L2/00-283
IDN Working Group
Internet Draft
L2/00-283

Edmon Chung & David Leung
Neteka Inc.

<draft-ietf-idn-dnsii-mdnr-00.txt>

August 2000

DNSII Multilingual Domain Name Resolution

STATUS OF THIS MEMO

Abstract

The Internet-Draft for DNSII-MDNP was focused purely on discussing the ultimate packet protocol that is being sent around the Internet for multilingual domain names. This paper complements the previous paper by outlining the contemplated resolution process with the DNSII protocol throughout the DNS name resolution process.

The DNSII-MDNR outlines a resolution process that forms a framework for the resolution of multilingual domain names. Whether the DNSII protocol is implemented exactly as specified in DNSII-MDNP is less relevant, rather it is the idea of having a multilingual packet identifier and the fall back process that matters. The DNSII-MDNR successfully addresses the transitional issues at each node of the DNS resolution process providing a clear migration path and strategy for the deployment of a multilingual enabled DNS system. It also outlines the conformance levels required for basic or complete implementations for applications, resolvers and name servers.

This document also introduces a tunneling mechanism for the short-run to transition the system through to a truly multilingual capable name space.

Table of Contents

1. Introduction

This Internet-Draft describes details of the contemplated resolution process after the deployment of DNSII-MDNP, or other multilingual domain name efforts containing a bit flag multilingual packet identifier or otherwise InPacket identifications for that matter.

The reader is assumed to be familiar with the concepts discussed in the DNSII-MDNP Internet-Draft <draft-ietf-idn-dnsii-mdnp.txt>.

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].

A number of multilingual characters are used in this document for examples. Please select your view encoding type to Big-5 (Traditional Chinese) for them to be displayed properly.

1.2 Multilingual Domain Name Resolution

The original specifications for the DNS were designed to be open enough for simple implementation of a multilingual naming system with slight adjustments as laid out in DNSII-MDNP. The transition and especially its resolution process during migration is however a tricky problem. Several things that MUST be kept in mind is that there is a initial phase, an intermediate phase and an ultimate steady state phase. DNSII-MDNP only described the ideal protocol at steady state, with incorporated flexibility for transition from the present to multilingual as well as again towards future unknown grounds.

