Bond
Internet Systems

REPORT ON THE ASSESSMENT OF SECURITY AND STABILITY IMPLICATIONS OF THE USE OF DNAME RESOURCE RECORDS IN THE ROOT ZONE OF THE DNS

Prepared for ICANN

by Bond Internet Systems SL

TABLE OF CONTENTS

Executive summary

This report describes the results of a study on the impact of using DNAME DNS records to support isomorphic TLDs at the top levels of the DNS tree (TLDs) on the Internet.

A captive test environment has been created using a local copy of the current root zone as well as a current copy of a sizeable gTLD (based on the .org TLD), which have been modified for the tests with the introduction of appropriate records as well as the use of locally generated DNSSEC keys to enable tests with signed and unsigned zones.

The captive environment comprised 13 root servers, a diverse set of servers for the test TLDs and a pair of additional servers for a second level domain, as well as a set of DNS resolvers that query this infrastructure.

Several different DNS server implementations were used during testing to assess their respective behaviour in each function, as applicable.

The impact on DNSSEC resolution has also been examined as part of this study.

The outcome of these tests shows that commonly used DNS server software employed on the Internet today for root and TLD authoritative service correctly serves the DNS information, though with variations. The fallback mechanism for DNAME defined in the standard (CNAME synthesis) allow the resolution process and, where available, the DNSSEC validation process to succeed.

Recursive DNS servers exhibit broader differences in handling of the DNS messages containing DNAME records and their fallback CNAME records, but in no instance do these differences prevent resolution of names or lead to incorrect results.

Introduction

The option of using DNAME Records [1] in the root zone as a means to create isomorphic TLDs without data replication has raised questions about its potential impact in actual operations, which from a technical point of view would be due to the behaviour of DNS resolver/caching servers and their treatment of DNAME Records or synthesised CNAME records.

The DNAME DNS Resource Record definition provides an elegant way to cross-link two sub-trees of the DNS tree by redirecting searches from one to the other for all names corresponding to searches below the label where the DNAME occurs. The name where the redirection takes place remains visible at all times.

The specification for the DNAME record also provides a fallback mechanism to provide support for the receivers of DNS responses that may not yet understand this new record type. In this mechanism, a CNAME record is synthesised by the DNS server (either authoritative or caching) providing the cross-link for the queried name, on a name by name basis. While this provides a means to ensure compatibility, it also lacks the general redirection properties of the DNAME record.

Additionally, the specification in RFC2672 calls for the TTL of the synthesised CNAME to be set to 0, preventing caching of the CNAME.

Recent work, in progress [2], aims to modify this behaviour to allow caching of the synthesised records.

Both options have a potential for impact in operations that is dependant on actual use of the names to be redirected.

A TTL=0 record will decrease the efficiency of DNS caches since it forces re-issue of queries every time the name is used. The impact will then be increased load on the DNS servers and increased latency experienced by the DNS client.

A non-0 TTL will alleviate these issues in exchange for an increase in the rate of utilisation of memory in the cache.

Concerns have also been raised linking the use of DNSSEC signatures, now in production at the root and in several TLDs and SLDs, and the ability to follow the trust paths when being redirect by a DNAME record, as well as the ability to sign the synthesised CNAME records.

This report describes tests undertaken to document observed behaviour in a replicated controlled environment that mimics the Internet DNS infrastructure.

Test scenarios

The test environment has been configured to test different scenarios with and without the use of DNAME records in the root zone and the interaction with DNSSEC.

In all cases, given that the current production DNS root zone is DNSSEC signed, the local root zone used for testing has been DNSSEC signed using the same parameters used in the production root zone: One 2048-bit RSA SHA256 KSK and one 1024-bit RSA SHA256 ZSK. Both keys have been generated locally for testing purposes only.

As test TLD a copy of the .org TLD zone obtained from its registry for research purposes has been used. From this zone file a simple modification yields a new test TLD, in this case .misc.

As in the case of the root zone above, DNSSEC keys are generated locally for test purposes only with parameters in accordance with those documented and observed for the .org TLD. Of note in this case is the fact that NSEC3 opt-out is used for signing the zones as this is the most commonly observed situation for DNSSEC signing in bigger TLDs.

