The Ecology of DNS Open Resolvers

Yuuki Takano, Ruo Ando, Satoshi Uda, Takeshi Takahashi, and Tomoya Inoue

***Abstract*—The DNS amplification attack is a DDoS attack by abusing DNS open resolvers as reflectors. Because the attack can amplify attacker's traffic dozens of times by reflectors, attackers can efficiently launch DDoS attack even though they have only low bandwidth network. Therefore, in this paper, we investigated DNS open resolvers abused by the DNS amplification attack from 3 aspects of wide active probing, silent monitoring and deploying open resolvers. As a result of our active probing, we found about 30 millions of DNS servers on the Internet, and 25 millions of them are open resolvers. Then, by silent monitoring and deploying open resolvers, we found that A record requests are often used by probing DNS servers, and ANY record requests are often used to launch DDoS attack. Moreover, we reveal that source network of the attackers could be traced by combining these results and BGP's routing information.**

Information

This paper was originally published in IEICE transactions, written in Japanese, and translated into English with the Google translation.

```
@article{IEICE:ytakano2014,
  author    = {Yuuki Takano and
               Ruo Ando and
               Satoshi Uda and
               Takeshi Takahashi and
               Tomoya Inoue},
  title     = "{The Ecology of DNS Open Resolvers}",
  journal   = {IEICE Transaction B},
  volume    = {J97-B},
  pages     = {873--889},
  number    = {10},
  year      = {2014},
}
```

I. Introduction

The Domain Name System (DNS) [1], [2] is a distributed database or name resolution system used to convert human readable host names and IP addresses, which is essential for today's Internet operation. DNS continues to be used with various updates up to the present in 2014, with specifications developed in RFC 1035 in the early days of the Internet in 1987. DNS spoofing attacks that make bogus DNS responses, and cache poisoning attacks of DNS cache servers have become major problems in DNS security so far.

The DNS spoofing attack is an attack using the fact that there is no authentication function in the DNS protocol, and it was actually reported that the censorship system called China's great fire wall has performed DNS spoofing attack [3]. DNS spoofing attacks have become a problem because they can be done relatively easily, and DNSSEC [4] thus has been proposed to solve this problem. Currently, many DNS server software implemented DNSSEC. Also, a very practical attack, called Kaminsky attack, as an attack on the DNS cache was reported in 2008 [5], which have also become a big problem.

Furthermore, following DNS spoofing attacks and DNS cache poisoning attacks, DDoS attacks that exploit DNS servers called DNS amp attacks [6] have gained attention. For example, 75 Gbps DDoS attacks using DNS amp attack received by Spamhous, a non-profit organization of anti-spam mail in 2013, became a big topic [7]. Originally, although DNS had a limitation of 512 bytes in packet size, it has become possible to transmit large inquiry and response packets exceeding 512 bytes by specification update [8]. As a result, DNS servers (open resolver) can be used as amplifiers for reflective DDoS attacks.

Therefore, in this research, we investigated open resolvers used as refrectors of the DNS amplification attack in detail from the three aspects of active scan, silent monitor observation, and actual open resolvers operation. Section II describes related works. Section III, we explain the active scan system of the DNS server we designed and implemented. Section IV and V show the results of the result of active scan. As a result of the active scan, it was revealed that about 30 million DNS servers exist on the Internet, of which about 25 million were open resolvers. Furthermore, by requesting VERSION.BIND to these servers, we obtained information of DNS server software type and version. Also, in Section VII, we show the result of observation by the silent monitor, and show the result of actual open resolvers operation in Section VIII. With the observation by the silent monitor, we observed many DNS server scan packets, and it became clear that DNS amp attacks were steadily performed as a result of the open resolver operation. In addition, this paper shows that by combining information obtained from the silent monitor, the actual open resolver we operated, and BPG, it is possible to specify the origin networks abused by attackers.

II. Related Work

In this section, we explain the DNS amplification attack and DNS server survey researches, and describe the differences between related works and this research.

A. DNS amplification attack

The DNS amplification attack is a type of DDoS attack that is performed by making use of the fact that there is a large difference in the size of the DNS query packet and response packet. According to the classification of RFC 4732 [9], it is considered to be a type of amplification attack like Smurf attack, TCP amplification attack, etc.

The DNS amp attack is performed by sending a query that misrepresents a source address as an attack target to a

Y. Takano†, R. Ando, and T. Takahashi are with National Institute of Information and Communications Technology. (†ytakano@wide.ad.jp)

Y. Takano, S. Uda, and T. Inoue are with Japan Advanced Institute of Science and Technology

plurality of DNS servers existing on the Internet. In general, DNS query packets has a size of several tens of bytes, but the response packets for a specific query such as ANY query is 3,000 bytes or more. Therefore, the traffic transmitted by the attacker is amplified by dozens of times and reaches the network to be attacked. The DNS amplification attack is performed using the DNS server on the Internet as amplifiers. Such DNS servers that exists on the Internet and answers recursive queries without any restrictions are commonly called DNS open resolvers.

DNS Open resolvers, which are the steppingstones of DDoS attacks, are considered to be distributed throughout the world because anyone can operate freely, but their ecology has not been grasped in detail. Also, to study the characteristics of the DNS amplification attacks and defense methods, it is important to grasp the ecology of open resolvers.

B. DNS Server Survey Researches and Projects

Open Resolver Project [10] is a project that actively investigates DNS servers worldwide from around March 2013 and regularly provides open resolver statistics on the web. From the page of this project, you can browse information of open resolvers from March 2013 until January 2014. Shadowserver [11] is an organization that conducts surveys related to Internet security, and they also conducts active measurements of open resolvers. The existence of this project could be found by operating the silent monitor described in Section VII. Similarly, Steve Sntorelli [12] also conducted investigation report of open resolves, but unlike the former two, it is not doing a wide area survey.

The former two projects roughly report the address distribution as a result of scanning the IPv4 address space. In this research, similarly, the result of scanning the DNS server on the Internet is described, and further detailed report on the software type, version information of DNS server, IP address inversion report information which the former two have not reported. In addition to the active scan, we operated the silent monitor and the actual open resolver, so in this paper we also report findings obtained from these operations.

III. Active Scan System for DNS servers

This section describes the system and architecture for actively scanning DNS servers on the Internet. Figure 1 is the system architecture of the DNS server scan system that we designed and implemented. This system consists of the following four components: DB, the DNS Prober, the Reverse Lookuper, and the Statistical Analyzer.

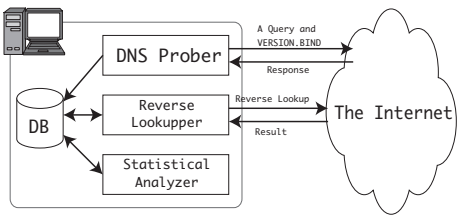


Fig. 1. System Architecture of DNS Server Crawler

DB

In this system, retrieved results and statistical information are stored in NoSQL database. The reason for adopting the NoSQL database is that NoSQL database can be easily used by rapid prototyping, and flexible records can be added regardless of restrictions such as column definition like SQL database. Moreover, MapReduce [13] adopted by some NoSQL databases can be used for data processing easily. In our implementation, we adopted MongoDB [14], which is equipped with MapReduce and high affinity with C++, Python etc.

DNS Prober

The DNS Prober is a component that sends a DNS A record request to the port 53 of all IPv4 global addresses and scans for DNS servers on the Internet. When scanning, by sending a query packet with RD flag [2] meaning recursive query request turned off to avoid loading the Internet as much as possible. If a response with the RA flag [2], which indicates that recursive inquiry is possible, is returned for this A record request, an open resolver exists at the address that returned the response. Conversely, when returning a response whose RA flag is turned off, it means that there is a DNS server that is not an open resolver.