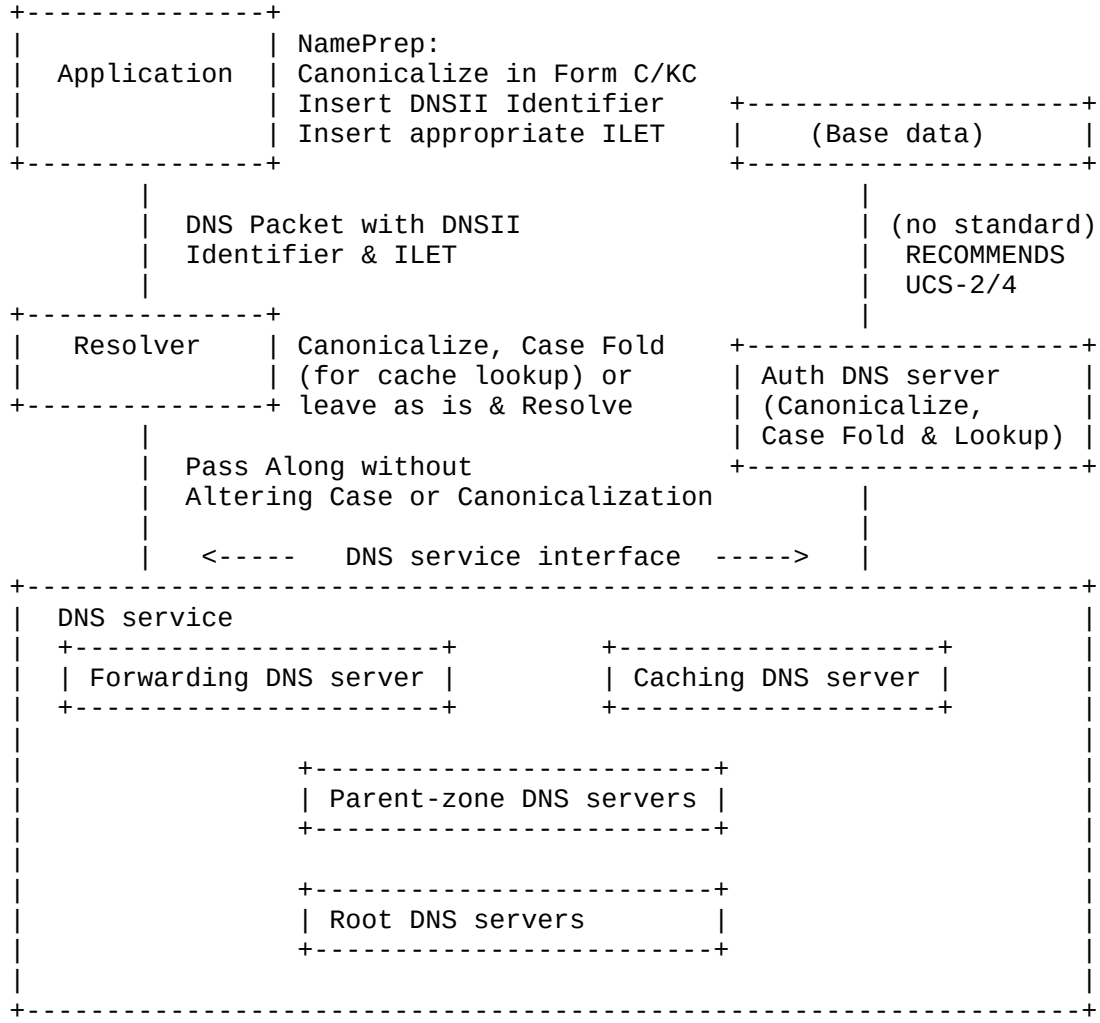
It is important to remember that the ultimate form SHOULD be determined and then the transition scheme laid out. While an ASCII translation system might seem favorable in the short-run, it effectively creates an alternative universe which is counter to the spirit of the DNS. However an ASCII translation is implemented, it immediately creates a "human-multilingual" universe and a "machine-ASCII" universe. This document introduces a tunneling mechanism to transition the DNS from today's monolingual system, through an 8-bit or 7-bit migration scheme towards a truly sustainable multilingual name space, arriving at a DNSII type system.

1.2 DNSII-MDNR

While DNSII-MDNP describes the framework for the ultimate protocol format of a multilingual DNS, DNSII-MDNR will discuss how the packet SHOULD be transported and interpreted throughout the DNS. The document will describe both the intended resolution process as well as part of the transition strategy from the existing DNS to a DNSII enabled system.

2. DNSII Proposal with respect to the DNS Layers

The following diagram illustrates the use of DNSII-MDNP at a steady state. Section 3 will discuss the fallback strategies while Section 4 will talk about issues on conformance levels.



Please note that at each level, the domain name is being canonicalized. This is to ensure that at the end, characters that can be represented by a single code point will not be otherwise compared resulting in inconsistent reply to a humanly identical name. It is RECOMMENDED that applications SHOULD conduct canonicalization while servers MUST. Duplicating the process is fine because if a character is already composed, it will not compose again with another character.

This arrangement is very similar to the ASCII case folding experienced in the DNS today. In the original specifications, it was RECOMMENDED that query sent be left as they are and case folding done only at the server end. Some application implementations however do perform the case folding at the user end. As the query arrives at the server, it is still case folded.

Case folding for multilingual domain names should follow the existing implementations for ASCII names, where the application SHOULD and the server MUST.

