in another Native Language Internet Address table even if the Native Language Internet Address implies the same meaning.
[7] The categorization of the table is defined by that of Unicode.
[8] A Native Language Internet Address in a given table is case-insensitive. For example, "abc" and "ABC" cannot exist in the same table. If both exist in the same table, one of them is ignored.

6. Structural Overview

6.1 Interoperable view

\[
\begin{array}{ccc}
+--------+ & +---------+ & +--------+
| Individual User & Query & Query & Query \\
| ---------------+---------| NLIAS   | NLIAS   \\
| Resource at ISP & Servers <---| Root Server \\
+--------+ & +---------+ & +--------+
\end{array}
\]
When the user types in the multi-lingual Native Language Internet Address to the application, the Native Language Internet Address will be converted to the Unicode character string and then transmitted to the Native Language Internet Address Server, say A. Native Language Internet Address server A forwards the query to one of many root Native Language Internet Address Servers, say B.

The corresponding root server B will respond with the resource by identifying the character set string based on the unicode. After this, Native Language Internet Address Server A caches the result so that A can respond for the the subsequent query for this Native Language Internet Address. The query should include at least the following information.

[1] Native Language Internet Address
[2] Resource Information (application information)
[3] Language Information

6.2 Client Side

+-------------+ Local Code +------------+ Unicode +-------------+
| User        | Local String | Resolver | Unicode String | NLIAS Server |
| Application | Resource     | Server   | Native Language Internet Address Table |

The end user application should be changed to accommodate the internationalized native Native Language Internet Address service. In other words, the resolver should take into account the multilingual characters. This includes the locale information of the client and the queried protocol information (ex: http, ftp, irc) as well as the encoding of the Native Language Internet Address itself. The definitions on character encoding schemes are defined in Unicode Technical Report 17.

6.3 Server Side

+------------+ Packets +-------------+ Packets +-------------+
| Cache      | ROOT     | NLIAS Server | Native Language Internet Address Table |
| NLIAS Server | Resource | Server | Table |
+-------------+ Resource +-------------+
Note) Packets should include not only the unicode Native Language Internet Address but also the locale and the type of the protocol.

The packet transmitted by the resolver includes the information about the language used by a user or an application. Therefore, it transmits the query string to the authoritative server for the corresponding locale. The authoritative server (Root server) searches a Native Language Internet Address table for the matching string and returns the resource information if exists.

Native Language Internet Address Server follows these steps for the Native Language Internet Address.

[1] It searches the currently maintained caches; If a matched resource is found, it returns the resource information. Otherwise, it asks for the root server query.

[2] Root Server returns the resource information for the queried Native Language Internet Address. If no matching resource is found, "NOT FOUND" will be returned instead.

6.4 Cache Server and Root Server

Cache Server refers a NLIAS server with no authority on the Native Language Internet Address zone file and does not refer to a separate system. This server caches the response information for some finite time and responds directly to the same Native Language Internet Address query. But it turns off the flag which indicates the response comes from the authoritative server. If the queried Native Language Internet Address is not cached, it queries to the root server.

6.5 Usage in application programs.

There are many Internet applications. Each of these applications require somewhat different return values for the Native Language Internet Address. For the web browser, the expected result is an URL. For the mail client, an e-mail address should be returned for the Native Language Internet Address queries. This means that each application may require different identifiers for the same Native Language Internet Address. Soon there may be some telephony system which dials up automatically just by saying the "Native Language Internet Address". NLIAS should be designed to accommodate these applications.

In order to satisfy these requirements, Native Language Internet Address System should be mapped to an information group and the information groups should be designed to be easily extended for the future applications as well as the existing applications.
6.6 Consideration on Internationalization

Native Language Internet Address System is not limited to be serviced on English speaking regions. Various languages should be supported for the international service. Until now, Unicode is used as an encoding scheme to support the international service so far, but the Native Language Internet Address System should support the local regional codes so that it can be more extensible. Also Native Language Internet Address Service should be based on the local service. It can be supported best in the corresponding local region by the traditional trademark principles. Native Language Internet Address System should consider the legal aspect since the legal issues as well as the technical issues must be developed.

7. Conclusion

As mentioned above, Native Language Internet Address Service should be another connection service that makes it easier to type and memorize the various resources without any modification nor change on the existing service.

8. Author’s Address

Jinhyun Bae
Netpia.com, Inc.
35-1, Youngdeungpo-Dong 8-ga, Youngdeungpo-Gu Seoul, Korea 150-038
Tel : +82-2-2165-3060
Fax : +82-2-668-4913
E-mail: piano@netpia.com

Panjeong Lee
Netpia.com, Inc.
35-1, Youngdeungpo-Dong 8-ga, Youngdeungpo-Gu Seoul, Korea 150-038
Tel : +82-2-2165-3044
Fax : +82-2-668-4913
E-mail: pjlee@netpia.com
9. definition

[1] LDH : Letters, Digits, and Hyphen

10. Reference

March 1982, K. Harrenstien

November 1987, P. Mockapetris.

November 1987, P. Mockapetris.

[DIRECTORY] "Definitions for talking about directories".
draft-alvestrand-directory-defs-02.txt.

[DNSROLE] "Role of the Domain Name System".
draft-klensin-dns-role-01.txt.
May 2001, J. Klensin.

[DNSSEARCH] "A Search-based access model for the DNS".
draft-klensin-dns-search-01.txt.

[ARROUYE] "Keyword Lookup Systems As a Class of Naming Systems"
draft-arrouye-kls-00.txt.
August 1, 2001, Y. Arrouye and V. Parikh and N. Popp

[SLS] "Service Lookup System".
draft-mealling-sls-00.txt.

June 2001, N. Popp, M. Mealling, and M. Moseley.

[UNICODE] The Unicode Consortium, "The Unicode Standard". Described at
http://www.unicode.org/unicode/standard/versions/.

[UTR15] "Unicode Normalization Forms", Unicode Standard Annex #15,
http://www.unicode.org/unicode/reports/tr15/, 2000-08-31,
M. Davis and M. Duerst, Unicode Consortium.

[UTR21] "Case Mappings", Unicode Technical Report #21,
http://www.unicode.org/unicode/reports/tr21/, 2000-09-12,
M. Davis, Unicode Consortium.