If a response to the A record request is received, the DNS Prober subsequently sends a TXT record request with the VERSION.BIND key to the responded address. In some DNS server implementations, the software version of the DNS server is returned in response to the TXT record request of VERSION.BIND. Actually, this behavior can be confirmed by using the dig command as `$ dig @127.0.0.1 -t TXT -c CHAOS VERSION.BIND`.

We implemented the DNS Prober in C++. DNS Prober uses Boost [15], libevent [16], MongoDB C++ driver, and Catenaccio DPI [17], and all obtained results are saved in MonboDB. In addition, since the DNS Prober scans for all IPv4 global addresses, if it makes sequential inquiries, it takes a great deal of time to complete the process. Therefore, in this implementation, the query issuing process and the response process are separated to speed up the scan.

Reverse Lookupper

The Reverse Lookuper performs a reverse lookup of the IP address obtained by the DNS Prober and acquires the domain name associated with the IP address. We implemented the Reverse Lookuper in C++, and the Reverse Lookuper also uses Boost, libevent, MongoDB C++ driver, and Catenaccio DPI. The number of reverse addresses becomes tens of millions. Thus, we also optimized the Reverse Lookupper to handle lots of queries asynchronously using libevent. Similarly, all obtained domain names are stored in MongoDB.

Statistical Analyzer

Results obtained from the DNS Prober and the Reverse Lookuper are analyzed by the Statistical Analyzer. In our implementation, we used MapReduce provided by MongoDB for analysis. Since MongoDB provides an interface for MapReduce by JavaScript, JavaScript is thus mainly used in the implementation of the Statistical Analyzer, and Python is supplementarily used.

We used this system to investigate DNS servers in all IPv4 address space from 17:26 on July 5, 2013 to 19:38 on July 6 (JST). As a result, it became clear that there were more than 30 million DNS servers in the IPv4 address space, and nearly 25 million DNS servers were open resolvers. Also, DNS software and version information could be obtained from 7 million DNS servers. Details of these results are described in Section IV and V.

IV. DNS Server Software and Regional Distribution

This section describes the result of DNS server scan using the system described in Section III.

A. Distribution of DNS Server Software

We classified the obtained IPv4 addresses into each region Internet registry (RIR) [18], and classified the results obtained by the VERSION.BIND inquiry into server types using regular expressions as shown in Table I.

TABLE I
Regex for Detecting DNS Server Types

DNS type	regular expression
BIND 9.x	^9(\.[0-9])+
BIND 8.x	^8(\.[0-9])+
BIND 4.x	^4(\.[0-9])+
Dnsmasq	^dnsmasq
Nominum Vantio	^Nominum Vantio
Nominum ANS	^Nominum ANS
PowerDNS	^PowerDNS
Unbound	^unbound
NSD	^NSD
Windows series	.*Windows

Table II is the distribution for each RIR of DNS server types. The "can't detect" row shows the servers which could not be classified by the regular expressions, and the row of "no version info" shows the servers that returned an error for the VERSION.BIND inquiry.

As a result of our measurement, the number of servers responded to the VERSION.BIND inquiry was 15,357,412, of which the number of servers classified by the regular expressions was 7,075,527. However, since the respond to the VERSION.BIND inquiry can be set freely at the time of server operation, it may not always be the correct version or software. However, it is also noted that in many cases, the default setting will respond with its own software name and version. DNS servers are classified into servers that manage zones (content server, slave server), and servers (full service resolver, forwarder) answering requests from clients. In this paper, the former is called an authoritative DNS server, and the latter is called a resolving DNS server in order to clarify the discussion.

BIND series [19], Nominum ANS [20], PowerDNS [21], NSD [22] are DNS server software that can build an authoritative DNS server. Dnsmasq [23], Nominum Vantio [24], Unbound [25] are DNS server software for constructing only a resolving DNS server. From Table II, it reveals that NSD and Nominum ANS specialized only for authoritative DNS server had very small percentage of open resolvers. In the BIND series, which can be used for both authoritative and resolving DNS servers, and PowerDNS, the proportion of open resolvers was high, and many DNS servers dedicated for resolving DNS server were also open resolvers. Details of each server software are discussed in Section V.

B. Distribution of DNS Servers for each RIR

In this subsection, we discuss the distribution of DNS servers and open resolvers for each RIR. We totally obtained 30,285,322 DNS server addresses, and 24,971,990 (about 82.5 % of addresses) responses with the RA flag indicating open resolvers. Approximately 62.3 % of DNS servers on the Internet were present in APNIC and RIPE NCC, and of these, about 87 % were open resolvers. LACNIC and AFRINIC had fewer DNS servers than others, 5,149,451 and 1,205,748 addresses, respectively, which are considered to be due to the number of Internet users. About 96 % of DNS servers existing in LACNIC and AFRINIC were open resolvers, that is, almost all DNS servers existing in LACNIC and AFRICIN were open resolvers. There were 3,139,392 DNS servers in ARIN, but only about 54.8 % (1,720,185) of them were classified as open resolver. In comparison with other RIRs, the percentage was relatively low.

Figure 2 shows the distribution of DNS open resolvers, where latitude and longitude are calculated using is calculated from latitude and longitude GeoIP Lite [26] from obtained IP adresses. This figure reveals that open resolvers were distributed over a very wide range around the world.

V. Distribution of DNS Server Software and Software Version

In this section, we discuss the distribution of DNS server software and software version of BIND series, PowerDNS, Dnsmasq, Unbound, NSD, Nominum Vantio, and Nominum ANS.

A. BIND Series

BIND is the most famous DNS server software and is available for both authoritative and resolving DNS servers. In our measurements, we obtained 417, 86, and 71 software versions on the BIND 9.x, 8.x, and 4.x series, respectively. The version distributions of each BIND series are shown in Figure 3, 4, 5. Here, the package of BIND provided by RedHat is counted as another version. This is because the BIND software distributed for RedHat Linux is backported by RedHat Linux. At the time of measurement (July 2013), the latest versions of the BIND 9.x series were 9.9.3-P2, 9.8.5-P2, 9.7.7 (EOL), 9.6-ESV-R9-P1 [27], and the final versions of BIND 8.x and 4.x are 8.4.7 and 4.9.11.

Table II reveals that many of the BIND series were open resolvers, but because BIND can be operated only as an authoritative DNS server, half or more BIND series DNS servers were not open resolvers.

Although BIND 4.x and BIND 8.x are very old implementations and it was officially announced that BIND 8.x entered the

TABLE II
DISTRIBUTION OF DNS SERVER TYPES

Type of DNS	#	Total %	APNIC #	RIPE #	ARIN #	LACNIC #	AFRINIC #	other #
BIND 9.x	4268442	(14.1%)	806357	1530177	1126501	169268	121556	514583
†	1851362	(6.1%)	551458	781954	176399	94385	117906	129260
BIND 8.x	35218	(0.1%)	4588	21348	6663	974	32	1613
†	30444	(0.1%)	4202	18958	5186	854	31	1213
BIND 4.x	3486	(0.0%)	121	2751	440	43	0	131
†	2765	(0.0%)	93	2256	348	11	0	57
Dnsmasq	1308653	(4.3%)	692042	216273	75201	226880	32676	65581
†	1308381	(4.3%)	692026	216028	75196	226877	32676	65578
Nominum Vantio	968041	(3.2%)	553404	284852	20142	21205	70861	17577
†	967044	(3.2%)	552650	284782	20125	21200	70736	17551
Nominum ANS	687	(0.0%)	18	34	79	42	2	512
†	13	(0.0%)	2	0	0	11	0	0
PowerDNS	373588	(1.2%)	14215	329994	14360	2952	91	11976
†	372684	(1.2%)	14207	329116	14354	2952	91	11964
Unbound	71781	(0.2%)	16230	43507	6941	1510	1585	2008
†	23220	(0.0%)	3281	14398	4638	315	312	276
NSD	33933	(0.1%)	1731	11077	17182	322	13	3608
†	17	(0.0%)	5	5	2	1	0	4
can't detect	8281885	(27.3%)	4012525	2367711	429450	690618	279903	501678
†	7658656	(25.3%)	3911886	2118455	244682	670597	278183	434853
Windows series	11698	(0.0%)	184	1077	85	10312	0	40
†	11342	(0.0%)	129	865	67	10257	0	24
no version info	14927910	(49.3%)	3457029	4505928	1442348	4025325	699029	798251
†	12746062	(42.1%)	3050589	3465814	1179188	3919438	668399	462634
Total	30285322	(100.0%)	9558444	9314729	3139392	5149451	1205748	1917558
†	24971990	(82.5%)	8780528	7232631	1720185	4946898	1168334	1123414

