DNS Transport Size Issues in IPv6 Environment

Kenji Rikitake
Security Laboratory
KDDI R&D Laboratories, Inc.
2-1-15 Ohara, Kamifukuoka City
Saitama 356-8502 JAPAN
kenji@kddilabs.jp

Hiroki Nogawa
Cybermedia Center
Osaka University
1-5 Mihogaoka, Ibaraki City
Osaka 567-0047 JAPAN
nogawa@cmc.osaka-u.ac.jp

Toshiaki Tanaka
Security Laboratory
KDDI R&D Laboratories, Inc.
2-1-15 Ohara, Kamifukuoka City
Saitama 356-8502 JAPAN
tl-tanaka@kddi.com

Koji Nakao
Information Security Department
KDDI Corporation
3-10-10, Iidabashi, Chiyoda-ku
Tokyo 102-8460 JAPAN
ko-nakao@kddi.com

Shinji Shimojo
Cybermedia Center
Osaka University
1-5 Mihogaoka, Ibaraki City
Osaka 567-0047 JAPAN
shimojo@cmc.osaka-u.ac.jp

Abstract

The nature of data exchanged between DNS (Domain Naming System) servers and resolvers has been changed since the introduction of IPv6 (Internet Protocol version 6) because of the increased address length and other technical enhancements. These changes and recent security concerns raise questions against fundamental design principles of the DNS protocols, especially to the transport layer and the payload size. In this paper we review how DNS protocols will be affected by the requirements and production-level deployment of IPv6, and how the new requirements for larger-size transfer should be satisfied by enhancing and/or redesigning the existing DNS transport layer protocols.

1. Introduction

DNS has been one of the core functions of Internet and will be in the coming era of IPv6. As IPv6 is going to be adopted to many consumer appliances such as Internet telephone and home electronic devices, the number of devices which depends on DNS will largely increase, and the DNS will become more complex.

On the other hand, DNS transport protocols, which has been defined by Mockapetris [11, 12] in 1987 and updated by a later Internet Standard [15] in 1989, have not been changed for more than 15 years. Since the functionality of database query and answer of DNS is no more than a retrieval of a set of mostly-static global distributed database, it has been solidly and reliably working, even under various systematic distributed DoS (Denial-of-Service) attacks such as the one occurred on October 21, 2002 [20].

The entire DNS database query protocol, however, is designed upon an assumption that the length of data exchanged in each query and answer mostly does not exceed a few hundred bytes. While this assumption works on the current Internet infrastructure, the recent changes such as the deployment of IPv6 whose addresses must be specified by an AAAA RR (Resource Record) for each 16-byte address, introduction of security enhancements such as DNSSEC [7], and the exchange of arbitrary texts using the TXT RRs for various purposes, sets a trend for exchanging larger amount of data over DNS database queries and answers.

Proceedings of the 2004 International Symposium on Applications and the Internet Workshops (SAINTW'04)
0-7695-2050-2/04 $20.00 © 2004 IEEE

In this paper, we first review the current DNS transport protocols and the limitation on Section 2, and show and some examples of the real-world DNS database exchange traffic in Section 3. We later discuss the analysis of the sampled traffic and possible change of DNS transport protocols to meet the demands of the trend for larger DNS database exchange in Section 4, and conclude this paper in Section 5.

2. DNS Transport Protocols

In this paper we call the set of communication protocols used between two DNS programs as the *DNS Transport Protocols*. The protocols are collectively defined by the two DNS RFCs of RFC1034 [11], which specifies the architecture, and RFC1035 [12], which specifies the implementation details. One of the Internet Host Requirement RFC, RFC1123 [15], also specifies the DNS Transport Protocol usage. Both UDP [13] and TCP [14] are used for the Transfer Protocols,

DNS Transport Protocol has the following two functions:

Zone Transfer: the retrieval action of Zone database files, which is sets of master information for RRs of a specific domain. The Zone Transfer is commonly used between DNS servers to obtain redundancy against a possible server failure. This function is performed solely over TCP.

RR Queries and Answers: the exchange between the servers and resolvers for retrieving the DNS RRs, which is essential for the application software, such as to resolve domain names to IP addresses. Most of the queries is over UDP, though TCP is also allowed, supported and used.

RFC1035 Section 4.2.1 restricts the maximum size of UDP queries and answers to 512 bytes. If this limit is exceeded, the answer is truncated and the TC bit in the Header section defined in RFC1034 Section 4.1.1. Since the answer is truncated, the answer is given solely for the notification to retry the query by TCP, which is capable to return longer answers.

In this paper we focus solely on the AAAA RRs [18] for IPv6 address resolution method from domain names. While the other address resolution method using A6 and DNAME RRs has been proposed [5], the proposal is now considered experimental and the AAAA records are considered preferable for the production deployment of IPv6 (RFC3363 [4] Section 2.a)) after an extensive discussion in IETF DNSEXT and NGTRANS working groups [1].

