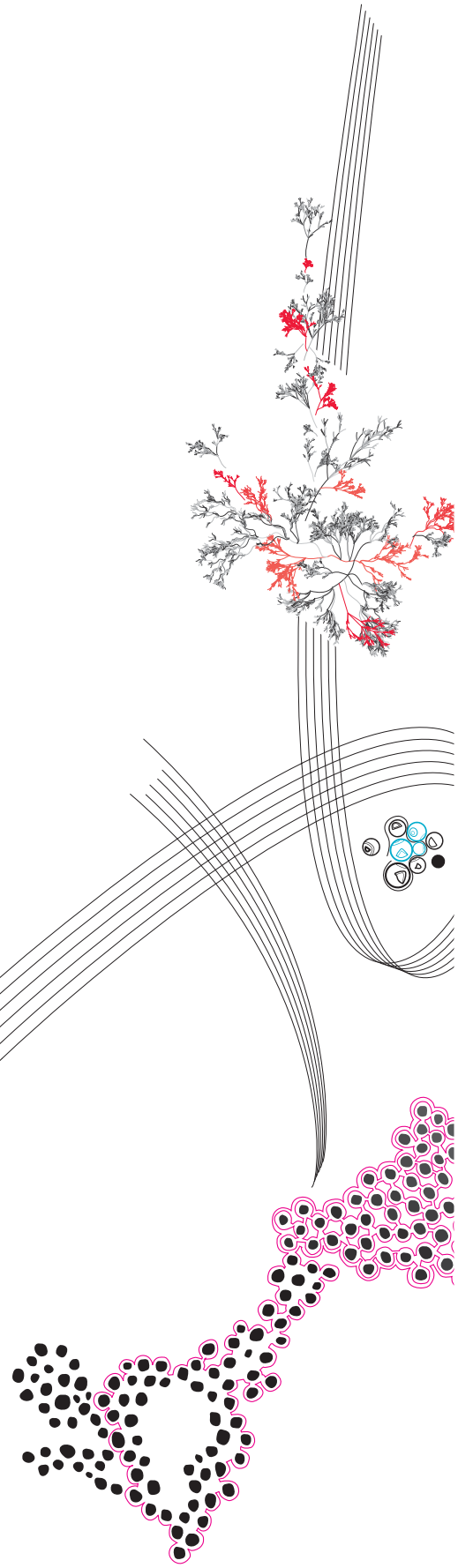

UNIVERSITY OF TWENTE

MASTER THESIS

Resilience of DNS Infrastructure Against DDoS Attacks

Arnab Chattopadhyay

March 26, 2022

Committee:	prof.dr.ir. R.M van Rijswijk - Deij dr.ir. Mattijs Jonker dr. A. Abhishta
Faculty:	Electrical Engineering, Mathematics and Computer Science (EEMCS)
Chair:	Design and Analysis of Communication Systems (DACS)

UNIVERSITY OF TWENTE.

Resilience of DNS infrastructure against DDoS attacks

Arnab Chattopadhyay
a.chattopadhyay@student.utwente.nl

Faculty of Electrical Engineering, Mathematics and Computer Science

University of Twente

March, 2022

Abstract

The Domain Name System (DNS) performs the crucial task of mapping human-readable domain names to machine-readable IP addresses. Disruption in this service, can lead to widespread latency or failure in accessing websites, web applications, email services, etc. Denial of Service (DoS) attacks on the DNS infrastructure disrupt operations on victim name servers either by choking its network or exhausting its computational resources. Prior studies have investigated well-known DoS attack incidents, including the defences within DNS for mitigation, either via simulated experiments or measurements on a small subset of the DNS landscape. However, a joint study of the effects of DoS attacks observed in the wild, together with the effectiveness of deployed countermeasures in the global DNS infrastructure is still missing.

This paper presents an overview on the operations of the global DNS infrastructure and efficacy of its defenses in the face of DoS attacks detected in the wild. To achieve this, we fuse data from two independent large-scale measurement systems, one inferring attack activity from a large network telescope and another consisting of active DNS measurements. Our data spans a period of one year, thus providing large-scale observations of DoS attacks, allowing us to gauge at the visible degradation of DNS resolution times consequent to attacks, as well as the effectiveness of resilience techniques deployed in DNS.