†: DNS open resolvers
2013/07/05 17:26 - 2013/07/06 19:38 (JST)

end of life in August 2007 [28], but our measurement reveals that BIND 8.x and BIND 4.x were still exist on the Internet. Furthermore, Table II reveals that BIND 8.x and 4.x exist in RIPE NCC a lot. Note that, when we announced this result at GreHack 2013 conducted in France [29], one participant pointed out that many BIND 4.x series were being operated on RIPE NCC for honeypots.

B. PowerDNS

Figure 6 shows the version distribution of PowerDNS, and, we were totally obtained 22 different versions in our measurement. PowerDNS is distributed as an authoritative DNS server called PowerDNS Authoritative Server and a resolving DNS server called PowerDNS Recursor. This figure shows only 10 different versions, because PowerDNS has responded to VERSION.BIND inquiries from the most recent release, version 3.0 and later [30]. As of July 2013, the latest versions of PowerDNS Authoritative Server and PowerDNS Recursor were 3.3 and 3.5.2, respectively. From this figure, in the case of PowerDNS, you can see that the latest version was used most.

You can see from Table II that many servers classified as PowerDNS were open resolvers. Although PowerDNS can be used as an authoritative DNS server, it is thought that most of it was used as a resolving DNS server.

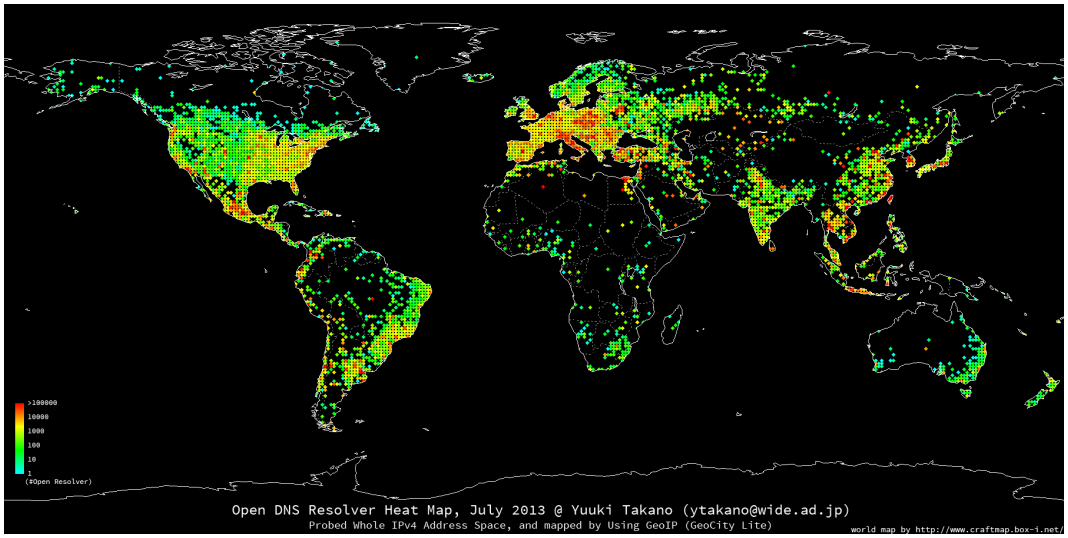
C. Dnsmasq

Dnsmasq is a lightweight DNS forwarder and DHCP software implementation for small networks. Although Dnsmasq is not assumed to be used as a large-scale resolving DNS server, Table II reveals that many Dnsmasq servers were open resolvers. Note that most open Dnsmasq servers on the Internet were open resolvers. Under NAT, DNS forwarders are indispensable in order to effectively utilize UDP and TCP ports, and Dnsmasq is often adopted by a home router that realizes NAT. Home routers and the like are installed in each household, and if this was an open resolver, it is difficult to restrict by Internet carriers. This fact reveals that it is difficult to fudamentaly solve the open resolver problem. Figure 7 shows the version distribution of Dnsmasq. In our measurements we totally obtained 86 versions. As of July 2013, the latest version of Dnsmasq is 2.66, but 2.66 is not shown in the figure.

D. Unbound and NSD

Unbound and NSD, both of which are being developed by NLnet Labs [31], are a DNS cache and an authoritative DNS server, respectively. Figure 8 and 9 show the version distribution of Unbound and NSD. In our measurement we totally obtained 30 Unbound versions and 42 NSD versions, and, as of July 2013, the latest version of Unboud, and NSD were 1.4.20, and 3.2.15, respectively. In addition, Figure 8 reveals that beta versions of NSD, 4.0.0b4 and 4.0.0 _imp_5, were operated on the Internet.

Table II shows that there were Unbound's open resolvers, but its proportion was lower than other server software. In most cases, Unbound accepts only inquiries from the local in the default installation state. In addition, even if you bind UDP



https://github.com/ytakano/pictures/blob/master/dns/open_dns_resolver_heatmap_201307.png

Fig. 2. World Heatmap of DNS Open Resolvers

socket used by Unbound to ANY address (0.0.0.0), unless you make special setting to respond to inquiries from all IP addresses, Unbound respond to only inquires from local addresses. Thus, in order to make Unbound an open resolver, a somewhat cumbersome procedure is required, so it seems that this resulted. On the other hand, since NSD only works as an authoritative DNS server, the proportion of being an open resolver was very small.

E. Nominum Vantio and Nominum ANS

Nominum Vantio and Nominum ANS are commercial DNS cache servers and authoritative DNS servers developed by Nominum. Figure 10 and 11 show the version distribution of Nominum Vantio and Nominum ANS, but these source codes are not published, so the version information is not disclosed. However, from our measurements, we can predict that the latest versions of Nominum Vantio and Nominum ANS were 5.3.3.1 and 5.3.1.0, respectively.

Table II reveals that there were nearly one million open resolvers by Nominum Vantio. However, Nominum Vantio and Nominum ANS are very expensive and it is unlikely that so many open resolvers were Nominum Vantio. Dnsmasq and other DNS forwarders can be configured to forward VERSION.BIND queries as they are, and thus, some servers of Nominum Vantio (or other DNS servers) should be DNS forwarders such as Dnsmasq.

F. Others

The rows of “can’t detect” and “no version info” in Table II show the number of server addresses that responded indistinguishable response or did not respond to VERSION.BIND. Generally, as a security measure, it may be recommended to keep the software type and version information private, but, as can be seen from Table II, most servers that do not respond correctly to VERSION.BIND were open resolvers. Thus, it can be seen that the server that keeps the version information private did not imply that it is secure. Domestic routers, home terminal terminals, etc. have DNS server forwarders that perform resolving services, and these are implemented by their own implementations or Dnsmasq. Such DNS servers of these devices may return special version information ,such as ISP name, or may not respond to version information queries. For example, it has been found from our investigation results that there are many DNS servers that return ISP names in Italy as version information [1].