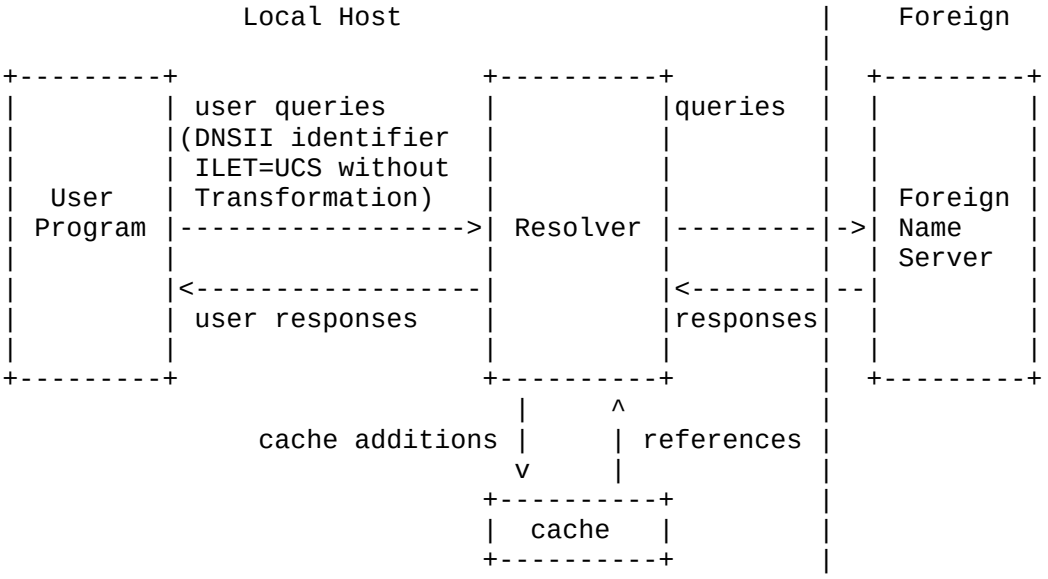
3. The Resolution Process

It is inevitable that at the end of the day, the entire DNS chain SHOULD be updated in order for multilingual domain names to be as efficiently resolved as names under the current DNS. DNSII strives to provide a schema that ultimately brings the system to a desirable steady state while carefully giving considerations to all the transition issues. These include the considerations that at the application end, there is already a preference and an installed base of character encoding that may or may not conform to the desires of the server end operations. The use of ILET is therefore highly desirable and essential.

3.1 Steady State Resolution

At steady state, the resolution process of multilingual domain names SHOULD be identical to the existing system. Additional steps of going through alphanumeric translation are unnecessary and undesirable.

With DNSII, the contemplated steady state process resembles the well-known DNS model laid out in RFC1035.



Eventually, an ISO 10646 UCS without transformation is desired as the common format. The benefits of having a uniform byte length encoding

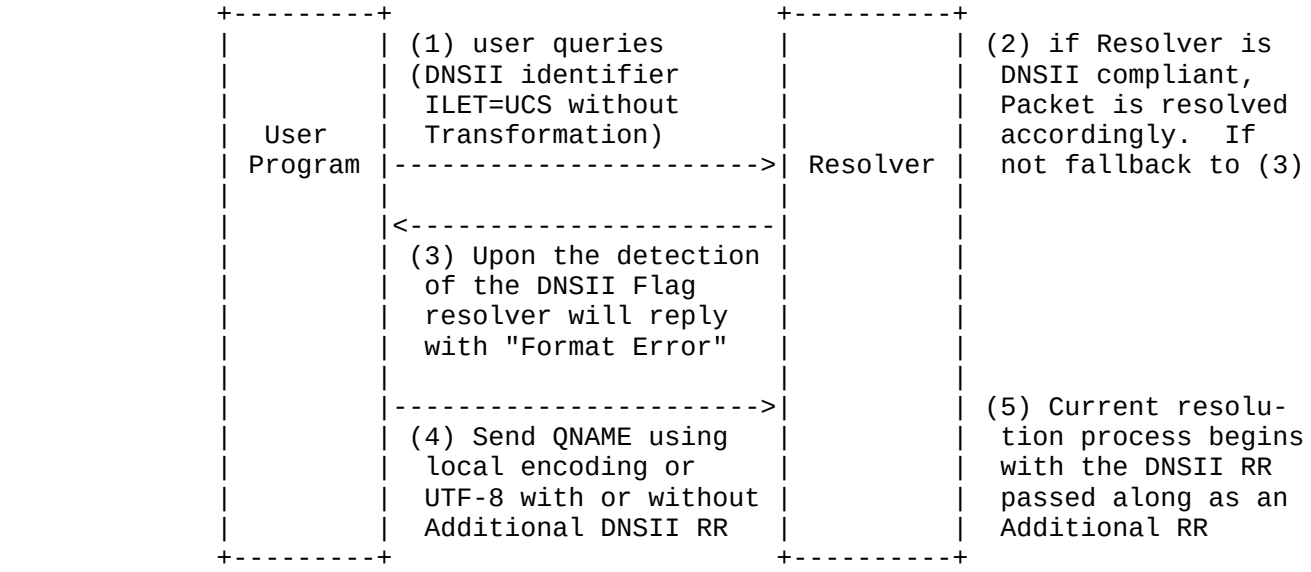
far exceeds the seemingly easier transformation solution. Especially

considering that the DNS requires a label count that should reflect the number of characters in a label. Of course there exists combination characters in the ISO 10646 specifications as well, but after canonicalization, only the ones that must use combinations remain and they are usually meaningful depictions.