1 Introduction

The Domain Name System (DNS) is one of the most critical services of the internet. It is a globally distributed infrastructure that provides mappings for human-readable domain names like "example.com" to machine-readable addresses such as "93.184.216.34" as its core functionality. Authoritative Name Servers of a domain, store several such mappings in the form of Resource Records (RRs) [1]. Every application, web browser, email service uses a combination of these RRs to function and communicate across networked devices. Therefore, disruption in DNS operations can have wide-ranging effects from latency to complete loss of functionality for web-based applications. Attacks that aim to cause such disruptions in the DNS infrastructure, overwhelm authoritative name servers by sending a large volume of traffic, such as the case of DNS flooding, or exhausting its resources, such as the case of Phantom Domain and NXDomain attacks [2]. Attackers send rogue DNS requests either through a software program that generates packets with spoofed source IP addresses [3], or by using a botnet to launch an attack from several geographically spread zombie machines [4]. In both these circumstances, a Denial of Service (DoS) is perpetrated on the target name server, thereby rendering it incapable of serving legitimate DNS requests. Attacks on popular DNS service providers such as those on DYN in 2016, OVH in 2016, and AWS in 2019 revealed the prevalent threat of such DoS attacks on the DNS infrastructure [5, 6, 7].

There are various methods by which resilience can be achieved in the DNS infrastructure against DoS attacks. Anycast is one such mechanism that has been widely adopted as a DoS countermeasure [8, 9]. An Anycast network consists of several nodes that share the same IPv4 address. When traffic is destined to this IPv4 address, anycast distributes the traffic based on a prioritization methodology, usually to the node closest to the traffic originator in order to reduce latency. In the context of DoS attacks, Anycast can absorb the attack traffic within one node, thereby protecting other Anycast nodes. It can also withdraw routes to other nodes, thereby distributing both legitimate & attack traffic, and reducing the severity of impact. Moura et al. investigated the root server attack[10] and the DYN attack [6] to reveal that the underlying infrastructure (responsible for name resolution), and the use of resilience techniques such as anycast and longer TTL (Time To Live) values for RRs, attributed to varied resolution time of domains in each attack instance [11, 9]. Experiments on smaller portions of the DNS infrastructure have also established name server redundancies [9], Autonomous System (AS) level interconnections [12], and the use of Anycast [8] as resilience techniques against DoS attacks.

A majority of prior studies investigated either the effects of DoS attacks [13, 14, 15, 16] or the resilience techniques deployed to mitigate such attacks [17, 18], exclusive of one other. Inclusive studies have involved investigations either via simulated experiments or on a small subset

of the DNS landscape [19, 20]. Moreover, these studies have mainly focused on major DoS attack activities, because they were well-known attacks. Although these instances shed light on the larger problem of DoS attacks on the DNS infrastructure, there is little quantitative data about the prevalence of these attacks and their impact on the operations of the global DNS infrastructure.

We advocate that a long-term study of a significant part of the DNS infrastructure is necessary to understand both the visible degradation of DNS resolution times consequent to attacks and the efficacy of resilience techniques deployed. Our paper focuses on this investigation by answering the following main research question: How resilient is the DNS infrastructure against DoS attacks? We use two datasets for our analysis, one for inferring attack activity and one for performing active DNS measurements, both spanning a period of one year. Our main findings are:

- During the course of one year, we saw **564** attacks on various name servers, with a majority of them being low-intensity attacks, followed by about **19%** of them being high-intensity attacks. The impact on the degradation of resolution time for domains was the highest for high intensity attacks.
- A majority of the domains in the *.com* TLD, that experience a degradation in resolution time due to attacks, usually have only one or two authoritative name servers. For the case of high-intensity attacks, the severity of this degradation increased with an increase in the proportion of domains' name servers under attack.
- Using Anycast and distributing authoritative name servers across several ASNs[1] are prudent resilience techniques with the former reducing the degradation of resolution time of domains by 15-30% and the latter by 10-20% with each additional ASN.