In this way, many open resolvers are also found in the row of “can’t detect” and “no version info” were considered to be caused by household routers and home termination terminals which were not properly set up.

VI. Domain Distribution of DNS Open Resolvers

We performed reverse lookup obtained 30 million IPv4 addresses obtained to examine domain names.

At first, we implemented software for reverse lookup using locally installed Unbound, but we gave up this method because it was estimated that it takes more than 2 months by the end of every reverse lookup. Therefore, we implemented the Reverse Lookuper in C++ that can perform reverse lookup asynchronously using libevent. Consequently, the reverse lookup of about 30 million addresses was finished in a short period of about 5 days with only one PC.

Figure 12 shows the domain name distribution of the open resolvers up to the third-level domain, and Figure 13 shows the third level domain name distribution of the open resolvers with JP TLD.

Figure 12 reveals that many open resolvers exist in domains from which spam mails are sent. For example, 163data.com.cn and hinet.net are the bulk sources of spam mails.

[1]Because of this, as shown in Figure 2, it is thought that the area around Italy is red.

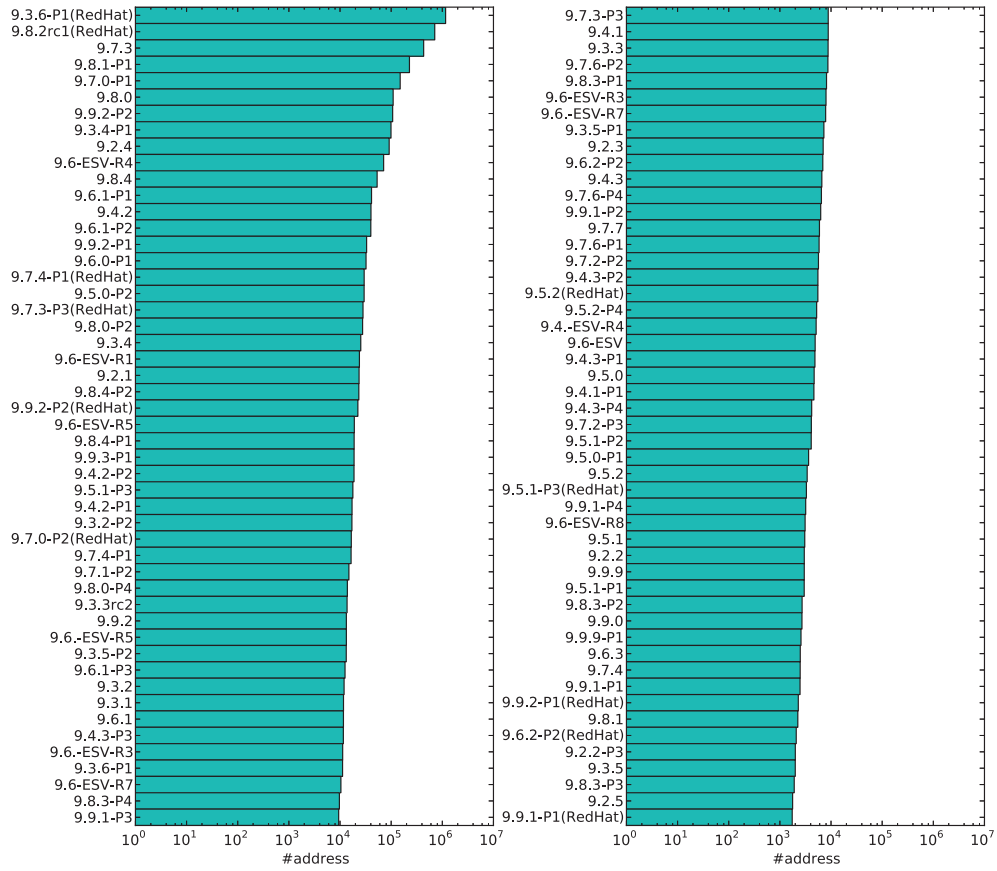


Fig. 3. Version Distribution of BIND 9.x Series (Top 100)

The open resolvers of JP TLD, according to our measurements, was obtained a total of 381,387 addresses. Figure 13 reveals that there were the most open resolvers in JP TLD on OCN (ocn.ne.jp) by NTT Communications since it is the biggest Internet service provider (ISP) in Japan. From this fact, it is understood that it is difficult to control with open resolvers by ISP because anyone can operate freely.

VII. Observation of Scan Packets with Silent Monitor

In order to perform DNS amp attack, it is necessary to use open resolvers, and in order to use open resolvers, it is necessary to find open resolvers distributed in worldwide. To find the open resolver, as we did, there is a method of scanning the IPv4 address space, and the scan packets can be observed by monitoring the port 53, which is used by DNS servers. We installed a silent monitor that monitors port 53, observing scan packets for DNS servers from September 28, 2013 to January 6, 2014 for about 3 months.

Table III shows the query types of the DNS packets found in the observation for about 3 months, and most DNS queries were A record queries. Table IV is a statistic of the query names of the scan packets. We totally observed 12,255 scan packets, and further observed 7,099 domain names out of them. As a result, it revealed that various domain names were used for inquiries.

We observed that the scan packets were being constantly transmitted from the same network. Particularly remarkable was the scan packets from dnsscan.shadowserver.org, openresolverproject.org, and AS 29073. These scan packets were observed at least every few days or almost every day. The project of dnsscan.shadowserver.org was founded by the

TABLE III
Query Types of Probe Packet

type	#
A	11,972
ANY	268
PTR	7
NS	5
TXT	1

TABLE IV
Query Names of Probe Packet

query name	#
www.ujiaoban.com.	1,145
vip3.gfdns.net.	666
dnsscan.shadowserver.org.	593
www.iana.org.	143
pay.13hp.com.	84
.	72
ghmn.ru.	48
loo1.ru.	29
isc.org.	28
fkfkfkfa.com.	21

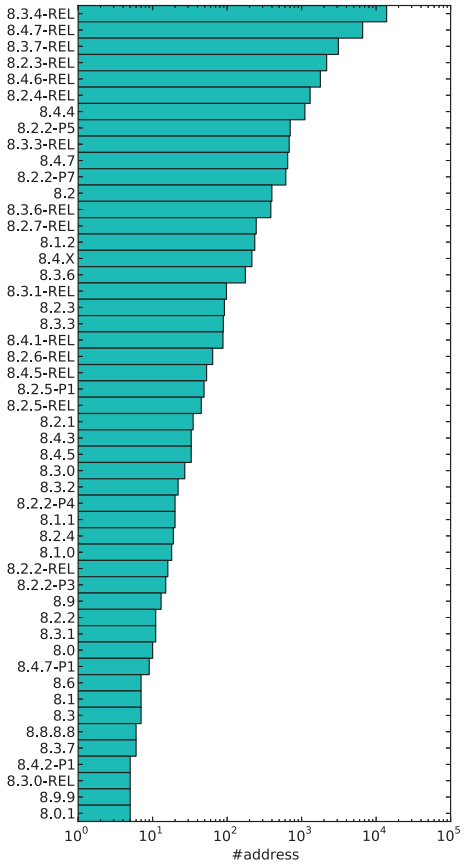


Fig. 4. Version Distribution of BIND 8.x Series (Top 50)

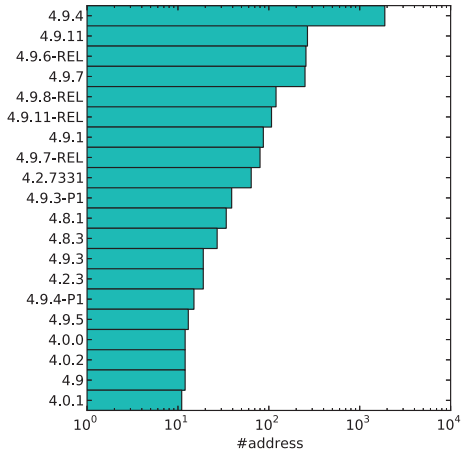


Fig. 5. Version Distribution of BIND 4.x Series (Top 20)