As second level domains in these zones, the bondis.org domain has been used (and the corresponding mirror bondis.misc synthesised domain)

Authoritative DNS servers used in these tests both at the root and TLD level are:

BIND 9.7.x and NSD 3.2.x, both being the most common versions of software in use today to provide service at the root and TLD level (in the case of the root, exclusively).

Other DNS server platforms used in TLD service, such as Verisign's Atlas or Neustar Ultra DNS could not be tested in this environment as they are not generally available.

While DNS server software used for serving the root zone is well known and so is the distribution of DNS Server at the TLD and lower levels (see for instance the ISC Domain Survey [3]), there is no quantitative information on the installed base of recursive DNS servers. Our operational experience indicates that the list of DNS servers below is well aligned with usage on the Internet.

Recursive DNS servers used in this tests are:

ISC BIND 8.4.7, 9.6.X and 9.7.3-P2

NLNetLabs Unbound 1.4.9

Windows 2008 server R2

djbdns dnscache 1.0.5

Additionally, DNAME processing has been tested on Nominum's recursive server outside the testbed, using a subdomain of a real domain, introducing a DNAME record in the bondis.org zone, as could be done anywhere else in the DNS tree today. DNAME processing has been found to follow the same pattern as NLNetLabs' Unbound server described later in this report.

Case 1 - Discrete TLD delegations - unsigned TLD

This is the baseline case, where no DNAME records are seen in the captive DNS tree.

TLDs and subsequent domains are delegated from the root downwards using only classical DNS NS delegations.
Additionally, no DNSSEC features are used **below** the root level in the test environment.

The root zone itself is DNSSEC-signed, with DNSKEY and RRSIG parameters equal to those observed in the production root zone, as this is the observed reality in the DNS today.

This scenario provides a “plain DNS” scenario that is useful to validate any differences in the other, feature-rich, scenarios in the other tests.

A copy of the .org TLD obtained for research purposes has been slightly modify to point delegation of specific SLDs to the test environment.

Separately, a copy of this modified zone file is again modified to change the TLD from .org into .misc (a fictitious TLD existing only in this test environment) leaving all other data untouched.

Case 2 - Discrete TLD delegations - signed TLD

This study case builds on the previous one by adding DNSSEC features to the test TLDs. Delegation from the root are still performed using standard NS records, into the TLDs, then from TLDs to SLDs, etc.

The test TLDs used for this case have been signed using NSEC3 and opt-out features as observed in most big TLDs that make use of DNSSEC today, and in particular the .org domain, whose zone file has been used as the basis for these tests.

Case 3 - DNAME used - unsigned TLD

Case 3 is the first to study the introduction of DNAME records in the root zone as a means of redirecting an entry that would look like a TLD into an existing TLD.

In order to separate the effects of the introduction of the DNAME from those of its interaction with DNSSEC, this case makes use of unsigned TLD zones.

The zone file used for the .org TLD is the same as in Case 1. There is no zone file for the .misc TLD as it is not a TLD in the sense of being a discrete part of the DNS tree. Rather, the DNAME record for .misc is introduced in the root zone.

Case 4 - DNAME used - signed TLD

This last case brings together the introduction of DNAME records in the root zone, just as in Case 3, but looks into TLDs that are signed and allows inspection of the validation chain in this scenario.

Test results

In this section we present the results of observations made in the previously described study cases using different combinations of DNS server and client software.

The responses provided by DNS servers where recorded and examined for correctness.

The caching behaviour of the DNS servers for each response was also examined.

Finally, where applicable, the process of DNSSEC validation and the integrity of the trust chain is also examined.

Case 1 - Discrete TLD delegations - unsigned TLD

Correctness

As expected, all DNS servers behaved as expected in this case, with both TLDs operating normally and searches proceeding down each branch of the DNS tree, as normally observed on the Internet.

