Bundle DNS Name Redirection
draft-yao-dnsext-bname-01.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 maps a single node of the DNS name space, from the DNAME which maps the subtree of the DNS name space to another domain.

Status of this Memo

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

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/ietf/1id-abstracts.txt.

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

This Internet-Draft will expire on July 6, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal
Provisions Relating to IETF Documents
(http://trustee.ietf.org/license-info) in effect on the date of
publication of this document. Please review these documents
carefully, as they describe your rights and restrictions with respect
to this document. Code Components extracted from this document must
include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the BSD License.

This document may contain material from IETF Documents or IETF
Contributions published or made publicly available before November
10, 2008. The person(s) controlling the copyright in some of this
material may not have granted the IETF Trust the right to allow
modifications of such material outside the IETF Standards Process.
Without obtaining an adequate license from the person(s) controlling
the copyright in such materials, this document may not be modified
outside the IETF Standards Process, and derivative works of it may
not be created outside the IETF Standards Process, except to format
it for publication as an RFC or to translate it into languages other
than English.
Table of Contents

1. Introduction .................................................. 3
  1.1. Terminology ............................................. 3
2. Motivation ..................................................... 3
3. The BNAME Resource Record ...................................... 3
  3.1. Format ................................................... 4
  3.2. The BNAME Substitution .................................. 4
  3.3. The BNAME Rules ......................................... 4
4. Query Processing .............................................. 4
  4.1. Processing by Servers .................................... 5
  4.2. Processing by Resolvers .................................. 7
5. IANA Considerations ........................................... 8
6. Security Considerations ....................................... 8
7. Acknowledgements .............................................. 9
8. Change History ................................................ 9
  8.1. draft-yao-dnsext-bname: Version 00 ..................... 9
  8.2. draft-yao-dnsext-bname: Version 01 ..................... 9
9. References ..................................................... 9
  9.1. Normative References .................................... 9
  9.2. Informative References .................................. 10
Authors’ Addresses .............................................. 10
1. Introduction

More and more internationalized domain name labels [RFC3490] appear in the DNS trees. Some labels [RFC3743] are equivalent in some languages. The internet users want them to be identical in the DNS resolution. For example, color.example.com==colour.example.com. 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 [RFC2672] and [RFC2672bis] do 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. 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 BNAME at that name. BNAME is a singleton type, meaning only one BNAME is allowed per name.

1.1. Terminology

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

2. Motivation

In some languages, some characters has the variants, which look differently or very similar but are identical in the meaning. For example, Chinese character U+56FD and its variant U+570B look differently, but are identical in the meaning. If Internationalized Domain Label" or "IDL" [RFC3743] are composed of variant characters, we regard this kind of IDL as the IDL variant. If these IDL variants are put into the DNS for resolution, they are expected to be identical in the DNS resolution. More comprehensible example is that we expect color.example.com to be equivalent with the colour.example.com in the DNS resolution. The BNAME Resource Record and its processing rules are conceived as a solution to this equivalence problem. Without the BNAME mechanism, current mechanisms such as DNAME or CNAME are not enough capable to solve all the problems with the emergence of internationalized domain names. The internationalized domain names may have alias or equivalence of the original one.

3. The BNAME Resource Record
3.1. Format

The BNAME RR has mnemonic BNAME and type code xx (decimal). It is not class-sensitive. 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].

<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 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. It means that if a BNAME RR is present at a node N, there MUST be no other data at N and no data at any descendant of N. This restriction applies only to records of the same class as the BNAME record.

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. 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 reach the correct DNS data. The provided synthesized CNAME RR, 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 as described above 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 into the answer section. If QNAME is not found in the cache but a BNAME record is present at 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 in section 3. An implementation might take advantage of this limitation by stopping the search of step 3c or step 4 when a BNAME record is encountered.

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. 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. IANA Considerations

IANA is requested to assignment the number to XX.

6. Security Considerations

There will have more discussions related to DNSSEC [RFC4033], [RFC4034] and [RFC4035]in the future version.
7. 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 mailing list. Thanks a lot to all in the list. Many important comments and suggestions are contributed by many members of the DNSEXT and DNSOP WG. The authors especially thanks the following ones:Niall O’Reilly, Glen Zorn for improving this document.

8. Change History

[[anchor13: RFC Editor: Please remove this section.]]

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

- Bundle DNS Name Redirection

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

- Improve the algorithm
- Improve the text

9. References

9.1. Normative References


9.2. Informative References

[RFC2672bis]
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: lee@cnnic.cn

Paul Vixie
Internet Software Consortium
950 Charter Street
Redwood City, CA
Phone: +1 650 779 7001
Email: vixie@isc.org