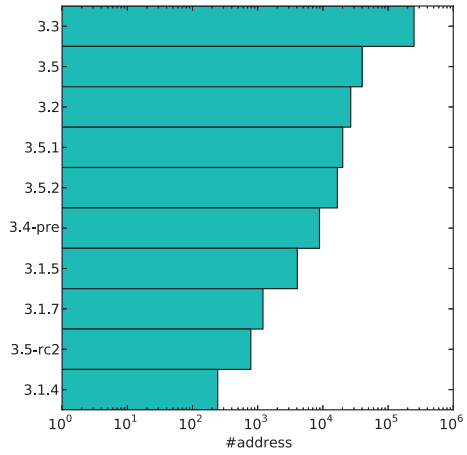


Fig. 6. Version Distribution of PowerDNS (Top 10)

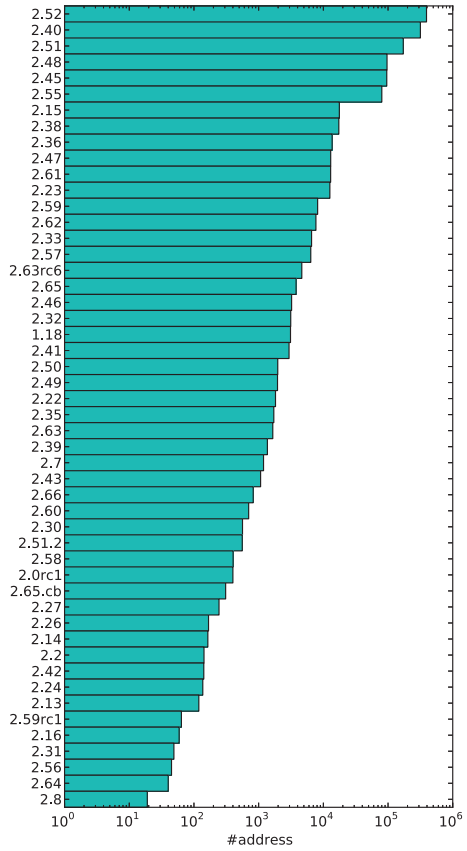


Fig. 7. Version Distribution of Dnsmasq (Top 50)

observation. Besides that, we observed DNS scan packets from PlanetLab [32], which is an open distributed network verification platform, and from some hosting servers.

VIII. Observation of Reflection Attacks with Actual DNS Open Resolvers

In order to know exactly how open resolvers are used as reflectors, we prepared multiple DNS open resolvers from September 28, 2013 to January 6, 2014 for about 3 months. When observing, we prepared two open resolvers (two IP addresses); one responded to only inquiries from AS 29073 described in Section VII, but another one no restriction. The IP address ranges of AS 29073 were determied from a route advertisement of border gateway protocol (BGP), and set it iptables OUTPUT filter to drop the obtained IP address ranges. The DNS server software used for the observation was Unbound. As a result, there was a big difference in how to use it between the former and the latter open resolvers.

Table V shows the types of DNS queries that came to

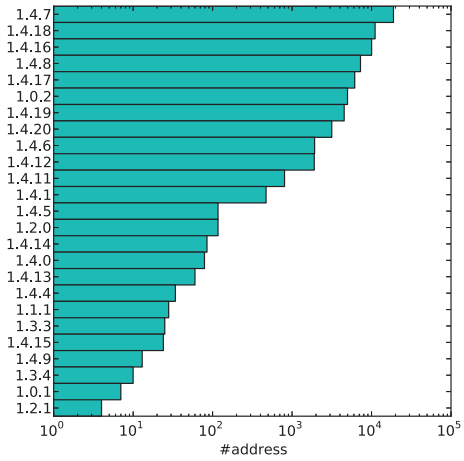


Fig. 8. Version Distribution of Unbound (Top 25)

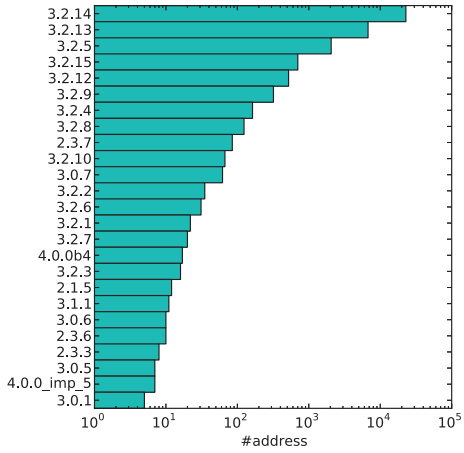


Fig. 9. Version Distribution of NSD (Top 25)

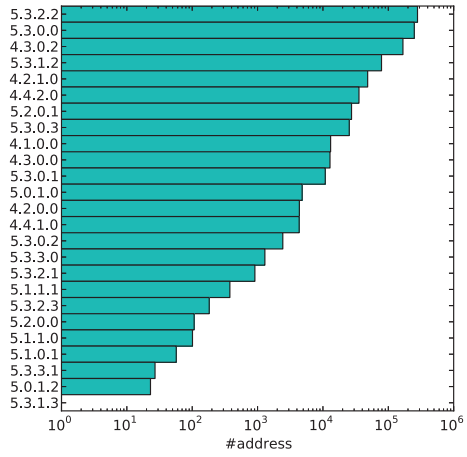


Fig. 10. Version Distribution of Nominum Vantio (All)

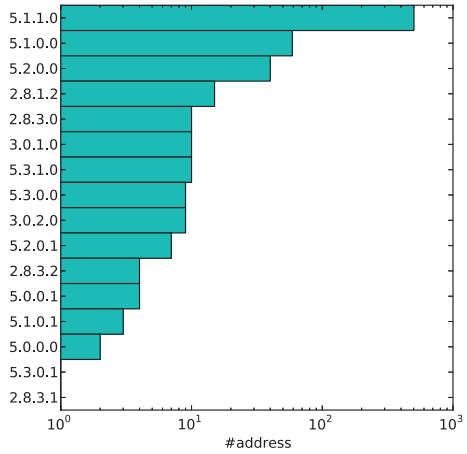


Fig. 11. Version Distribution of Nominum ANS (All)

the installed open resolver. Since no difference was found between the two open resolovers, data of representative one data is shown in the table. Table V reveals ANY queries were used very often. It implies that the open resolvers we prepared has a high possibility of being used as reflectors for the DNS amplification attack because ANY query is very often used for the DNS amplification attack. The next most common was A queries. In the case of a domain name to which a large number of IP addresses are assigned, this also increases the amplification factor, so it is also used for the DNS amplification attacks. Table V also reveals that the possibility that our open resolvers were abused for attack by A queries is high.

Next, let's look at the breakdown of ANY and A queries. Table VI and VII show the breakdown of the ANY and A queries, These tables reveals that the number of the ANY query of pkts.asia. and the A query of reanimator.in. were very small in the openresolvers rejecting inquiries from AS 29073, but these URLs were appeared in the open resolver with no restriction.

Figure 14 and 15 show the time transitions of the ANY and A queries that came to the prepared open resolver. Figure 14 reveals that the possibility that the DNS amplification attack using ANY queries was being steadily performed is extremely high. On the other hand, from Figure 15, the DNS amplification attack using A queries was ways of being used intensively in terms of time. Furthermore, similarly, these figures reveal that when AS 29073 was restricted, it was less used as a reflector.