The importance of this count value is further demonstrated by scrambling efforts to extend the size of this field or to compress character encoding to accommodate multilingual characters. With DNSII, this no longer constitutes an issue because any language will be able to share the same number of characters thanks to the use of ISO 10646. As a matter of fact, the desire to use uniform byte length characters formed part of the original intent of the ISO 10646 initiative anyway.

3.2 Client-End or Inquirer Transitional Fall-Back Strategy

For a DNSII aware Client, it will be able to create DNSII packets which provides precise character data of the domain name in question. However, if it encounters a non-compliant resolver, it MUST be able to fallback to a format acceptable by current resolvers.



3.2.1 Tunneling MDNP through DNSII RR

An additional DNSII RR would be tunneled through using the format of a TXT RR with the RDATA part containing the multilingual labels in a DNSII compliant format. The TTL of the RR MUST be set to zero to avoid caching.

It is not feasible to have a new RR type just for DNSII because the resolver might reject RRs with unknown types. For the name used in

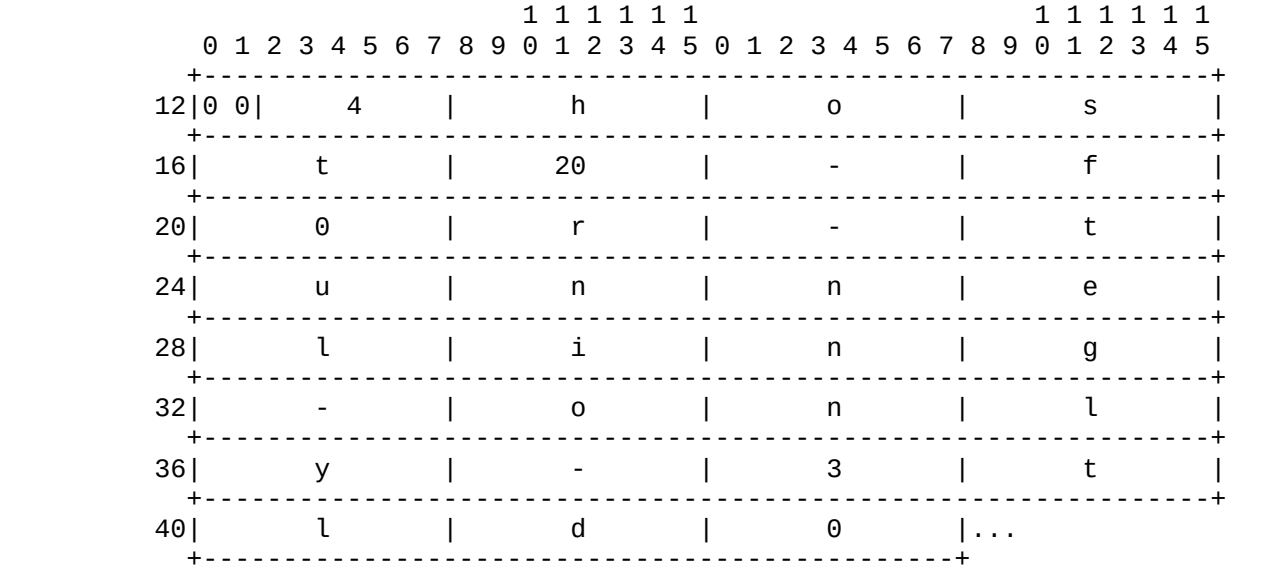
the QNAME as well as the NAME field of the DNSII RR UTF-8 MAY be used as the default fallback encoding. However, an arbitrary ASCII name MAY also be used just for the purpose of tunneling. The TTL for responses to tunneled requests MUST be set to zero to avoid caching at any level in the DNS. More detailed description in Section 3.4.