The remainder of this paper is structured as follows. In Section 2, we discuss related work. Section 3 and Section 4 present the data sets used in our analysis and the methodology respectively. The results are discussed in Section 5. Limitations to our research are discussed in Section 6. Ethical considerations made during our research are described in Section 7, with conclusions and future work presented in Section 8.

2 Related Work

DoS activity inference has been the subject of several studies focusing on large-scale network measurements. Moore et al. [22] and Fachka et al. [23] measured attacks by analysing backscatter traffic from network telescopes. Krämer et al. [24] and Bailey et al. [25] used honeypots to monitor attack activity. The impact of such attacks on the DNS infrastructure has also been investigated in prior studies. Abhishta et al. [16] studied the effects DDoS attacks on customers of Managed DNS (MDNS) service providers, to find that customers usually rely heavily on the authoritative name servers of MDNS providers prior to attacks but switch to other servers in the aftermath. Moura et al. examined the effects of two DoS incidents - the root DNS servers [10] and DYN [6]. Investigation on both these instances revealed the use of anycast, redundancies, and longer TTL values as being the reasons for the varied effects. Other DDoS defense mechanisms have also been studied in the past. Jonker et al. [26] confirmed an increasing adoption of DPS (DDoS Protection Service Providers) in the top three TLDs, while in a separate paper, Jonker et al. [27] studied BGP blackholing and confirmed its effectiveness as a defense against DoS attacks. Anycast has also been widely adopted as a resilience technique against DDoS attacks. Sommese et al. [28], developed a methodology that uses anycast to detect prefixes that use anycast and furthermore showed the adoption of anycast in the major TLDs and SLDs [8].

3 Data Sets

In this section we describe the datasets that we use in our study. To study DoS attacks targeted towards the DNS infrastructure, we use attack-activity data from the UCSD Network Telescope [29] and active DNS measurements from OpenINTEL [30], over the period of 1 year. We examine these measurements to study the effects of such attacks on DNS, along with the resilience techniques used to mitigate these attacks. The measurement apparatus and the resulting datasets are described next.

3.1 Attack Measurement

We investigate a class of DoS attacks [31] - Randomly and Uniformly Spoofed Denial Of Service (RSDOS) attacks in which the attacker uses of a software program or a botnet to generate thousands of packets with spoofed source addresses and sends these packets to the victim. Under normal circumstances, the victim is not capable of differentiating a normal packet and an attack packet and hence sends a response packet in either case. Since this response packet is destined to an originally spoofed IP address, it becomes part of the Internet Background Radiation (IBR) or backscatter traffic. Network Telescopes monitor large address blocks and receive backscatter packets which not only consist of traffic from attack events but also traffic from the automated spread of Internet worms and viruses, scanning of address space by attackers or malware looking for vulnerable targets, and various misconfigurations (e.g., mistyping an IP address). As a result, backscatter observed on the network tele-

[1]Autonomous System Number (ASN) is a globally unique identifier that defines a group of IP prefixes under a single or sometimes multiple network administrations [21]

scope can be used to infer a conservative lower bound of denial of service activity as Moore et al. do in their seminal paper [22]. Figure 1, shows an example of attack activity inference using a network telescope. An attacking host Y generates 4 packets with randomly and uniformly spoofed source IP addresses and sends them to victim X. Replies from victim X constitute unsolicited traffic, one of which is observed by the network telescope, as it monitors $1/4^{th}$ of the IPv4 address space.

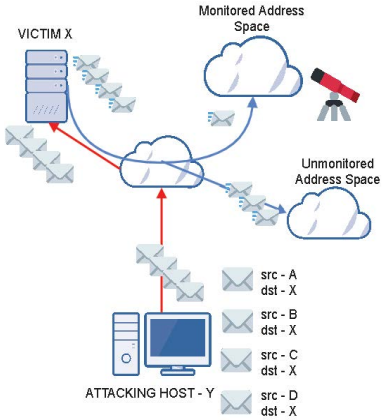

Figure 1: Backscatter traffic from an RSDOS attack. The network telescope monitors 1/4th of the address space (in this example), and receives 1 out of the 4 uniformly spoofed attack packets sent to the victim.