TABLE V
QUERY TYPES TO OPEN RESOLVER

type	#	# (drop AS 29073)
ANY	33,564,934	14,359,798
A	2,910,108	727,044
TXT	38,292	32,618
RRSIG	4,719	0
MX	1	0
SOA	1	0
SRV	1	1
TYPE0	0	78,088

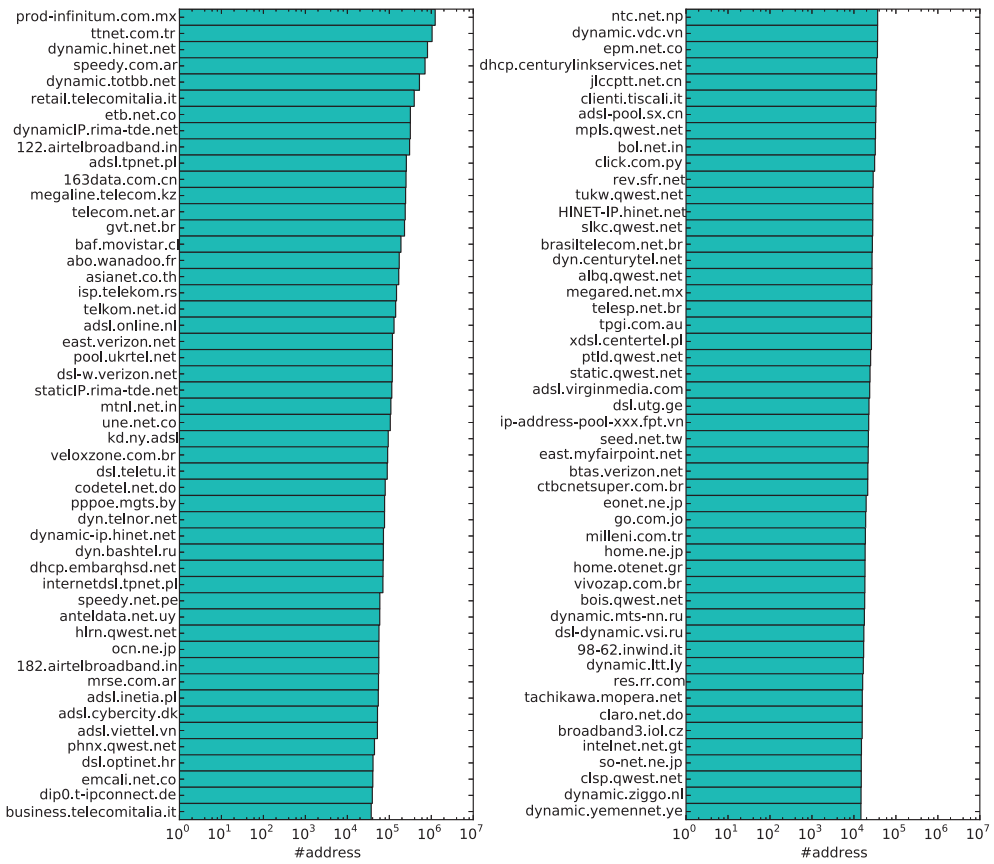


Fig. 12. Domain Distribution of DNS Open Resolvers (Top 100)

TABLE VI
ANY QUERY STATISTICS

query	#queries	#queries (drop AS 29073)
pkts.asia.	10,971,788	331
isc.org.	10,926,653	9,369,123
.	2,758,851	2,668,052
krasti.us.	1,974,023	3
fkfkfkfa.com.	1,701,773	20
ym.rctrhash.com.	916,113	1,346,231
ghmn.ru.	689,708	3
x.slnm.info.	650,039	1
lrc-pipec.com.	515,806	21,557
eschenemnogo.com.	452,167	444,744

TABLE VII
A QUERY STATISTICS

query	#queries	#queries (drop AS 29073)
reanimator.in.	873,583	11
ilineage2.ru.	863,495	6
eschenemnogo.com.	711,605	711,703
txt.fwserver.com.ua.	219,073	2
lrc-pipec.com.	210,354	9
ghmn.ru.	18,393	14,894
1x1.cz.	6,798	1
doc.gov.	5,963	0
aa.10781.info.	191	178
dnsscan.shadowserver.org.	101	91

IX. DISCUSSION

In this section, we discuss attack packets used for the DNS amplification attack, countermeasures, and the results of the open resolvers operation.

A. Attack Packets and DNSSEC

The reason why the DNS amplification attack is established is that DNS servers transmits response packets amplified by dozens of times as compared with the inquiry packets for specific DNS querys. For example, as of August 2013, the ANY query inquiry packets for isc.org and ripe.net were 64 and 65 bytes, respectively, but the responce packets were 3,245 and 2,669 bytes, respectively. This fact can be confirmed by using the dig command as `$ dig any isc.org + bufsize = 4096`.

Table VIII shows the breakdown of the response packets for ANY queries against isc.org and ripe.net. From this table, it can be seen that the RRSIG, DNSKEY, NSEC records related to DNSSEC [33], [34], [35] occupies the majority of responses to ANY queries. This suggests that an increase in response packets by DNSSEC triggers the DNS amplification attack. DNSSEC is a technology that guarantees that it is a valid DNS server response to Internet users. For example, in a China's censorship system called great firewall, anonymous authors reported that they have been performing DNS packet spoofing

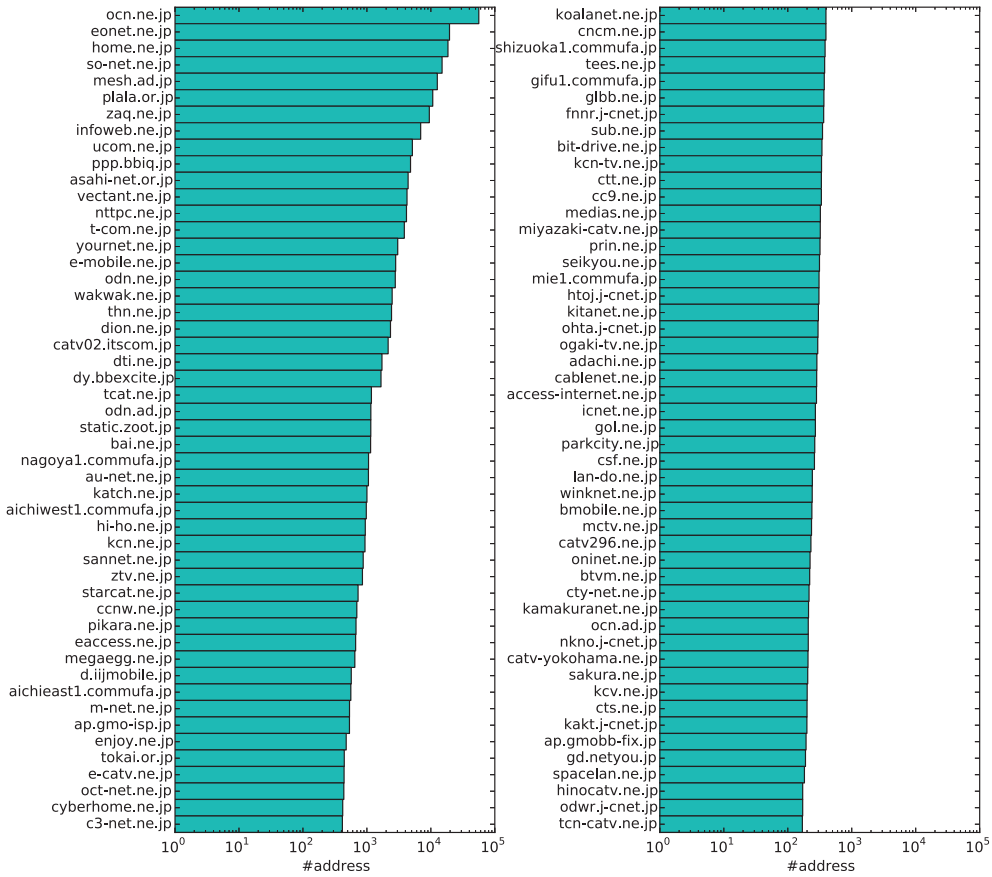


Fig. 13. Domain Distribution of DNS Open Resolvers in JP TLD (Top 100)

attacks at AS level [3], but, using DNSSEC, such spoofing attacks can be prevented. However, the extension of the DNS protocol to support DNSSEC [8] also has a negative aspect that enabled DDoS attacks.