When a major transition from IPv4 to IPv6 occurs, the DNS Transfer Protocols will be affected due to the size increase of the RR represents IP addresses. Two major changes should be addressed as follows:

- For resolution from domain names to IP addresses, while the A RR for IPv4 has only a 4-byte RDATA field (RFC1035 Section 3.4.1), the AAAA RR for IPv6 has a 16-byte RDATA field (RFC1886 Section 2.2).
- For the reverse-lookups from IP addresses to domain names, while a domain name under the `in-addr.arpa` zone used for performing reverse-lookups of IPv4 (RFC1034 Section 5.2.1) takes 28 bytes maximum per name (such as `123.234.111.222.in-addr.arpa`), the IPv6 name under the domain `ip6.arpa` (RFC3152 [3]) takes 40 bytes maximum per name (such as `f.e.d.c.b.a.9.8.7.6.5.4.3.2.1.0.ip6.arpa`).

The size increase due to the introduction of AAAA RR and the `ip6.arpa` reverse-lookup namespace will affect the volume of data transferred through the DNS Transfer Protocols. The volume increase will not largely affect the Zone Transfer functions for the following reasons:

- No limit for the transferred zone data has been defined in the protocol specification of RFC1034 and RFC1035.
- Duplication of a gigabyte-size file is not a difficult task over the current production-level Internet, while the size of Zone Data files do not exceed a few hundred megabytes even on a large-scale domain such as `.com`, and in most of the cases the size is much smaller.

On the other hand, the volume increase will directly cause increase of data for the RR queries and answers as follows:

- Change of the reverse-lookup namespace from `in-addr.arpa` to `ip6.arpa` will increase the size of each reverse-lookup query by 12 bytes maximum per each RR, though the message compression (RFC1035 Section 4.1.4) will help reducing the size of reducing the upper-level portion of domain names shared inside the same query and answer messages.
- Adding or replacing a new RR for a new IPv6 address for existing hostname increases the size of DNS database answers as follows:
  - To replace an IPv4 address (A RR) to an IPv6 address (AAAA RR), 12 more bytes required; (from 4- to 16-byte address data)
  - To add another IPv6 address (AAAA RR) to an entry which has an IPv4 address, at least 28 more bytes required.
    (2 for compressed domain name index, 10 for the rest of RR header, and 16 for the IPv6 address data for the AAAA RR)

Recent introduction of security and other types of data into DNS database even make the transaction data length larger. For example, Gudmundsson [8] claims in RFC3226 Section 2.1 that DNSSEC will put a public key signature on each RR set, whose size ranges in size of 80 to 800 bytes, though most will be smaller or equal to 200 bytes.

The 512-byte limitation imposed to the UDP queries and answers of DNS now severely conflicts with the demand for more increased size. One of the well-known problem due to this limitation is the expandability of The Root Zone. The answer to the non-recursive Root Zone SOA RR request to a Root Server has already reached the limit; the returned 1 SOA, 13 NS and 13 A RRs for the 13 Root Servers consumes 493 bytes in total.

If another server were added to the Root Zone the SOA answer to the current `root-servers.net` zone with one letter prefix such as `z.root-servers.net` (which is non-existent in the real world), 15 more bytes for the NS RR and 16 more bytes for A RR would be added to the answer even if the message compression works, so the answer would become 524 bytes and exceeds the 512-byte limit.

If the existing 13 Root Servers added an AAAA record for the allocated IPv6 addresses, 28 more bytes would be added to each Root Server, so the answer would become $(493 + 28 \times 13) = 857$ bytes. This is obviously exceeding the 512-byte limit and must be handled by TCP transport of DNS.

To alleviate the 512-byte limit and allow longer UDP packets to be used on DNS Transport Protocols, Vixie [19] proposed an extension mechanism called EDNS0, which includes a pseudo-RR called OPT in RFC2671 for indicating the capability of accepting and sending longer UDP packets than 512 bytes. Though this proposal is effective for the compatible programs, old implementations still need to follow the fallback rule to TCP.

3. A Real-World Example of DNS Messages and A Simulated Result on Adding AAAA RRs

In this section, we analyze some real-world example of DNS Messages sampled from a campus network traffic. The samples are taken during 1149-1305JST of August 18, 2003, mirrored from the packet flow between the internal and external networks of the campus network. Excluding the error and non-answer DNS messages, we collected 285866 DNS UDP answer messages during the sampling period.

Using the Snort [10] and Tcpdump [17] packet analysis tools, we first measured the message size which is equal to the UDP payload size, the number of A RRs contained in each record.