DNSSEC validation

The only zone signed in this first test is the root zone, which validated correctly when resolver software was provided with the appropriate trust anchor (the locally generated KSK for the test version of the root zone) except in the following cases:

- Windows 2008 R2 Server does not include support for DNSSEC algorithm #8 (RSA/SHA-256) at the time of this report. Therefore it cannot be used as a DNSSEC-validator name server with a starting point at root zone trust anchor.
- DJBDNS dnscache does not implement DNSSEC validation.

Caching behaviour

The effectiveness of DNS caching was observed for comparison with the later test cases. As expected, all records where cached following normal DNS behaviour.

Case 2 - Discrete TLD delegations - signed TLD

Correctness

As expected, all DNS servers behaved as expected in this case, with both TLDs operating normally and searches proceeding down each branch of the DNS tree, as normally observed on the Internet.

DNSSEC validation

The DNS caching servers capable of performing DNSSEC validation from the root behaved correctly with regards to DNSSEC validation, following the trust chain down to the requested records in the final zones (specific records inside an SLD within each TLD).
This test aims to provide a baseline test to verify that the test zones are correctly signed and the trust chain is correctly established through the use of correct DS, corresponding to the test keys in use.

Caching behaviour

As expected, all records where cached following normal DNS behaviour.

The effectiveness of DNS caching was observed for comparison with the later Case 4.

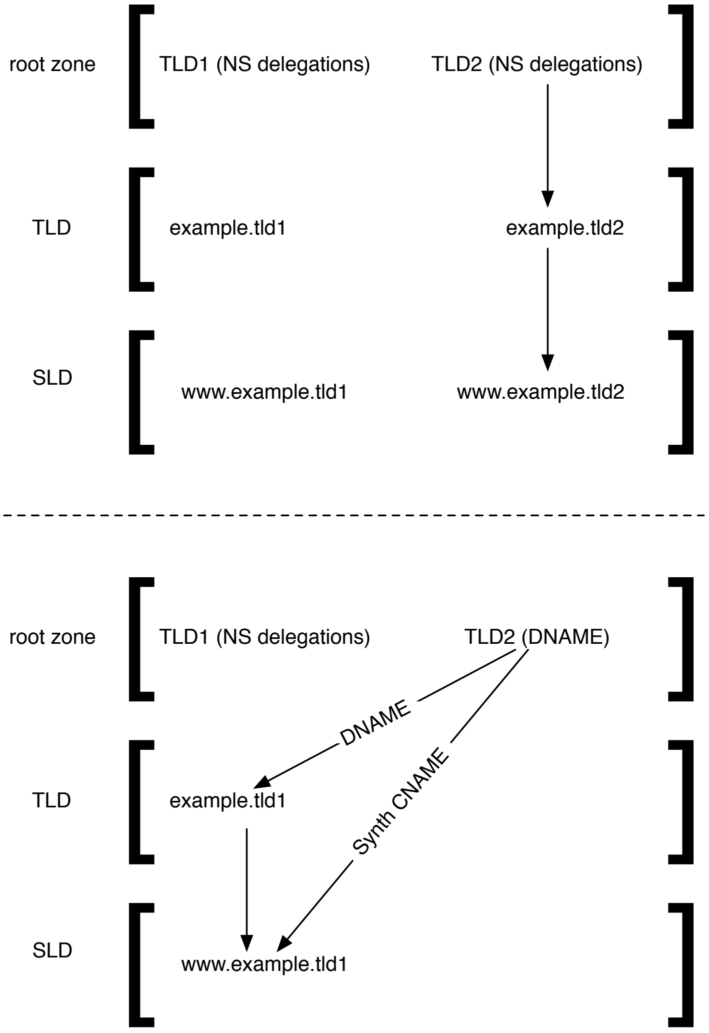
Case 3 - DNAME used - unsigned TLD

Correctness

The fallback mechanism that accompanies the DNAME definition allows for all tested DNS resolvers to process individual queries successfully and correctly.

The fallback mechanism is always active in authoritative server responses due to the specification in the standard, which calls for CNAME synthesis in the absence of EDNS in the query received at the server or an EDNS version of 0. Given that currently only version 0 of EDNS is defined all name server software falls in one of these categories.