We use the UCSD network telescope for our study, which consists of two globally routed networks - a /9 and a /10 address block, covering approximately 1/341 of the IPv4 address space. The attack dataset is generated, by collecting raw telescope data of response packets sent by the victims of RSDOS attacks and then processing it in 5-minute intervals. We categorise a victim as being "under an attack" when the network telescope receives at least one flow (5-tuple[2]) of packets with proven thresholds of packet rate, packet count, and attack duration [22, 27]. Once the telescope observes an attack on an IP address, it collects various metrics during this 5-minute window, such as number of packets/bytes sent, the maximum packet rate in number of packets per minute (PPM), the number of unique attackers, the targeted ports, etc. These metrics represent an attack vector uniquely identified by the IP address under attack and the start-time of the attack. The observed metrics on the telescope provide us with a lower bound of attack activity on the victim. In the above example, the victim received 4 uniformly spoofed packets, however, we observed only 1 packet on the telescope due to its 1/4 address coverage. The metrics for each attack vector are updated once the telescope receives backscatter traffic from the target IP, therefore, the observed value of metrics on the telescope is always lower than the inferred/total metrics. Additionally, if an attack persists over multiple windows, all attack vectors are continuously updated across these 5-minute windows and the measurement stops, once the flow of packets ends. All metrics pertaining to the attack are then cumulated to generate the RSDOS dataset. A full list of all the metrics can be found on the official website of the network telescope.[3].

The measurements in our data cover a period of 1 year from August 2020-2021. We also use flow specific measurements for two months - November and December 2020. This gives us the characteristics of attack flows targeted towards DNS, which is representative of the remainder of the measurement period. These measurements contain exact values of target and attacker ports instead of cumulative statistics of the previous measurements. Analysing both these types of measurements gives us detailed record of DoS activity aimed at DNS.

3.2 Active DNS Measurements

Rijswijk et al. implemented a system for large-scale active DNS measurements [30]. This measurement platform, called OpenINTEL, measures domains for major Top Level Domains (TLDs) such as *.com,.net,.org*. It queries a set of RRs for these domains and repeats the measurements daily. The response of these queries along with their round trip times (RTT), are stored in a dedicated Hadoop cluster. In addition to the domain list of each zone, the measurement system also collects authoritative name server IP addresses explicitly from **NS** queries and their subsequent resolution of **A** queries. We use this measurement system to record the state of the DNS infrastructure and analyze its behaviour in response to DoS attacks. Once the data arrives in the Hadoop cluster, we leverage Spark's[4] large-scale data processing and distributed computing to perform data analysis in batches. Since the measurement snapshots have daily granularity, we use batch sizes of one day. We focus our investigations on the *.com* TLD because it is the largest zone covering a significant portion of the namespace, therefore our analysis is representative of the global DNS infrastructure. As future work, we plan to integrate our analysis with other TLDs as well, such as *.net* and *.org*.

3.3 Complementary Datasets

We also use two complementary datasets for our investigations. First, we use data from Rapid7[5], which contains monthly snapshots of responses to zmap probes performed against common TCP services in the entire

[2]A 5-tuple represents 5 different values that uniquely identify a TCP/IP connection. It includes the source IP address, source port, destination IP address, destination port and the protocol in use.
[3]https://www.caida.org/data/passive/telescope-daily-rsdos.xml
[4]https://spark.apache.org/
[5]https://opendata.rapid7.com/

IPv4 space. We combine this with our flow-specific measurements to determine if the target port was "open" on the probed IP during the attack. Second, we use the MAnycast^2 census to detect name servers that have anycast enabled [28]. MAnycast^2 probes IP addresses with ICMP echo-requests, from all of its anycast testbed nodes as source. It then labels the probed IP as being unicast, if the ICMP echo-reply is routed back to a single node on the testbed and, being anycast if the ICMP echo-reply is received on multiple nodes. Combining our anycast census data along with DNS measurements and attack-activity measurements, enables us to evaluate the efficacy of anycast as a resilience technique against DoS attacks.

4 Methodology

We use the datasets described in Section 3 to infer RSDOS attacks on the DNS infrastructure. To study the characteristics and effects of these attacks along with the resilience techniques deployed in DNS, we follow the steps described in the next sections.