For older DNS servers, requests with a non-empty additional information section MAY produce error returns, however since the deployment of DNSSEC, especially for TSIG considerations, this no-longer constitutes a problem. Basic security prepared servers such as BIND 4 or 8 is already capable of bearing the tunneled DNSII RR.

It is possible to use ACE/RACE type translations at this level, however it is more advisable to use a truly arbitrary label such as "-for-tunneling-only-". So doing requires only reserving one arbitrary name while ACE/RACE creates one more arbitrary name for each new multilingual name registered, which will eventually contribute to the fracturing of the DNS.

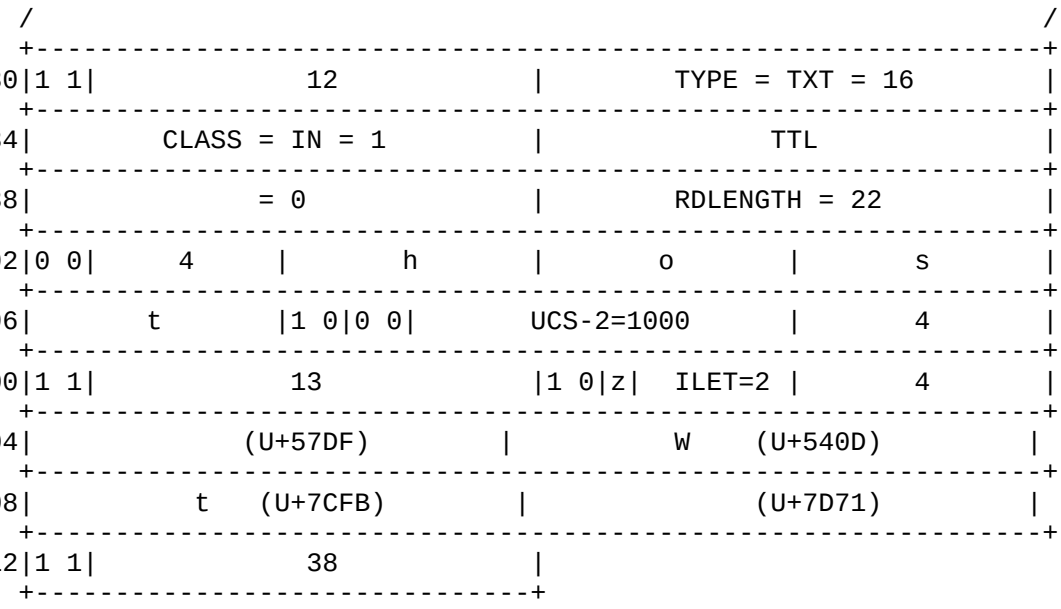
As an example, a tunneling packet for the domain name: host. Wt .tld. (4host4??Wt3 tld0) will be represented as:

(in the QNAME field)



(The Additional DNSII RR Tunneled in TXT RR form)

: :



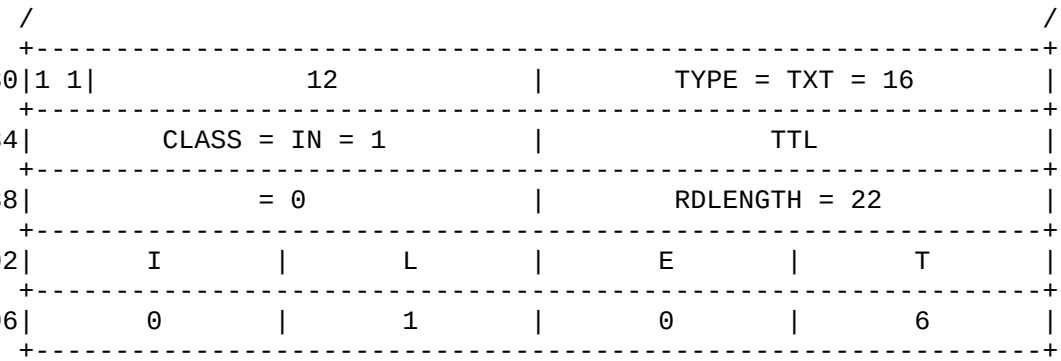
The reason a DNSII RR is attached is to alert the authoritative DNS server that the query is DNSII compliant while being able to tunnel the request through non-compliant resolvers without any loss of information.