There is, however, great variability in responses obtained from name servers when using DNAME to redirect a TLD to another. This exhibited variability is mainly on whether then DNAME record is passed through the DNS server to the client or downstream caching server and on the TTL of the associated synthesised CNAME records.
The values of the TTL on the CNAME record have an impact on the caching efficiency in the DNS system.



Resolving www.example.tld2 in a normal NS delegate zone and in a zone redirected with DNAME.

Authoritative servers

In the case of ISC's BIND, early versions with support for DNAME handling synthesised CNAME records with an associated TTL of 0 seconds, thus preventing caching of these records.
However, starting with versions BIND 9.6.2 and 9.7.0 (two family versions that run in parallel), the TTL of the synthesised CNAME is set to the TTL of the DNAME Record in the server.
This means that the CNAME will be cached, and re-used during the specified lifetime. As with any other records, if the CNAME is being obtained from a cached instance, its TTL, as reported in the query, will have been decremented from the value in the authoritative server during its presence in the cache.
This behaviour is described in an internet draft that aims to succeed RFC 2672 [2].

All versions of NSD with support for DNAME records and CNAME synthesis generate TTL=0 responses for the CNAME, preventing caching.

As a result of this, the efficiency of caching for a given DNAME-redirected domain (in this case a TLD) is highly dependant on the mix of DNS server software used to provision the DNS service for that TLD. If using one of the versions that generate non-0 TTL responses, caching will be enable, otherwise they will not be.
It must be said that the presently valid RFC defining DNAME does require a 0-TTL response and that departure from this behaviour is a leap of faith from the implementer's side given the non-final status of the update. Nonetheless, the change in behaviour does not cause interoperability concerns and does have the potential to mitigate concerns on the load generated by the CNAME synthesis process.

Recursive caching servers

The main differences between the tested recursive caching DNS servers lay in whether they have support for DNAME records, in which case they ignore the synthesised CNAME provided by the authoritative servers and will generate their own in response to client queries, or they lack support for DNAME records, in which case they are not used when received and only the CNAME is used (e.g .djbdns dnscache).
Some recursive servers that do not implement support for DNAME records will report a SERVFAIL error when querying directly for the DNAME'ed TLD at the root. This is however not a query that would take place in normal operation during DNS resolution, being more of a debugging query, which would require the DNS administrator to query the authoritative server directly using common DNS examination tools such as DiG.

DNSSEC validation

The DNS caching servers capable of performing DNSSEC validation with a trust anchor corresponding to the test KSK in the root zone are able to verify the records in the root zone, including the DNAME record itself.
As the TLDs are not signed, no further validation work is done in this study case.
Direct queries for the DNAME record will provide a fully signed answer. Queries for names below the DNAME will result in responses including a signed DNAME record and an unsigned CNAME record, as the authoritative servers provisioning the DNS server do not have the capability to generate “online signatures”.

Caching behaviour

In the case of the recursive DNS server shipped with Windows Server 2008, which does not currently support DNAME records, the CNAME is used and cached, even when the original query contained a 0 TTL CNAME record. Investigation indicates that the minimum SOA TTL value is applied to the CNAME record and is with that TTL that the record is cached.
If the authoritative server from which the CNAME is obtained is a version of BIND that sets the CNAME TTL value to that of the DNAME record then this non-0 value is used for caching.

Unbound and BIND, which understand the semantics of the DNAME record, will cache the record. When using the cached record to synthesise CNAMES, they will follow different rules for the assigned TTL, with BIND using the remaining TTL on the cached DNAME to provide the TTL for the CNAME, and Unbound always setting the CNAME TTL to 0.

This results in very different, and difficult to predict caching efficiency patterns for the records, depending on which authoritative server they were originally sourced from and which caching server is handling the client (or lower tier caching server) queries.

Case 4 - DNAME used - signed TLD

Correctness

As with case 3, the DNAME to CNAME fallback mechanism allows all DNS server software to resolve correctly by following the CNAME chain for the cases where the resolvers do not support the DNAME record directly.

The differences between this case and case regard DNSSEC validation which is described below.

