Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 24, 2016

J. Yao
X. Lee
CNNIC
P. Vixie
CNNIC-Farsight Joint Laboratory
May 23, 2016

Bundled DNS Name Redirection
draft-yao-dnsext-bname-06.txt

Abstract

This document defines a new DNS Resource Record called "BNAME", which provides the capability to map itself and its subtree of the DNS name space to another domain. It differs from the CNAME record which only maps a single node of the DNS name space, from the DNAME which only maps the subtree of the DNS name space to another domain.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on November 24, 2016.

Table of Contents

1. Introduction

For some names, the internet users may want them to be identical in the DNS resolution. For example, exmaple.ong==example.ngo, cnnic.cn==cnnic.net.cn. The BNAME represents for bundle names. This document defines a new DNS Resource Record called "BNAME", which provides the capability to map an entire tree of the DNS name space to another domain. It means that the BNAME redirects both itself and its descendants to its owner. The DNAME [RFC6672] does not redirect itself, only the descendants. The domain name that owns a DNAME record is allowed to have other resource record types at that domain name. The domain name that owns a BNAME record is not allowed to have other resource record types at that domain name unless they are the DNSSEC related resource record types defined in [RFC4033], [RFC4034], [RFC4035] and [RFC5155]. A server MAY refuse to load a zone that has data at a sub-domain of a domain name owning a BNAME RR or that has other data except the DNSSEC related resource record types and BNAME at that name. BNAME is a singleton type, meaning only one BNAME is allowed per name except the DNSSEC related resource record types. Resolvers, servers and zone content administrators should be cautious that usage of BNAME or its combination with CNAME or DNAME may lead to form loops. The loops should be avoided.

1.1. Terminology

All the basic terms used in this specification are defined in the documents [RFC1034], [RFC1035] and [RFC2672].

2. Motivation

CNAME can redirect itself to other name. DNAME can rediret its children to other name. In practice, many names need redirect itself and its children to another name. For example, we expect exmaple.TLD1 to be identical with the example.TLD2 in the DNS resolution. Without the BNAME mechanism, current mechanisms such as DNAME or CNAME are not enough capable to redirect itself and its children to another name at the same time.

3. The BNAME Resource Record

3.1. Format

The BNAME RR has mnemonic BNAME and type code xx (decimal). Its RDATA is comprised of a single field, <target>, which contains a fully qualified domain name that must be sent in uncompressed form [RFC1035], [RFC3597]. The <target> field MUST be present. The presentation format of <target> is that of a domain name [RFC1035]. The wildcards in the BNAME RR SHOULD NOT be used.

```
<owner> <ttl> <class> BNAME <target>
```

The effect of the BNAME RR is the substitution of the record's <target> for its owner name, as a suffix of a domain name. This substitution has to be applied for every BNAME RR found in the resolution process, which allows fairly lengthy valid chains of BNAME RRs.

3.2. The BNAME Substitution

A BNAME substitution is performed by replacing the suffix labels of the name or the whole name being sought matching the owner name of the BNAME resource record with the string of labels in the RDATA field. The matching labels end with the root label in all cases. Only whole labels are replaced.

3.3. The BNAME Rules

There are two rules which governs the use of BNAMEs in a zone file. The first one is that there SHOULD be no descendants under the owner of the BNAME. The second one is that no resource records can co-exist with the BNAME for the same name except the DNSSEC related resource record types. It means that if a BNAME RR is present at a node N, there MUST be no other data except the DNSSEC related resource record types at N and no data at any descendant of N. This restriction applies only to records of the same class as the BNAME record.

3.4. BNAME Examples

The table below shows some examples of the BNAME usage.

QNAME	owner	BNAME	target	result
com.	example.com.	example.net.	<no match>
com.	com.	net.	net.
example.com.	example.com.	example.net.	example.net.
a.example.com.	example.com.	example.net.	a.example.net.
a.b.example.com.	example.com.	example.net.	a.b.example.net.
ab.example.com.	b.example.com.	example.net.	<no match>
bar.example.com.	example.com.	example.net.	bar.example.net.
a.b.example.com.	b.example.com.	example.net.	a.example.net.
a.example.com.	example.com.	b.example.net.	a.b.example.net.

Table 1. BNAME Usage Examples.

If the owner name of the CNAME RR is regarded as the target of the MX RR, it may cause some problems. Some mail servers may directly connect the owner name of the CNAME instead of the name pointed by CNAME for mail delivery and cause the undelivery of the mails. BNAME has the similar problems with CNAME. This document specifies that the owner name of the BNAME should not be the targets of some RRs such as MX, SRV and PTR. In case that the owner name of the BNAME RR is the target of some RRs, it should be cautious.