3.2.2 Tunneling ILET RRs

Another fallback strategy is to tunnel just the ILET instead of the entire DNSII label. If UTF-8 or a local encoding is used as the QNAME, then the arbitrary ASCII label is no longer necessary. The tunneled RR therefore need only consist of an ILET indicating the encoding format used.

Within the RDATA of an ILET RR masked as a TXT RR the first 4 bytes will be used to indicate that it is an ILET and the following 4 bytes to reflect the MIBenum of the encoding used.

Following the example given in 3.2.1, the QNAME would be in UTF-8 (MIBenum = 106), while the additional ILET RR would be in the form:



3.3 Resolvers & Server-End Transitional Fallback Strategy

The tunneling scheme described in Section 3.2 assumes that the authoritative server is fully DNSII compliant. This assertion is laid out in Section 4.3 where it is discussed that only fully compliant servers SHOULD serve multilingual names directly under their authoritative zone. In any other case, the arbitrary domain "-for-tunneling-only-" should result in an NXDomain response.

For security aware servers, an NXT RR of the last name wrapped by its first name in the zone records will be returned because of the leading "-" for the tunneling label.

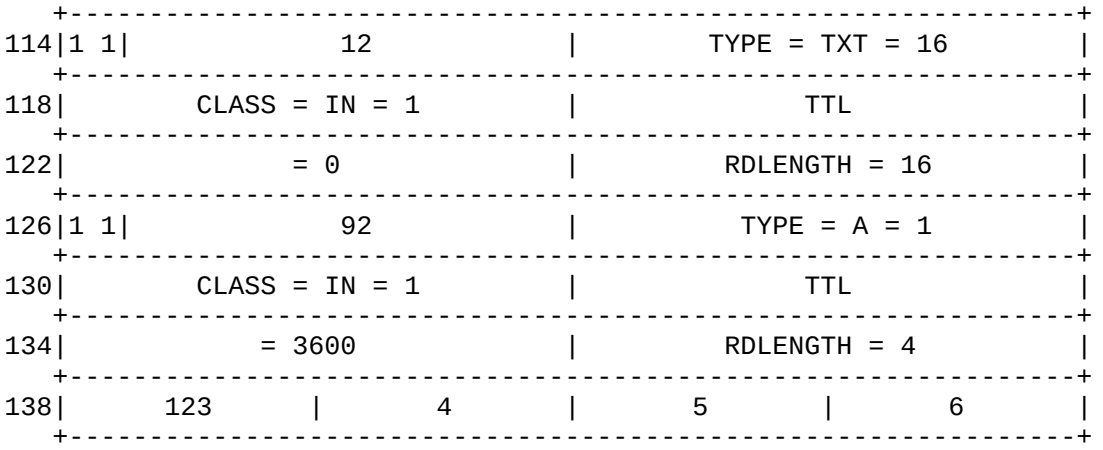
If the application end is not DNSII compliant, the fallback resolution strategy for resolvers would simply be to pass along the labels in their 8-bit format and determine the existence of the requested name as usual. If a tunneled DNSII RR is detected, by way of a label constituting entirely of "-for-tunneling-only-" and the existence of a valid DNSII RR, the resolver should attempt to resolve the name according to the DNSII specification and tunnel the answer back to the inquirer.

3.3.1 Tunneling MDNP Responses through DNSII ANS RR

To tunnel a DNSII compliant answer through a non-compliant resolver, another DNSII ANS RR is tunneled. Also using the TXT RR format as a mask. TXT RRs are best used because it is a valid RR and its RDATA is an unrestricted byte stream determined only by the RDLENGTH. The RDATA for a DNSII ANS RR would be the entire content of a regular response RR attached to a DNSII format name.

Continuing on the example given in Section 3.2, suppose an A record is requested and the IP address returned for the domain host.Wt

.tld. is 123.4.5.6, then an additional DNSII ANS RR (TXT) in the following form will be included:



Note that compression is available in the DNSII RRs. While the tunneling TXT mask uses the ASCII tunneling name and therefore points back to the QNAME at offset 12, the tunneled A Record response uses the offset corresponding to the DNSII compliant labels at 92. While the TTL of the TXT mask MUST be zero, the tunneled A Record RR contains a regular TTL, in this case 3600.