4.1 Extracting DoS attacks on DNS infrastructure

Figure 2 depicts our analysis workflow of extracting DoS attacks aimed at the DNS infrastructure. For a given day, we uniquely identify every attack instance based on a target IP and the start time of an attack from the RSDOS dataset. If the start time of consecutive attacks on the same target IP overlap, we consider this as one large attack instance instead of several smaller ones. This effectively reduces the dataset from approximately 7.5M data points to 1.8M attack instances over the period of one year.

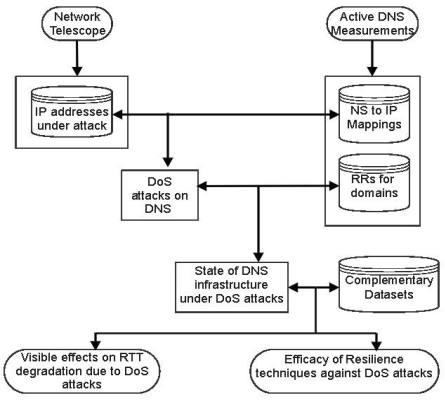

Figure 2: Analysis Workflow. Fusion of RSDOS dataset along with data from active DNS measurements, allows us to evaluate the state of DNS infrastructure consequent to DoS attacks

To identify name servers within these attack instances, we apply the following procedure. We start with NS to IP address mappings from OpenINTEL as described in Section 3.2. We track the occurrence of these name server IP addresses in the RSDOS dataset to yield around 500 attack instances on DNS infrastructure over the course of one year. We eliminate clear cases of misconfigurations among identified attack instances. Furthermore, certain attack instances on name servers continued for more than a day. In these situations, we take the attack into account, for the day when it ends. For example, if an attack on 1.2.3.4 starts at 10 PM on 23-05-2021 and ends at 5 AM on 24-05-2021, we include this attack instance in the analysis for 24-05-2021, thus eliminating duplicate calculations over multiple days. With our approach, we were able to identify both large-scale attacks - usually on well-known DNS service providers, as well as small-scale attacks targeted towards smaller name servers.

4.2 Defining Attack Intensity

Once attacks are identified, we focus on categorising them based on their intensity. We advocate, that higher intensity attacks on the name server infrastructure to cause higher impacts on name resolution for its registered domains. Since our dataset was unlabelled in terms of intensity, we chose the approach to cluster attack instances using K-Means. Further analysis revealed that packet-rate and number of distinct attacker IPs contribute 92-96% of variance in the dataset. Consequently, we use these metrics to generate the K-value to be 3 clusters. We then cluster attacks based on their intensity into three categories - High Intensity (HI), Moderate intensity (MI), and Low Intensity (LI) attacks.

We perform our clustering using the following procedure. On any given day, we report name servers under attack, along with all the metrics for the attack, by adopting the methodology described in Section 3.1. We then generate our K-Means model based on these metrics and run the model daily over the course of our measurement period. Table 1 presents the distribution of DoS attacks with varying intensity on the DNS infrastructure. As can be seen from the table, a majority of attacks are LI attacks, followed by HI and MI attacks. For HI attacks, 51 unique ASNs were targeted across 115 attack instances over the course of our measurement period of one year.

Intensity	# unique target IPs	# unique ASNs	# attack instances	# domain names
Low	245	120	432	53M
Moderate	16	11	17	4.4M
High	87	51	115	31M

Table 1: DoS attacks on the DNS infrastructure during our measurement period. A majority of DoS attacks on the DNS infrastructure are Low Intensity attacks

After characterizing the attacks on the DNS infrastructure, we now have the basis for evaluating the effects of

similar intensity attacks as well as the resilience techniques used to mitigate such attacks. The various steps involved in quantifying these effects and resilience techniques are described next.