TABLE VIII
BREAKDOWN OF RESPONCE OF ANY QUERY

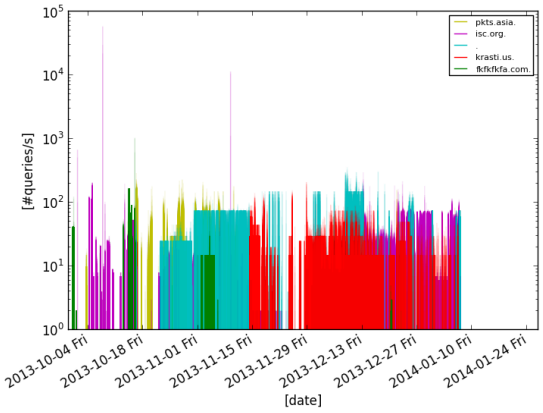
	isc.org	ripe.net
RRSIG	1,965	1,304
DNSKEY	427	848
NSEC	53	38
SPF	112	-
TXT	181	-
NS	97	136
NAPTR	46	-
A	16	16
AAAA	28	28
MX	24	50
SOA	54	52
Total	3,005	2,472

(bytes)

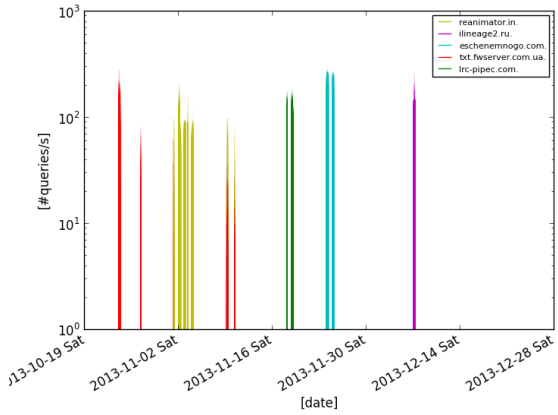
There are other DNS packets actually used for attack, other than DNSSEC being a factor. For example, pkts.asia or fkfkfkfa.com shown in Table VI had a large number of A records assigned to a single domain, and the response of ANY query increased becaused of the A records. Also, in the case of krasti.us, massively assigned MX records and TXT records were the cause of amplification. As can be seen from these facts, simply abandoning the ANY query can not fundamentally prevent the DNS amplification attack. In fact, we can see that it is possible to attack using RRSIG or A query, etc., based on the operation results of our open resolvers. Therefore, in order to take countermeasures against the DNS amplification attack, a more radical solution should be considered.

B. Countermeasure

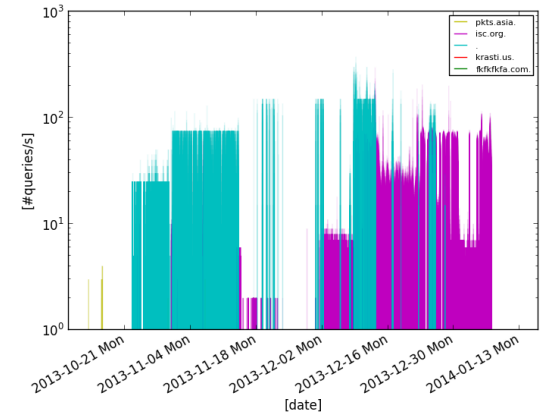
Improving and updating the DNS protocol itself is the most fundamental solution to the open resolver problem. In order to prevent the DNS amplification attack performed by misrepresenting the source address of UDP packets, it is necessary to incorporate the function of checking the validity of the sender in the DNS protocol. This can be achieved by using TCP instead of UDP. Since TCP is a connection-oriented protocol, at the beginning of communication, a connection is established by a three-way handshake. Therefore, unlike UDP, it is not easy to disguise the source and communicate. However, using TCP increases the response time of DNS (1 RTT for UDP, at least 2 RTT for TCP). This is a problem for DNS that needs to shorten the response time as much as possible. However, this problem can be alleviated by using a method that realizes high-speed TCP connection establishment, such as TCP Fast Open [36] or ASAP [37], rather than using pure TCP.



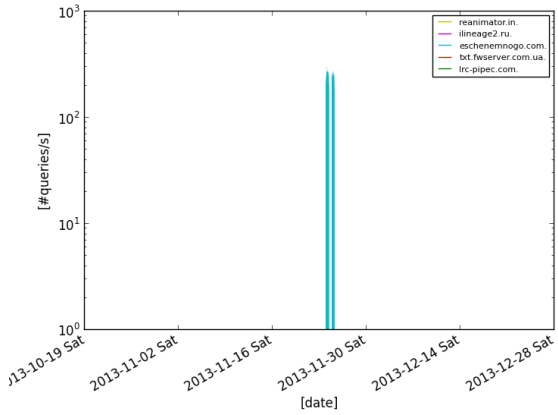
(No Restriction)



(No Restricion)



(Drop AS 29073)



(Drop AS 29073)

Fig. 14. Time Transision of ANY Queries

Fig. 15. Time Transision of A Queries

It is another fundamental solution to stop or correctly set up all open resolvers that are on the Internet and have nearly 25 million units. However, this method is expected to be difficult due to the nature of the Internet that is managed in an autonomous decentralized manner. Even if we can eradicate all open resolvers in Japan, it is only part of the whole, so its effect is light. In addition, as mentioned in Section VI and V, the fact that there are many open resolvers in the network from which spam mails originate and BIND 4, and the old implementation of 4.x and 8.x BIND servers are still operational, it is expected to be difficult to control all DNS servers.

Another solution is to make each ISP set an appropriate egress filter so that packets that misrepresented the source address can not be sent over the Internet. This method is effective not only for DNS amp attack, but also for other attacks that spoofed the sender address, and furthermore it is possible to take measures at the core network level of ISP. However, setting appropriate egress filter requires a reasonable cost. uRPF [38] is a technique applied to filtering the routing table, collates the source address with the routing table, and discards the packet if the source address does not exist as a route. Application of uRPF is highly effective against source address spoofing and is considered to be relatively realistic. However, there still remains the problem that it can not be forced to all ISPs.

C. Identify the Attack Origin

In Section VIII, it revealed that there was a big difference in how the open resolvers were used by blocking communication to the sender network of the DNS server scan packets observed by the silent monitor. There are three hypotheses to be considered for this cause as described below.

(1) It happened by chance. The way the usage of the open resolver differs greatly is the hypothesis that there is no causal relation with the blocking of network communication, and that our open resolvers happened to be differently used by chance. In fact, we doubted that such a result happened accidentally, so we also operated and observed open resolvers for about a week in several different places, and the open resolver rejecting against AS 29073 was used less often. Therefore, the probability that this hypothesis is the cause is considered to be low.

(2) It happened by someone sent a malicious packet to our prepared servers for the purpose of damaging the data acquired by our open resolvers. However, this is also considered to

be unlikely due to the short-term open resolvers operation as discussed above.

(3) We do not know whether the owner of the network of AS 29073 or the network of AS 29073 was abused, but attackers scanned DNS servers using networks of AS 29073 and performed DDoS attacks. From the results of Section VII and VIII, this hypothesis is considered to be influential.

