DNS authoritative server misconfiguration and a countermeasure in resolver

draft-fujiwara-dnsop-bad-dns-auth-01.txt

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of section 3 of RFC 3667. By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with RFC 3668.

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 April 18, 2005.

Copyright Notice

Copyright (C) The Internet Society (2004).
Abstract

This memo describes misconfigurations of DNS authoritative name server and its effect of increasing the load in DNS resolver server. In some cases we recommend re-checking DNS authoritative servers with a viewpoint of current RFC and propose corresponding changes in DNS resolver server implementations to protect it. And more, we point open issues.

The recommendations made in this document are based on analysis of abnormal DNS resolver server load at large ISP cache server which has many customers.
1. Introduction

This memo describes that combination of misconfigurations at authoritative servers can create significant overloads on resolver servers. Specifically, the combinations of large response size, non-use of EDNS0 option and TCP filtering at authoritative servers may increase the number of TCP SYN_SENT status in resolver servers and the load on the servers. This behavior was found through the observation of query traffic to/from ISP resolver servers [TOYAMA04].

Response sizes from DNS authoritative servers may grow as the use of IPv6 or DNSSEC spread [RFC3226]. In ENUM and SIP, many NAPTR resource records may be written to a domainname.

Thus the above combination and the anomalies in resolver servers will frequently occur.

While there are reports on the observations of query traffic to root or top-level domain servers and the recommendations to the resolver servers to reduce anomalies on the servers [WESSELS04], this memo intends to notify to the operators of authoritative servers that their configuration can lead overload on resolver servers.

In the following sections, we provide a detailed explanation of the problem. We then recommend to re-check the configurations of authoritative servers to avoid the problem. We also show an example of resolver sever modification in case that operators of authoritative servers are not cooperative or, in extreme, malicious in that they intentionally attempt to increase the load of resolver servers.

2. Problem Description

DNS message size is limited to 512 octets in UDP packet[RFC1035]. However, some response can exceed the limitation. A typical case observed is a response with PTR RRsets for an IP address which is assigned for many domain names [TOYAMA04]. Besides, spread of IPv6 and DNSSEC may also increase the cases.

If the authoritative server who returns such response do not support EDNS0 option [RFC2671], the server returns truncated response (TC bit = 1) to the query sent by a resolver server. Then the resolver server tries to get whole message by using TCP connection. A problem occurs if the authoritative name server filters TCP DNS port. In that case, because the resolver server cannot establish a TCP connection to the authoritative server, it keep the TCP SYN_SENT states for some interval. Because keeping many TCP states increase
the load of the resolver server, this phenomenon can significantly impact on the resolver server.

When there are multiple authoritative name servers for the record, the resolver server repeats the sequence for all the authoritative name servers, depending on the implementation of the resolver server (At least, we found that the BIND follows this sequence).

Finally, the resolver server responds with ServFail to stub resolver, which is not cached by both the resolver server and the stub resolvers.

3. Authoritative servers

In the viewpoint from resolver servers, authoritative servers MUST be configured correctly.

3.1 RRSet size

DNS responses which fit in 512 octet are carried by UDP packet. \[RFC1035\] This case is safe and light for DNS resolver servers. Larger responses are carried by TCP virtual circuit or EDNS0 UDP packet only.

3.1.1 Recommendation

DNS zone authors SHOULD write RRSet as small as possible and SHOULD NOT write useless RRs. And if they must write large RRSet which response packet size is larger than 512 octet, they MUST be especially careful to setup authoritative servers described in section 3.2 and 3.3.

3.2 TCP query issue

There are many authoritative servers which filter or reject TCP queries. There are many administrators who want to close DNS authoritative server TCP port. Many of them compared the server’s security and the issues caused by closing TCP port and they decide filtering TCP port.

But filtering DNS authoritative server TCP port may causes problems described in section 2. According to RFC1123 section 6.1.3.2 \[RFC1123\], DNS servers MUST be able to service UDP queries and SHOULD be able to service TCP queries.

3.2.1 Recommendation

DNS server administrator SHOULD re-check DNS authoritative server TCP
3.3 EDNS0

EDNS0 [RFC2671] relaxes the DNS message size limit.

As noted in RFC1123 Section 6.1.3.2, UDP queries have much lower overhead, both in packet count and in connection state.

To reduce TCP query cost, EDNS0 support is necessary.