4.3 Measuring Effects on Domains

We measure the impact on registered domains of name servers under attack, in terms of their degradation in name resolution times. We use the Round Trip Time (RTT) from our DNS measurements, specifically, we use the RTT of **SOA** records. We use the **SOA** record because, for a given domain name, **SOA** records are the first to be queried by our measurement system, followed by other records such as **A**, **AAA**, **NS**, etc. However, this choice is independent of our results. Investigations on the RTT for other records show a similar trend as we see in **SOA** records. Because our measurement system queries a given domain once daily, we compare the RTT on the day of the attack, with the RTT on the previous day (stable state) to infer the degradation of resolution time due to an attack. For a given domain, we have the impact factor/impact intensity as follows:

$$\textbf{Impact Factor} = \frac{\textbf{RTT on attack day}}{\textbf{RTT on previous day}}$$

A limitation to our research is that the DNS measurement times from OpenINTEL do not always overlap with the attack duration as seen on the telescope. We will discuss this further in Section 6. To account for this limitation, we had two approaches. In the first approach, for a given set of name servers under attack, we measured the resolution time of its registered domains, beginning from 5 minutes prior to the start of the attack till 5 minutes after the attack ends based on the timestamps we saw on the telescope. We used 5-minute buffers because the network telescope measures responses from victims in 5-minute windows. Increasing our buffer in the first approach to more than 5 minutes, would already have been recorded as an attack on the telescope. We then compared the recorded RTT with the measurements done for the same set of domains on a day prior to the attack. With the first approach, we were only able to measure a small set of the domains whose measurement times overlapped within the timestamps of the attack recorded in the RSDOS dataset. We got a noticeable degradation in resolution times for domains, but due to the small number of domains measured, we ran into the limitation of detecting possible intermittent network latency between our vantage point and the name server under attack.

In the second approach, we measured the resolution time of all the registered domains for a given set of name servers under attack for an entire day, rather than measuring them just within the time period that we saw on the network telescope. We then compared these measurements with a day prior to the attack, i.e., the stable state of the domains. We make the argument here that although this technique captured several false positives, for example - when the domain was measured before an attack began on its name server or after the attack ended, we would still be able to see an increase in the average impact factor across the entire day, albeit not as significant as the previous approach. Moreover, this approach allowed us to circumvent the limitations discussed above, and also allowed us to make comparisons across several days of attacks with varying intensities.

An impact factor greater than 1 signifies a considerable degradation of resolution time for domains. We ascertain this hypothesis, by the following method. We take the domains whose name servers were not under attack for a period of one month. We track the variation of impact factor for these domains over the course of the month. In this scenario, the impact factor is the resolution time of the domain on any given day, compared to its resolution time on the previous day. The results of this analysis show that, for any given day, the average impact factor for all the domains measured - varied between 0.95-1.02. This variation is explained by intermittent latency between our vantage points and the name servers between these domains, or by intermediate infrastructure changes for these domains. Therefore, the stable state for a domain is defined to be an impact factor of 1, and a value greater than 1 demonstrates a significant deterioration in its resolution.

For a given domain, we also track the percentage of its authoritative name servers under attack during our measurement period. Consider the following example - the domain example.com has three authoritative name servers - ns1.example.com (1.1.1.1), ns2.example.com (2.2.2.2) and ns3.example.com (3.3.3.3). By analyzing the RSDOS data, we see that only ns1.example.com (1.1.1.1) is under attack. In this case, the domain has two backup servers to fall back on to continue its day-to-day operations. Thereby, in this case, we say that the domain is 33% under attack. For a given domain, we have the attack percentage as follows:

$$\textbf{Attack(\%)} = \frac{\textbf{\# name servers under attack}}{\textbf{total \# name servers}} * 100$$

Defining the above two metrics helps us in two ways. First, it allows us to evaluate the domains that experience a degradation in resolution times as well as the severity of this degradation. We use the impact factor to determine this. Second, it also shows us which kind of domains (in terms of the number of authoritative name servers) fall prey to such attacks on name servers. For example, if a domain has 5 authoritative name servers, it should experience a significantly lesser impact than a domain with just 1 authoritative name server, in the event of an attack. Moreover, domains with 1 authoritative name server are more prone to even a small attack, while domains with 5 name servers have the infrastructure to withstand smaller attacks.

4.4 Effects on the same IP subnet

Further investigations on our RSDOS dataset revealed that some attack instances were accompanied by a very high number of attacker IPs and attack packets. With such a high volume of attack flows, we hypothesize, that the impact of an attack on the name server is visible on other services hosted on the same network.