DNSSEC validation

The DNS caching servers capable of performing DNSSEC validation from the root behaved correctly with regards to DNSSEC validation, following the trust chain down to the requested records in the final zones (specific records inside an SLD within each TLD).
There are several things to note in the query returned by a caching validator in this scenario.
Below is a sample query output from a caching validator (unbound) for a query looking for a web site. The caching server is performing validation based only upon the trust anchor for the root zone in use.

```
; <<>> DiG 9.7.3 <<>> @127.0.0.1 www.bondis.misc. +dnssec
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 16953
;; flags: qr rd ra ad; QUERY: 1, ANSWER: 6, AUTHORITY: 3, ADDITIONAL: 3

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
```

```
;; QUESTION SECTION:
;www.bondis.misc.       IN    A

;; ANSWER SECTION:
misc.                   172770 IN    DNAME  org.
misc.                   172770 IN    RRSIG  DNAME 8 1 172800 20110526183822
20110426183822 19791 . jMyfVfHSSGDV7l0HBA0fbGh4M5knUQ5f4SCq7vJiJRdX52VbXNW72ux1
cLYAKNjfXynRyfvENREkMevb1WpfDKA5/9sU94NB2KdsntdA+mWeXWni iFX/Q0xWxAEVhJwSw0O/qqVPnDMcv
+0I6hLFeAPIZs497wTjkt4leAnv Wjk=
misc.                   172770 IN    RRSIG  DNAME 8 1 172800 20110526183822
20110426183822 57378 . HP6tBrNKcqQ07OvNaLlxvHsj1yb83bQuAO1y3IM3o9Wr16gfZH5BB2Jk
ZjoIo7t1Wban+4kYOo1/SxEUaewmkmZj3SKWVYszCArlfcRLP1lzgIkG prY/281cthQNF/Bk+
+IPYsslHlO9gVapIR+a72jzsPriwhclJ34kTxhs Okc=
www.bondis.misc. 0      IN    CNAME  www.bondis.org.
www.bondis.org.         259   IN    A      194.176.119.250
www.bondis.org.         259   IN    RRSIG  A 5 3 300 20110515090000 20110414214221
40583 bondis.org. tJyzSRwv2kc3fKBQXglEZ3w1pr7tPCOAdeIqCjUXmwyGVnMrYjZHS872
mUXwBAj7bXwceXZ47+7A7ls7Vs2/awY1MLOTCx50LLaKmhW3wAykx2le
zVeoSs01KnHBTwH4iyRzXMtRbnrYuzFf4BNHzz2xrS0RGxKZqbn24fkz NKI=

...
;; MSG SIZE  rcvd: 994
```

The same query processed by a BIND caching validator yields similar results, with differing TTL for the CNAME record.

```
; <<>> DiG 9.7.3 <<>> www.bondis.misc a +dnssec
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 820
;; flags: qr rd ra ad; QUERY: 1, ANSWER: 6, AUTHORITY: 3, ADDITIONAL: 3

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
;; QUESTION SECTION:
;www.bondis.misc.       IN    A

;; ANSWER SECTION:
misc.                   172623 IN    DNAME  org.
misc.                   172623 IN    RRSIG  DNAME 8 1 172800 20110526183822
20110426183822 19791 . jMyfVfHSSGDV7l0HBA0fbGh4M5knUQ5f4SCq7vJiJRdX52VbXNW72ux1
cLYAKNjfXynRyfvENREkMevb1WpfDKA5/9sU94NB2KdsntdA+mWeXWni iFX/Q0xWxAEVhJwSw0O/qqVPnDMcv
+0I6hLFeAPIZs497wTjkt4leAnv Wjk=
misc.                   172623 IN    RRSIG  DNAME 8 1 172800 20110526183822
20110426183822 57378 . HP6tBrNKcqQ07OvNaLlxvHsj1yb83bQuAO1y3IM3o9Wr16gfZH5BB2Jk
ZjoIo7t1Wban+4kYOo1/SxEUaewmkmZj3SKWVYszCArlfcRLP1lzgIkG prY/281cthQNF/Bk+
+IPYsslHlO9gVapIR+a72jzsPriwhclJ34kTxhs Okc=
www.bondis.misc. 172623 IN    CNAME  www.bondis.org.
www.bondis.org.         124   IN    A      194.176.119.250
www.bondis.org.         124   IN    RRSIG  A 5 3 300 20110515090000 20110414214221
40583 bondis.org. tJyzSRwv2kc3fKBQXglEZ3w1pr7tPCOAdeIqCjUXmwyGVnMrYjZHS872
mUXwBAj7bXwceXZ47+7A7ls7Vs2/awY1MLOTCx50LLaKmhW3wAykx2le
zVeoSs01KnHBTwH4iyRzXMtRbnrYuzFf4BNHzz2xrS0RGxKZqbn24fkz NKI=
...
;; MSG SIZE  rcvd: 991
```