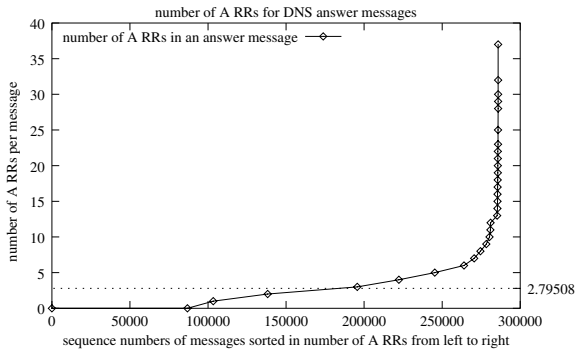

Figure 1. Number of A RRs in DNS Answer Messages

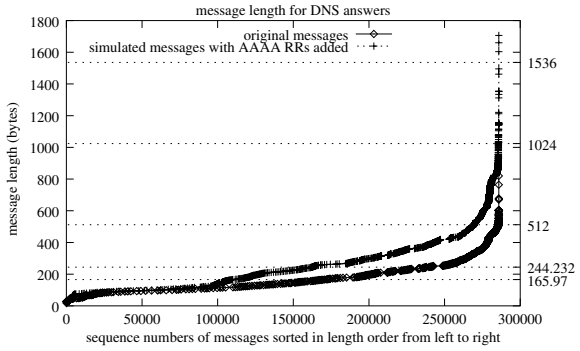

Figure 2. Simulation Result of DNS Message Sizes If an AAAA RR Added to Each A RR

Figure 1 shows the number of A RRs in the sampled DNS answer messages. The mean value of the 285866 samples was 2.79508, and the unbiased standard deviation was 2.79877. While ≈98% of the answers contained 9 or less A RRs, the maximum number of A RR in the samples was 37.

We decided to conduct a simple simulation of adding an AAAA RR for each A RR. During the IPv4-to-IPv6 migration period, many hosts would be requested to be reachable from both IPv4 and IPv6 networks. This indicates that a host which is only assigned an A RR should have at least one new AAAA RR assigned for the IPv6 reachability. The size of the entire DNS answer message increases 28 bytes for each a RR in the message, if the corresponding AAAA RR is added. While this could be an underestimation since many hosts would register more IPv6 addresses than the IPv4 addresses they had, this simulation suggests a practical maximum value of the DNS message size limit.

Figure 2 shows DNS answer message sizes of the samples and the simulated messages. If the message size fol-

	Sampled messages	AAAA-added messages
Mean value (μ)	165.97	244.232
Unbiased standard deviation (σ)	87.638	159.372
Probability of exceeding 512 bytes if the message size follows the normal distribution	≈0.04%	4.65%
Maximum value	820	1706

Table 1. Message Length in Bytes for The Sampled and AAAA-RR-added Simulated Messages

	DF bit set	DF bit *not* set
Number of packets	109075	176791
Percentage (%)	38.2	61.8

Table 2. Numbers and Percentages of The IP-Packet DF Bit Status for DNS Answer Messages

lows the normal distribution, we can estimate the probability of the message size exceeding 512 bytes by using the mean value μ and the unbiased standard deviation σ. Table 1 shows the statistically-analyzed values of the sampled and simulated message sizes.

We analyzed whether each UDP packet carrying the DNS message explicitly prohibits the IP-level fragmentation by setting the DF Bit. Table 2 shows the numbers and percentages of the DF Bit status.

During the DNS message analysis we notified that many DNS messages contained the OPT pseudo-RR proposed in EDNS0 [19]. Table 3 shows the result which contains the numbers of messages with the OPT pseudo-RRs, categorized by the payload size specified within.

4. Analysis and Protocol Modification Proposal for Larger DNS Messages

By the measurement and analysis in Section 3, we can suggest the following trends on the DNS answer messages:

- While most of the answers contained less than 10 A RRs, in some rare cases the number of RR reached up to 37. We also observed a case similar to this; a set of PTR records replied in a large DNS message carried by TCP, which has 1112 bytes of the payload, contained 42 PTR records for the reverse-lookup requests to the same IPv4 address, presumably for a virtual-domain Web service which shared the same IPv4 address by many domain names. While this sort of replying large number or RRs in the single DNS message is not a common practice yet, the DNS Transport Protocol should be able to reliably transfer the larger size messages. The value of 4096 bytes, shown in Table 3, will be practically sufficient.

Size value	# of messages
1280	29
2048	233
4096	63961

Total: 64223 messages
(22.5% of all messages)

Table 3. Numbers of The EDNS0 OPT Pseudo-RR-contained Messages and The Specified UDP Payload Size

- During the IPv4-to-IPv4 transition, the size of DNS messages will increase, since a globally-reachable host will be assigned at least one globally-reachable IPv6 address, which must be registered to the DNS database. As shown in Fig. 2 and Table 1, near 5% of the whole answers will exceed the 512-byte message size limit if all hosts become IPv4-and-IPv6 reachable, while currently less than 0.1% of the whole answers exceeds the limit. Since the DNS messages larger than 512 bytes are not handled by TCP, the number of IP packets for DNS database exchange will significantly increase due to this transition and impose some burden to the DNS servers, if no systematic solution such as the EDNS0 would be deployed Internet-wide.