3.3.1 Recommendation

DNS server administrator SHOULD support EDNS0 in their authoritative server if they write RRSet which response size exceeds 512 octets.

3.3.2 Consideration about EDNS0

RFC3226 [RFC3226] says all DNSSEC compliant servers and resolvers MUST support EDNS0.

When do we mandate EDNS0?

3.4 Considering TCP query necessity

There are many resolver servers and stub resolvers which does not support EDNS0, they cannot handle answers of 512 octet and more. Therefore, supporting TCP queries is mandatory now.

If EDNS0 is widespread and all of stub resolvers, resolver servers and authoritative servers support EDNS0, supporting TCP queries is still necessary because TCP is only method to send larger data more than path MTU when IP fragment is prohibited.


But RFC2460 "IPv6 Specification" [RFC2460] section 5 "Packet size issues" states that the use of such fragmentation is discouraged in any application that is able to adjust its packets to fit the measured path MTU. Furthermore, IP Fragmentation is one of weakpoints in IP.

There are some open issues.

EDNS0 requires IPv4 and IPv6 fragmentation support. Is it OK? Can we restrict TCP query when TC=1 in UDP in the future?
4. Resolver algorithm improvement

To avoid problems described in section 2, we tried to change existing resolver server not to query by TCP when truncation occurs and it reduces TCP sessions, but it cannot cache any data and it may violate RFC2181.

Then, we propose new resolver algorithm to cache misconfigured RRsSets and misconfigured servers and decrease originating TCP queries.

In iterative query, when resolver server gets truncated answer (TC=1), the resolver server should record the query’s information (servername, domainname, class, type, TC=1) and try to get whole message by using TCP connection. When the resolver server cannot establish a TCP connection to the authoritative server because of the authoritative server’s TCP port filtered, the resolver server should be marked as a bad authoritative server which does not answer TCP. And All the servers for a given domainname answer TC=1 and don’t answer by TCP, the resolver server answers RCODE=2 "Server failure" to the stub resolver and the domainname’s RRSets (domainname, class, type) could not be resolved. In this case, we propose that the domainname’s RRSets may be marked as a temporary unavailable for a while, such as 3600 second. The next query for the same RRSets will be answered as RCODE=2 "Server failure" when the cache entry exists and is marked as "a temporary available".

In efficiency, the resolver server should avoid TCP queries using the bad authoritative server information which does not answer TCP. This information’s lifetime may be 3600 second, for example.

This information is similar to the BIND 9’s EDNS0 unsupported host information.

Caching the unresolvable RRSets which is stored in the bad configured servers as "a temporary unavailable" for a while and answering RCODE=2 "Server failure" from cache may be protocol change.

5. Conclusion

In this document, we describe a observed anomaly of resolver servers caused by the combination of authoritative server misconfigurations; large RRSets, EDNS0 unsupport, and TCP filtering. Because size of RRSets tends to increase, which increase the frequency of this phenomenon, which can severely impact on resolver servers. Therefore, the operators of the authoritative server should re-check the configuration of their server. Meanwhile, we propose a modification of resolver server to protect against the phenomenon, which can be caused by intentional DoS attacks to the resolver
6. Security considerations

Misconfigurations of authoritative servers discussed in this document expose resolver servers to increased risk of intentional DDoS attacks.

Modification of the resolver servers discussed in this memo can reduce the risk.

References


[RFC2181] R. Elz and R. Bush, "Clarifications to the DNS

Authors’ Addresses

Kazunori Fujiwara
Japan Registry Service Co.,Ltd.
Chiyoda First Bldg. East 13F,
3-8-1 Nishi-Kanda Chiyoda-ku,
Tokyo 101-0065, JAPAN
Phone: +81-3-5215-8451
E-Mail: fujiwara@jprs.co.jp

Keisuke Ishibashi
Nippon Telegraph and Telephone Corporation
Information Sharing Platform Laboratories
3-9-11 Midori-cho
Musashino-shi, Tokyo 180-8585 Japan
Phone: +81-422-59-3407
E-Mail: ishibashi.keisuke@lab.ntt.co.jp

Katsuyasu Toyama
Nippon Telegraph and Telephone Corporation
Information Sharing Platform Laboratories
3-9-11 Midori-cho
Musashino-shi, Tokyo 180-8585 Japan
Phone: +81-422-59-7906
E-Mail: toyama.katsuyasu@lab.ntt.co.jp
Funding for the RFC Editor function is currently provided by the Internet Society.