If the third hypothesis is correct, it is possible to specify the attack origins by using a combination of data obtained from silent monitors, open resolvers, BGP, and it is possible to defense against DNS amplifier attacks.

D. Attack Targets

DNS amp attack is performed by misrepresenting the source address of UDP as an attack target, so if you examine the source address of the inquiry packets that came to the open resolver, you can know where the attack is for. Publishing the details of the attack targets obtained in our investigation may be disadvantages of the attack targets, so we will only briefly discuss in this section.

In the Section VIII, we showed that ANY queries with large amplification rates and A queries for specific domain names are used quite often as inquiries to open resolvers. Examining the sender addresses of these inquiries, it was found that there were many attacks against data centers and cloud or hosting services. This is probably due to the fact that many Internet services are operated on data centers and cloud services, as indicated by the name of denial of service attacks.

X. Conclusion

The DNS amplification attack is a DDoS attack using a resolving DNS servers that are open to the public as reflectors, called open resolvers, and it have become a big problem. In this research, we conducted a survey on the open resolver from a multifaceted viewpoint of active scan, silent monitor observation, actual open resolvers operation, and reported the details of open resolvers in this paper.

According to our active scan, we found that more than 30 million DNS servers existed on the Internet, of which about 25 million were open resolvers. We also conducted version survey of the DNS server. As a result, it revealed that very old software implementations such as BIND 8.x and 4.x series still existed, and Dnsmasq, a DNS forwarder for small networks, and PowerDNS were open resolvers. In addition, it alos revealed that a part of Unbound and BIND 9.x series were also open resolvers, although the ratio was small compared to others. Moreover, as a result of reverse lookup of the IP addresses of the open resolver obtained by the investigation, it revealed that many open resolvers existed in the domains of spam mail origins.

Furthermore, we show that DNS server scan packets observation can be done by a silent monitor, and as a result of the observation, we found scan packets from AS 29073, dnsscan.shadowserver.org, and openresolverproject.org. We also found that the A record query was often used for scanning. Next, we actually operated open resolvers and observed attack packets. As a result, it revealed that many ANY queries with high amplification rate and specific A queries were used for DDoS attacks. In addition, as a result of preparing and comparing two types of open resolver which discards inquiries from AS 29073 and open resolver which was not restricted anything, it revealed that the way of use differs greatly, and the network of AS 29073 was abused by the DNS amplification attack.

Acknowledgement

We thank Prof. Shinoda of the Japan Advanced Institute of Science and Technology, who supported this research. I also thank Kunio Akashi of the Japan Advanced Institute of Science and Technology, who had developed the network and PC environment used in this research.

References

[1] P.V. Mockapetris. Domain names - concepts and facilities. RFC 1034 (INTERNET STANDARD), November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 2065, 2181, 2308, 2535, 4033, 4034, 4035, 4343, 4035, 4592, 5936.

[2] P.V. Mockapetris. Domain names - implementation and specification. RFC 1035 (INTERNET STANDARD), November 1987. Updated by RFCs 1101, 1183, 1348, 1876, 1982, 1995, 1996, 2065, 2136, 2181, 2137, 2308, 2535, 2673, 2845, 3425, 3658, 4033, 4034, 4035, 4343, 5936, 5966, 6604.

[3] Anonymous. The Collateral Damage of Internet Censorship by DNS Injection. *SIGCOMM Computer Communication Review*, 42(3):21–27, 2012. http://conferences.sigcomm.org/sigcomm/2012/paper/ccr-paper266.pdf.

[4] M. Borella, D. Grabelsky, J. Lo, and K. Taniguchi. Realm Specific IP: Protocol Specification. RFC 3103 (Experimental), October 2001.

[5] Understanding Kaminsky's DNS Bug - Linux Journal. http://www.linuxjournal.com/content/understanding-kaminskys-dns-bug.

[6] US-CERT Alert(TA13-088A) DNS Amplification Attacks. http://www.us-cert.gov/ncas/alerts/TA13-088A.

[7] CloudFlare - The DDoS That Knocked Spamhaus Offline (And How We Mitigated It). http://blog.cloudflare.com/the-ddos-that-knocked-spamhaus-offline-and-ho.

[8] J. Damas, M. Graff, and P. Vixie. Extension Mechanisms for DNS (EDNS(0)). RFC 6891 (INTERNET STANDARD), April 2013.

[9] M. Handley, E. Rescorla, and IAB. Internet Denial-of-Service Considerations. RFC 4732 (Informational), December 2006.

[10] Open Resolver Project. http://openresolverproject.org/.

[11] The Shadowserver Foundation: DNS Scanning Project. https://dnsscan.shadowserver.org/.

[12] Steve Santorelli. The global open resolver picture. http://www.securityacts.com/securityacts03.pdf#page=29.

[13] Jeffrey Dean and Sanjay Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. In *OSDI*, pages 137–150. USENIX Association, 2004.

[14] MongoDB. http://www.mongodb.org/.

[15] Boost C++ Library. http://www.boost.org/.

[16] libevent. http://libevent.org/.

[17] Catenaccio DPI. https://github.com/ytakano/catenaccio_dpi.

[18] IANA IPv4 Address Space Registry. http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.txt.

[19] Internet Systems Consortium — BIND. http://www.isc.org/downloads/bind/.

[20] Authoritative DNS — Nominum. http://www.nominum.com/products/core-engines/authoritative-dns/.

[21] Welcome to PowerDNS. https://www.powerdns.com/.

[22] nlnetlabs.nl :: Name Server Daemon (NSD) ::. http://www.nlnetlabs.nl/projects/nsd/.

[23] Dnsmasq - a DNS forwarder for NAT firewalls. http://www.thekelleys.org.uk/dnsmasq/doc.html.

[24] Vantio Caching DNS — Nominum. http://www.nominum.com/products/core-engines/caching-dns/.

[25] Unbound. http://unbound.net/.

[26] GeoIP Products ¡¡ Maxmind Developer Site. http://dev.maxmind.com/geoip/.

[27] Internet Systems Consortium — BIND Software Status. http://www.isc.org/downloads/software-support-policy/bind-software-status/.

[28] BIND8 entering end of life. https://lists.isc.org/pipermail/bind-announce/2007-August/000222.html.
[29] Ruo Ando, Yuuki Takano, and Satoshi Uda. Unraveling large scale geographical distribution of vulnerable DNS servers using asynchronous I/O mechanism, 2013.
[30] PowerDNS Authoritative Server 3.0 Release Notes. http://doc.powerdns.com/html/changelog.html#changelog-auth-3-0.
[31] nlnetlabs.nl :: Home ::. http://nlnetlabs.nl/.
[32] PlanetLab — An open platform for developing, deploying, and accessing planetary-scale services. https://www.planet-lab.org/.
[33] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. DNS Security Introduction and Requirements. RFC 4033 (Proposed Standard), March 2005. Updated by RFCs 6014, 6840.
[34] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Resource Records for the DNS Security Extensions. RFC 4034 (Proposed Standard), March 2005. Updated by RFCs 4470, 6014, 6840, 6944.
[35] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. Protocol Modifications for the DNS Security Extensions. RFC 4035 (Proposed Standard), March 2005. Updated by RFCs 4470, 6014, 6840.
[36] TCP Fast Open (IETF Draft). https://tools.ietf.org/html/draft-ietf-tcpm-fastopen-04.
[37] Wenxuan Zhou, Qingxi Li, Matthew Caesar, and Brighten Godfrey. ASAP: a low-latency transport layer. In Kenjiro Cho and Mark Crovella, editors, *CoNEXT*, page 20. ACM, 2011.
[38] F. Baker and P. Savola. Ingress Filtering for Multihomed Networks. RFC 3704 (Best Current Practice), March 2004.