- The use of EDNS0 is fairly popular if not a common practice, according to the result shown in Table 3. The popularity of BIND [9] DNS program may contribute to this, since BIND supports the EDNS0 extension. The percentage of 22.5%, however, is not sufficient to consider that EDNS0 is fully supported Internet-wide.

- About 38% of DNS answer packets were DF-bit set as shown in Table 2. This suggests that the defragmentation issue could be controlled by the application programs, using the IPv6 Fragment Header (in RFC2460 [6] Section 4.5).

To handle the size-increasing trend of DNS database messages, we propose to implement the following protocol modifications:

- Promoting the implementation of RFC2671, which defines the EDNS0 enhancement, which eventually leads

to the raised status;

- Reconsidering RFC1035 Section 4.2.1 to allow larger packets than 512 bytes for UDP query and answer exchange, with the suggested value analyzed from the real-world Internet traffic and forecasting;

- Introducing a faster TCP message handling method as an alternative for the hosts which cannot efficiently handle the large UDP packets. Using the T/TCP [2] is a solution should be considered, since it reduces the number of packets exchanged for the multiple transactions. We published an analysis for introducing T/TCP to DNS [16].

5. Conclusion

In this paper, we discussed the DNS transport protocol issues due to the size limitation imposed on the UDP exchange, and estimated how it affects the IPv4-to-IPv6 transition. We also performed a simulation of adding AAAA RRs using a set of real-world traffic samples, and concluded that the current limitation of 512 bytes for the DNS message exchange on UDP will significantly affect the performance of DNS queries and answers, since the probability of exchanged data exceeding the limit would increase from ≈0.04% to 4.65%. We also proposed the protocol modification to manage the size-increasing trend of DNS database exchange, including deployment of EDNS0 implementation which allows the longer UDP messages, and a faster TCP transaction protocol such as T/TCP.

6. Acknowledgements

Our thanks go to Mr. Tohru Asami, the president and CEO of KDDI R&D Laboratories, Inc., for supporting our research activities. We also thank Toshikazu Akiyama of Cybermedia Center, Osaka University, for his constructive comments on the general IPv6 issues on DNS.

References

[1] R. Austein. Tradeoffs in Domain Name System (DNS) support for Internet Protocol version 6 (IPv6), 2002. RFC3364.
[2] R. Braden. T/TCP – TCP extensions for transactions functional specification, 1994. RFC1644.
[3] R. Bush. Delegation of IP6.ARPA, 2000. RFC3152.
[4] R. Bush, A. Durand, B. Fink, O. Gudmundsson, and T. Hain. Representing internet protocol version 6 (IPv6) addresses in the Domain Name System (DNS), 2002. RFC3363.
[5] M. Crawford and C. Huitema. DNS extensions to support IPv6 address aggregation and renumbering, 2000. RFC2874.
[6] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) specification, 1998. RFC2460.
[7] D. Eastlake. Domain name system security extensions, 1999. RFC2535.
[8] O. Gudmundsson. DNSSEC and IPv6 A6 aware server/resolver message size requirements, 2001. RFC3226.
[9] Internet Software Consortium. BIND. `http://www.isc.org/bind/`.
[10] M. Roesch et al. Snort. `http://www.snort.org/`.
[11] P. V. Mockapetris. Domain names – concepts and facilities, 1987. RFC1034 (also STD13).
[12] P. V. Mockapetris. Domain names – implementation and specification, 1987. RFC1035 (also STD13).
[13] J. Postel. User datagram protocol, 1980. RFC768 (also STD6).
[14] J. Postel. Transmission control protocol, 1981. RFC793 (also STD7).
[15] R. Braden (Editor). Requirements for Internet hosts – application and support, 1989. RFC1123.
[16] K. Rikitake, K. Nakao, H. Nogawa, and S. Shimojo. T/TCP for DNS: A Performance and Security Analysis. *IPSJ Journal*, 44(8):2060–2071, Aug. 2003.
[17] TCPDUMP Public Repository. tcpdump. `http://www.tcpdump.org/`.
[18] S. Thomson and C. Huitema. DNS extensions to support IP version 6, 1995. RFC1886.
[19] P. Vixie. Extension Mechanisms for DNS (EDNS0), 1999. RFC2671.
[20] P. Vixie, G. Sneeringer, and M. Schleifer. Events of 21-Oct-2002. `http://f.root-servers.org/october21.txt`.