4. Query Processing

To exploit the BNAME mechanism the name resolution algorithms [RFC1034] must be modified slightly for both servers and resolvers. Both modified algorithms incorporate the operation of making a substitution on a name (either QNAME or SNAME) under control of a BNAME record. This operation will be referred to as "the BNAME substitution".

4.1. Processing by Servers

For a server performing non-recursive service steps 3.a, 3.c and 4 of section 4.3.2 [RFC1034] are changed to check for a BNAME record, and to return certain BNAME records from zone data and the cache.

If the owner name of the bname is the suffix of the name queryed, or same, when preparing a response, a server performing a BNAME substitution will in all cases include the relevant BNAME RR in the answer section. A CNAME RR is synthesized and included in the answer section. This will help the client to reach the correct DNS data.

The server MUST know BNAME. The provided synthesized CNAME RR if there has one, MUST have

The same CLASS as the QCLASS of the query,

TTL equal to the corresponding BNAME RR,

An <owner> equal to the QNAME in effect at the moment the BNAME RR was encountered, and

An RDATA field containing the new QNAME formed by the action of the BNAME substitution.

The revised server algorithm is:

1. Set or clear the value of recursion available in the response depending on whether the name server is willing to provide recursive service. If recursive service is available and requested via the RD bit in the query, go to step 5, otherwise step 2.

2. Search the available zones for the zone which is the nearest ancestor to QNAME. If such a zone is found, go to step 3, otherwise step 4.

3. Start matching down, label by label, in the zone. The matching process can terminate several ways:

a. If the whole of QNAME is matched, we have found the node.

If the data at the node is a CNAME, and QTYPE doesn't match CNAME, copy the CNAME RR into the answer section of the response, change QNAME to the canonical name in the CNAME RR, and go back to step 1.

If the data at the node is a BNAME, and QTYPE doesn't match BNAME, copy the BNAME RR
and also a corresponding,
synthesized CNAME RR
into the answer section of the
response, change QNAME to the name carried as RDATA in the BNAME RR, and go back to step 1.

Otherwise, copy all RRs which match QTYPE into the answer section and go to step 6.

b. If a match would take us out of the authoritative data, we have a referral. This happens when we encounter a node with NS RRs marking cuts along the bottom of a zone.

Copy the NS RRs for the subzone into the authority section of the reply. Put whatever addresses are available into the additional section, using glue RRs if the addresses are not available from authoritative data or the cache. Go to step 4.

c. If at some label, a match is impossible (i.e., the corresponding label does not exist), look to see whether the last label matched has a BNAME record.

If a BNAME record exists at that point, copy that record into the answer section. If substitution of its <target> for its <owner> in QNAME would overflow the legal size for a <domain-name>, set RCODE to YXDOMAIN [RFC2136] and exit; otherwise perform the substitution and continue. The server SHOULD synthesize a corresponding CNAME record and include it in the answer section. Go back to step 1.

If there was no BNAME record, look to see if the "*" label exists.

If the "*" label does not exist, check whether the name we are looking for is the original QNAME in the query or a name we have followed due to a CNAME. If the name is original, set an authoritative name error in the response and exit. Otherwise just exit.

If the "*" label does exist, match RRs at that node against QTYPE. If any match, copy them into the answer section, but set the owner of the RR to be QNAME, and not the node with the "*" label. Go to step 6.

4. Start matching down in the cache. If QNAME is found in the cache, copy all RRs attached to it that match QTYPE or BNAME RR into the answer section. If QNAME is not found in the cache but a BNAME record is present at an ancestor of QNAME, copy that BNAME record into the answer section. If there was no delegation from authoritative data, look for the best one from the cache, and put it in the authority section. Go to step 6.

5. Use the local resolver or a copy of its algorithm (see resolver section of this memo) to answer the query. Store the results, including any intermediate CNAMEs and BNAMEs, in the answer section of the response.

6. Using local data only, attempt to add other RRs which may be useful to the additional section of the query. Exit.

Note that there will be at most one ancestor with a BNAME as described in step 4 unless some zone's data is in violation of the no-descendants limitation of the owner of the BNAME. If some descedants are found when a BNAME record is encountered, the server can stop search of step 3c or step 4.