3.3.2 Reinsertion of ILET and DNSII Identifier

When a resolver receives an incoming query with a tunneled DNSII/ILET RR, it SHOULD reconfigure the query into a fully compliant format before engaging in further resolution. If a "00" query is received, the resolver should convert the label into UTF-8, set the DNSII identifier "10" on and set the ILET to UTF-8.

In the scenario where the client end is not DNSII compliant, either a local encoding 8-bit stream or a UTF-8 encoded stream without the DNSII flag nor ILET will be transported. During the transition period, should this occur, the above forced UTF-8 conversion and ILET insertion would take place and it would be up to the authoritative server to determine the existence of the requested domain. InZone DNSII handling mechanism will serve to take care of these transitional exceptions.

4. DNSII Conformance Levels

DNSII is designed for a smooth transition from the existing ASCII based DNS to a multilingual capable DNS. Therefore, it is not necessary for all servers and applications to be switched to multilingual capable before starting the deployment.

4.1 Application Conformance Levels

The BASIC compliancy for applications would be to remove validity checks for domain names. The resolution process will determine a non-existing domain name, so there really is no need to prevent a DNS packet with multilingual labels to be sent through the wires.

The INTERMEDIATE compliancy for applications involves the insertion of the DNSII identifier as well as the ILET according to the local inputting and screen scheme. If a user is using a JIS format scheme, it should set the ILET to reflect it being used. At the same time, the tunneling mechanism discussed in Section 3.2 MUST also be in place.

FULLY compliant applications will send all DNS packets with the DNSII identifier and the ILET set to UCS-2/4. The fallback scheme discussed in Section 3.2 MUST also be in place. InZone DNSII mechanisms SHOULD also be available to deal with local encoding exceptions.

4.2 Resolver Conformance Levels

The BASIC compliancy for resolvers would be to allow an 8-bit clean approach to name resolution. Also, it should be made sure that the additional DNSII RR (TXT) will not be truncated during resolution.

The INTERMEDIATE compliant resolvers MUST understand how to process the DNSII identifier as well as not reject the ILET. Interpretation of the name is not required so it is NOT necessary for the resolvers to be able to map all or any of the ILET values (with the alternative approach in DNSII-MDNP, the ILET value corresponds to the byte length of the characters contained in the label, which makes the count workable even if the ILET value is not known by the resolver). In this scenario caching will be for exact comparison only (label to label with ILET intact). The important criteria is for the resolver to be able to pass along the DNS query to the corresponding authoritative server and obtain a correct response.

FULLY compliant resolvers will be able to process the DNSII identifer and know all the ILET values for full function name mapping. Cache name lookup will be fully enabled and inquiry fallback mechanism discussed in Section 3.2.2 SHOULD be performed in the event of encountering a non-compliant server.

4.3 Authoritative Server Conformance Levels

Authoritative servers MUST be fully compliant before attempting to serve multilingual sub-domains under its authoritative zone. They should however prepare for the transition towards a multilingual name space even if they do not intend to deploy it right away.

The BASIC compliancy for authoritative name servers is to allow an 8-bit clean approach towards sub-domains that are not directly under its authority (i.e. sub-sub-domains).

The INTERMEDIATE compliant name server will be able to process the DNSII identifier and ILET without rejecting the query. The authoritative zone as well as its direct sub-domains however SHOULD

not include the use of the DNSII flags because ILET values are not understood at this compliancy level.

FULLY compliant name servers will be able to handle DNSII compliant label formats at its sub-domain levels. That is, fully compliant root servers will be able to serve multilingual TLDs, fully compliant TLD servers will be able to serve multilingual SLDs, etc.

5. Transition Schedule & Strategy

DNSII is designed to allow a gradual adoption of multilingual domain names on the Internet. The transition strategy is therefore geared to be demand-pull instead of a technology-push incentive. However, to provide a platform for a demand-pull approach, it is required for operators to first safeguard their system. The simple approach as laid out in Section 4 is to propose that servers take a 8-bit clean approach on name resolution.

As discussed in DNSII-MDNP, it is reasonable for the deployment of DNSII-MDNP at the registry level first to draw demand for the service and let the host level DNSs with multilingual names to begin migration first. DNS operators around the world should however prepare for the transition and begin the deployment of DNSII depending on their interest in serving multilingual domain names, according to the conformance levels laid out in Section 4, beginning from BASIC compliancy for operators that are least interested to a FULLY compliant server for operators who wishes to provide multilingual capabilities for their users.