As expected, the synthesised CNAME is not signed, as the keys are not generally available to the authoritative servers (the root servers in this case) to enable them to generate signatures dynamically.
This does not affect the process as the validation process follows the redirection and is able to restart the validation based on the trust chain leading from the root to the .org TLD and further down.
The final result is clearly marked as Authenticated Data (the ad bit is set in the response from the validator) in the message header.

Conclusions

All the tests performed in this study indicate that the introduction of DNAME records as a means of creating isomorphic TLDs does not prevent name resolution using software in common use today.

This includes DNS resolution involving IDN names, which once provisioned in a DNS system ,are represented by ASCII strings generated by applying the Punycode algorithm (RFC 3492) to the normalised UTF-8 label with the pre-pending of the "xn--" prefix. This mechanism transform IDN labels into plain ascii representations that the DNS handles as any other label. This applies to all software tested in this report.

The observed behaviour of DNS servers varies in the handling of the DNAME record and the caching properties of the synthesised CNAMEs but this does not affect the outcome of the resolution process.

Differences in caching behaviour may generate different load profiles in the DNS servers but **do not prevent** resolution completion.

Performance profiling of recursive servers is nearly impossible to carry out in a deterministic and reproducible manner due to the fact that information gathering by the recursive server is largely dependent on external factors, such as network access and remote server responsiveness, rather than purely local factors.

However, concerns have been raised about the impact of the 0 TTL on the synthesised CNAME record, as that TTL value implies that the record is not cacheable.

At present, some servers would see an increase in load with the introduction of DNAME records in the root zone. However, this increase is not focused on a few servers but rather gets **distributed** at various places in the network as follows:

Recursive servers that understand DNAME will cache it locally and further queries from their clients do not get sent to the authoritative servers as the **synthesis will be performed locally** by these recursive servers, rather than forwarded to the authoritative servers while the DNAME remains in the cache, therefore distributing the load throughout the Internet.

Authoritative servers would therefore only see an increase in query load from the recursive servers that do not implement any processing for DNAME records.

This situation is likely to change in the near term as discussions are underway at the IETF [2] to modify the original specification of the DNAME record, modifying the rules regarding the TTL of the synthesised CNAME record to follow those of other DNS records. Under this new specification the TTL of the CNAME record is set to that of the companion DNAME record, both in the authoritative and recursive servers. Notably, ISC's BIND beginning with version 9.6.2 (in the 9.6 family) and 9.7.0 (in the 9.7 family) already implement this new behaviour.

With regard to DNSSEC validation, the tested servers that could perform DNSSEC validation using keys like the ones used in the Internet's root zone and TLDs have universal support for DNAME record processing and act appropriately by going up in the DNS chain after a signed DNAME redirection to continue the validation process on sound ground and **complete the validation process successfully**.

Finally, while hard to quantify without well known query patterns for the potential new TLDs (popularity, etc) the introduction of DNAME in the root zone is highly **unlikely** to generate any significant impact in the load of root servers.

References

[1] RFC 2672. Non-Terminal DNS Name Redirection,M. Crawford (1999)

[2] Update to DNAME Redirection in the DNS, S. Rose, W. Wijngaards. http://tools.ietf.org/html/draft-ietf-dnsext-rfc2672bis-dname-22. This is an internet draft and is cited for reference only. It is work in progress at the IETF.

[3] ISC Domain Survey. http://www.isc.org/solutions/survey