4.2. Processing by Resolvers

A resolver or a server providing recursive service must be modified to treat a BNAME as somewhat analogous to a CNAME. The resolver algorithm of [RFC1034] section 5.3.3 is modified to renumber step 4.d as 4.e and insert a new 4.d. The complete algorithm becomes:

1. See if the answer is in local information, and if so return it to the client.

2. Find the best servers to ask.

3. Send them queries until one returns a response.

4. Analyze the response, either:

   a. if the response answers the question or contains a name error, cache the data as well as returning it back to the client.

   b. if the response contains a better delegation to other servers, cache the delegation information, and go to step 2.

   c. if the response shows a CNAME and that is not the answer itself, cache the CNAME, change the SNAME to the canonical name in the CNAME RR and go to step 1.

   d. if the response shows a BNAME and that is not the answer itself, cache the BNAME and the CNAME if there has one. If substitution of the BNAME's <target> for its <owner> in the SNAME would overflow the legal size for a <domain-name>, return an implementation-dependent error to the application; otherwise perform the substitution and go to step 1.

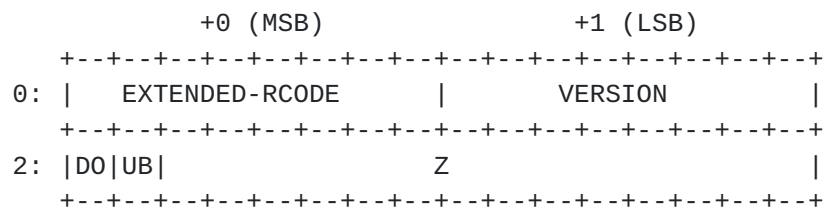
   e. if the response shows a server failure or other bizarre contents, delete the server from the SLIST and go back to step 3.

A resolver or recursive server which understands BNAME records but sends non-extended queries MUST augment step 4.c by deleting from the reply any CNAME records which have an <owner> which is a subdomain of the <owner> of any BNAME record in the response.

5. Signaling of BNAME

A new UB (Understand BNAME) bit in the EDNS flags field [RFC2671]can be used to signal that the resolvers can understand BNAME. BNAME aware resolvers can set the Understand-BNAME (UB bit) to receive a response without the synthesized CNAME or DNAME. The UB bit is part of the EDNS extended RCODE and Flags field[RFC2671]. Resolvers MUST set this in a query to know BNAME.

Below are Updated EDNS extended RCODE and Flags fields [RFC2671]:



6. BNAME in DNSSEC

6.1. BNAME-aware Resolvers

With the deployment of DNSSEC, more and more servers and resolvers will support DNSSEC. In order to make BNAME valid in DNSSEC verification, the DNSSEC enabled resolvers and servers MUST support BNAME.

The BNAME aware resolvers MUST set DO bit and UB bit when sending DNSSEC queries to servers. The synthesized CNAME in the answer section for the BNAME may not be signed if there has one. DNSSEC validators MUST understand BNAME, verify the BNAME and then checking that the CNAME was properly synthesized in order to verify the synthesized CNAME. The BNAME enabled resolver (validator) should do somewhat analogous to a CNAME for further query.

In any negative response, the NSEC or NSEC3 [RFC5155] record type bit map SHOULD be checked to see that there was no BNAME that could have been applied. If the BNAME bit in the type bit map is set and the query type is not BNAME, then BNAME substitution should have been done.

6.2. Compatibility with BNAME unaware resolvers

In order to have a compatibility with BNAME unaware resolvers, the BNAME aware servers receiving queries from BNAME unaware resolvers with DO bit set but no UB bit set should do the following things if BNAME is put into the response and the query type is not BNAME:

- Issue the corresponding CNAME signature when querying the same owner name with BNAME based on the question name, and put into the answer section.

- Issue the corresponding DNAME and its signature when querying the children of the same owner name of BNAME based on the question name, and put into the answer section.

In order to satisfy the BNAME query, the server should prepares the siganture of CNAME and DNAME of the owner name of the BNAME beforehand.

The BNAME unaware resolvers with DNSSEC enabled are supposed to neglect the BNAME RR. If the corresponding CNAME signature is found, the validators will use it to verify the CNAME. If the corresponding CNAME signature is not found, but the corresponding DNAME with signature is found, the validators will use it to verify the CNAME.

7. IANA Considerations

This document updates the IANA registry "DOMAIN NAME SYSTEM PARAMETERS" (http://www.iana.org/assignments/dns-parameters) in sub-registry "TYPES", by defining one new type. IANA is requested to assign the number to BNAME.

8. Security Considerations

CNAME, DNAME, and BNAME may form a loop chain, which will cause the unresolvable of some names. The BNAME should avoid point to some name which is the owner name of CNAME or DNAME RRs.

9. Acknowledgements

Because the BNAME is very similar to DNAME, the authors learn a lot from [RFC2672]. Many ideas are from the discussion in the DNSOP and DNSEXT mailling list. Thanks a lot to all in the list. Many important comments and suggestions are contributed by many members of the DNSEXT and DNSOP WGs. The authors especially thanks the following ones:Niall O'Reilly, Glen Zorn, Mark Andrews, George Barwood,Olafur Gudmundsson, Sun Guonian and Hanfeng for improving this document.