The root servers could easily be adjusted to be a BASIC compliant authoritative name server. Once the demand is proven and the stability of the system tested, they too could transition to fully compliant authoritative servers so that multilingual TLDs could be rolled out.

6. Summary of Discussion

This document introduces the contemplated transition and steady state resolution process for multilingual domain names with a DNSII compliant format. Two tunneling mechanisms using the TXT RR was introduced for the preservation of information during transitional resolution.

6.1 Client/Application Resolution Strategy

```
Send Query in DNSII format
IF RCODE = Format Error
     THEN send query in UTF-8/Local Encoding AND append DNSII RR
     IF RCODE = Format Error
          THEN send Query in ASCII with "-for-tunneling-only-" label
          AND append DNSII RR
          AND check for DNSII ANS RR in response
```

```
     ELSE proceed and check for DNSII ANS RR in response
ELSE proceed as usual
```

6.2 Resolver Resolution Strategy

```
(*)IF incoming request is in pure DNSII format
     THEN resolve according to ILET in cache and by recursive lookup
     IF encounter RCODE = Format Error
          THEN send query in UTF-8 AND append DNSII RR
          IF RCODE = Format Error
               THEN send query in ASCII with "-for-tunneling-only-"
                    label
               AND append DNSII RR
               AND check for DNSII ANS RR in response
          ELSE proceed and check for DNSII ANS RR in response
     ELSE proceed as usual with pure DNSII Format (*)
     AND respond in pure DNSII format
ELSE IF incoming request has tunneled MDNP information
     THEN resolve using the information in the appended DNSII RR
          Reset Query using DNSII Format and go through (*)
     AND convert back to tunneling format before responding to query
          With DNSII ANS RR appended to response
     AND set TTL for regular RRs in the Answer field to be = 0
ElSE IF incoming request is in the original "00" label format
     AND no tunneled information is included
     AND the label contains characters beyond A-z, 0-9 or "-"
     THEN force convert all labels to UTF-8
     AND query using DNSII Format and go through (*)
ELSE proceed as usual (existing ASCII based names)
```

6.3 Authoritative Name Server Resolution Strategy

```
IF incoming request is in pure DNSII format
     THEN resolve according to ILET
     AND respond in pure DNSII format
ELSE IF incoming request has tunneled MDNP information
     THEN resolve using the information in the appended DNSII RR
     AND convert back to tunneling format before responding to query
          With DNSII ANS RR appended to response
     AND set TTL for regular RRs in the Answer field to be = 0
ELSE use InZone DNSII mechanisms AND use 8-bit clean approach
```


7. Security Considerations

DNSII RRs will be secured through transaction authentication, while DNSII ANS RRs could have their own SIG RRs. In general, the DNSII-MDNR should not constitute any extra burden on DNS security.

8. Intellectual Property Considerations

It is the intention of Neteka to submit the DNSII protocol and other

elements of the multilingual domain name server software to IETF for review, comment or standardization.

Neteka Inc. has applied for one or more patents on the technology related to multilingual domain name server software and multilingual email server software suite. If a standard is adopted by IETF and any patents are issued to Neteka with claims that are necessary for practicing the standard, any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specifications under fair, reasonable and non-discriminatory terms.

Other DNSII related documents and discussions could be found at http://www.dnsii.org.

9. References

[DNSII-MDNP] E. Chung & D. Leung "DNSII Multilingual Domain Name Protocol", August 2000

[RFC1700] J. Reynolds, J. Postel, "ASSIGNED NUMBERS", RFC 1700, October 1994.

[ISO10646] ISO/IEC 10646-1:2000. International Standard -- Information technology -- Universal Multiple-Octet Coded Character Set (UCS)

[RFC1034] Mockapetris, P., "Domain Names - Concepts and Facilities," STD 13, RFC 1034, USC/ISI, November 1987

[RFC1035] Mockapetris, P., "Domain Names - Implementation and Specification," STD 13, RFC 1035, USC/ISI, November 1987

[RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels," RFC 2119, March 1997

Authors:

Edmon Chung
Neteka Inc.
2462 Yonge St. Toronto,
Ontario, Canada M4P 2H5
edmon@neteka.com

David Leung
Neteka Inc.
2462 Yonge St. Toronto,
Ontario, Canada M4P 2H5

david@neteka.com