DNS authoritative server misconfiguration and a countermeasure in resolver

draft-fujiwara-dnsop-bad-dns-auth-00.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 DNS authoritative name server misconfiguration and that results in a significant query cost 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 a combination of misconfigurations of authoritative servers may lead significant heavy load to resolver servers. This behavior was found through the observation of query traffic to/from resolver servers [NANOG32CACHE]. Here, the misconfiguration combinations are large RRset, EDNS0 unsupported and TCP filtering.

While there are reports on the observations of query traffic to root or top-level domain servers and recommendation to the resolver servers [1, 2], this memo intends to notify to the operators of authoritative servers that their configuration can lead heavy load on resolver servers.

As stated in [3], response size from DNS authoritative servers may grow as the use of IPv6 spreads. Thus the above combination and the anomaly in resolver server caused by the combination will frequently occur.

In the following sections, we describe the problem in detail. We then recommend to re-check configurations of authoritative servers to avoid the problem. We also show an example of resolver server 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 [RFC1035]. However, some response can exceed the limitation. One typical case is the response with PTR RRsets for an IP address which is assigned for many domain names. If the authoritative server who returns such response do not support EDNS0 option [RFC2671], the name 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 packets. In that case, because the resolver server cannot any establish TCP connections to the authoritative server, it hold the TCP SYN_SENT states for some interval. Because holding TCP SYN_SENT states increase the load of the resolver server, this phenomenon can significantly impact on resolver server.

When there are multiple name servers for the record, the resolver server repeats the sequence for all the 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 users with ServFail, which is not cached by both resolver server and 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, 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 setting and SHOULD configure the server to service TCP queries.

3.3 EDNS0

EDNS0[RFC2671] extends UDP payload size.

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 octet.

4. Resolver algorithm

4.1 TCP query necessity

There are many resolver servers and stub resolvers which do 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 than path MTU when IP fragment is prohibited.

4.2 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 servers. In iterative query, resolver server should cache truncated answers information (servername, domainname, class, type, TC=1) and bad authoritative server information which does not answer TCP. All the servers for a given domainname answers TC=1 and does not answer by TCP, the resolver server cache the domainname as "cannot resolve" for several period such as 3600 second.

5 Conclusion

In this document, we describe an observed anomaly of resolver servers caused by the combination of authoritative server misconfigurations; large RRset, EDNS0 unsupport, and TCP filtering. Because size of RRset tends to increase, which increase the frequency of this phenomenon, which can severly 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 server.

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


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
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

Acknowledgment

We would like to thank Ichiro Mizukoshi, Haruhiko Ohshima, Masahiro Ishino, Chika Yoshimura, Tsuyoshi Toyono, Hirotaka Matsuoka, Yasuhiro Morisita, and Bill Manning.
Funding for the RFC Editor function is currently provided by the Internet Society.