10. Change History

[[CREF1: RFC Editor: Please remove this section.]]

10.1. draft-yao-dnsext-bname: Version 00

- Bundle DNS Name Redirection

10.2. draft-yao-dnsext-bname: Version 01

- Improve the algorithm
- Improve the text

10.3. draft-yao-dnsext-bname: Version 02

- Add the DNSSEC discussion
- Improve the text

10.4. draft-yao-dnsext-bname: Version 03

- Update the DNSSEC discussion
- Update the IANA consideration

10.5. draft-yao-dnsext-bname: Version 04

- Improve the text

10.6. draft-yao-dnsext-bname: Version 05

- add section: bname examples
- add section: Signaling of BNAME

10.7. draft-yao-dnsext-bname: Version 06

- redesign with DNSSEC verification
- Issue the corresponding CNAME signature when querying the same owner name with BNAME based on the question name when UB is not set
- Issue the corresponding DNAME and its signature when querying the children of the same owner name of BNAME based on the question name when UB is not set

11. References

11.1. Normative References

[ASCII] American National Standards Institute (formerly United States of America Standards Institute), "USA Code for Information Interchange", ANSI X3.4-1968, 1968.

[EDNS0] Vixie, P., "Extension Mechanisms for DNS (EDNS0)", RFC 2671, August 1999.

[RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, DOI 10.17487/RFC1034, November 1987, <http://www.rfc-editor.org/info/rfc1034>.

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <http://www.rfc-editor.org/info/rfc1035>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2136] Vixie, P., Ed., Thomson, S., Rekhter, Y., and J. Bound, "Dynamic Updates in the Domain Name System (DNS UPDATE)", RFC 2136, DOI 10.17487/RFC2136, April 1997, <http://www.rfc-editor.org/info/rfc2136>.

[RFC2671] Vixie, P., "Extension Mechanisms for DNS (EDNS0)", RFC 2671, DOI 10.17487/RFC2671, August 1999, <http://www.rfc-editor.org/info/rfc2671>.

[RFC2672] Crawford, M., "Non-Terminal DNS Name Redirection", RFC 2672, DOI 10.17487/RFC2672, August 1999, <http://www.rfc-editor.org/info/rfc2672>.

[RFC3490] Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications (IDNA)", RFC 3490, March 2003.

[RFC3597] Gustafsson, A., "Handling of Unknown DNS Resource Record (RR) Types", RFC 3597, DOI 10.17487/RFC3597, September 2003, <http://www.rfc-editor.org/info/rfc3597>.

[RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", RFC 3629, November 2003.

[RFC3743] Konishi, K., Huang, K., Qian, H., and Y. Ko, "Joint Engineering Team (JET) Guidelines for Internationalized Domain Names (IDN) Registration and Administration for Chinese, Japanese, and Korean", RFC 3743, DOI 10.17487/RFC3743, April 2004, <http://www.rfc-editor.org/info/rfc3743>.

[RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, DOI 10.17487/RFC4033, March 2005, <http://www.rfc-editor.org/info/rfc4033>.

[RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Resource Records for the DNS Security Extensions", RFC 4034, DOI 10.17487/RFC4034, March 2005, <http://www.rfc-editor.org/info/rfc4034>.

[RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, DOI 10.17487/RFC4035, March 2005, <http://www.rfc-editor.org/info/rfc4035>.

[RFC5155] Laurie, B., Sisson, G., Arends, R., and D. Blacka, "DNS Security (DNSSEC) Hashed Authenticated Denial of Existence", RFC 5155, DOI 10.17487/RFC5155, March 2008, <http://www.rfc-editor.org/info/rfc5155>.

[RFC6672] Rose, S. and W. Wijngaards, "DNAME Redirection in the DNS", RFC 6672, DOI 10.17487/RFC6672, June 2012, <http://www.rfc-editor.org/info/rfc6672>.

11.2. Informative References

[RFC2672bis] Rose, S. and W. Wijngaards, "Update to DNAME Redirection in the DNS", ietf-dnsext-rfc2672bis-dname-17.txt (work in progress), 6 2009.

Authors' Addresses

Jiankang YAO
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing

Phone: +86 10 58813007
Email: yaojk@cnnic.cn

Xiaodong LEE
CNNIC
No.4 South 4th Street, Zhongguancun
Beijing

Phone: +86 10 58813020
Email: xl@cnnic.cn

Paul Vixie
CNNIC-Farsight Joint Laboratory
4 South 4th Street,Zhongguancun,Haidian District
Beijing, Beijing 100190
China

Phone: +1 650 489 7919
Email: vixie@fsi.io