To determine this, we adopt the following approach. We take the /24 prefixes both from the RSDOS dataset and the NS-IP mappings. The reasons for taking a /24 prefix are twofold. First, a more specific prefix will eventually be summarised during route advertisements to the upstream routers, so an effect discovered on an upstream /24 prefix will also make its way into the downstream specific prefixes. Choosing a more specific prefix for analysis, therefore will be inefficient as it will be very strict to the prefix and might obscure the impact of an attack on a major part of the upstream network. Second, choosing a more generic prefix for analysis such as a /21, or a /16 will lead to an overestimation of the impact of an attack. Organisational-based network segmentation usually occur at the /24 prefix level, with name servers and other services hosted in this prefix [32]. Therefore choosing a more generic prefix, will not guarantee that these services share the same network infrastructure.

Our goal is to determine if the resolution time degradation that is observed for registered domains of the attacked name server, is also visible on other services hosted on the same address block as that of the target name server. Therefore, we choose a /24 prefix for our analysis. Once we have the /24 prefixes, we track NS-IP mappings (prefix mappings now) in the RSDOS dataset to isolate the /24 prefixes that were under attack. We then track the name servers hosted in these /24 prefixes and evaluate the resolution time degradation of domains that use these name servers (other than the target name server). We argue that a resolution time degradation, implies a similar performance degradation for other services hosted in the same network. Therefore, based on the impact factor of these other domains, we determine, if the impact of attacks on name servers is visible on other services hosted in the same network.

4.5 Deriving Anycast Use

Anycast has been adopted in authoritative name servers across a majority of TLDs and SLDs as a defense against DoS attacks. Our goal is to quantitatively determine the efficacy of this resilience technique.

To this end, we use the anycast census - MAnycast[2] data repository to track the target IPs that were under attack and were also using anycast. We use our DNS measurement system, to evaluate the degradation of resolution times of the registered domains of attacked name servers. We compare the severity of RTT degradation for domains that use anycast name servers versus domains that use unicast name servers to determine the effectiveness (or lack thereof) of anycast, as a resilience technique against DoS attacks.

5 Results

In this section, we present the results of our research. We start by providing a general overview of DoS attacks on the DNS infrastructure, followed by a case study of two attack instances on a large European DNS service provider. We then discuss, based on our metrics, the impact on DNS operations consequent to DoS attacks detected in the wild. This is followed by an analysis of the efficacy of resilience techniques used in DNS, and we conclude the section with results from reactive measurements.

5.1 DoS attacks on the DNS infrastructure

We analyze DoS attacks targeted towards the DNS infrastructure over a period of 1 year from August 2020 to August 2021. We determine the intensity of attacks by following the steps outlined in Section 4.2.

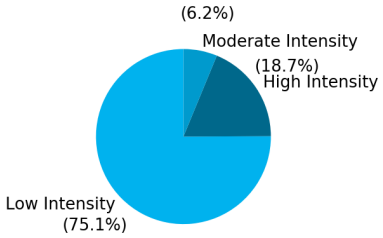

Figure 3: Attack Intensity distribution. Three quarter of attacks on DNS infrastructure were Low Intensity and 19% attacks being High Intensity.

Figure 3 shows the distribution of DoS attacks based on their intensity. During the course of the year, we saw **564** attacks on **143** organisations[6], with a majority of them being low-intensity attacks, followed by about **19%** of them being high-intensity attacks.

In Figure 4, we show the number of attacks targeted towards the name server infrastructure every month during our 1 year measurement period, as well as the number of their registered domains that were affected by these attacks. The figure is color-coded with darker shades depicting higher intensities of attacks. Referencing both the histograms in Figure 4, we found no correlation. For example, we saw more than **80** attacks on name servers

[6]Organisation here refers to a distinct ASN, whose name servers were under attack

in August 2020, whereas the total number of domains affected by these attacks was quite low compared to other months. Conversely, we also saw that about **16M** domains were affected in September 2020, whereas the number of attacks in that month was comparatively lower.

When a popular DNS service provider is the target of an attack, we see a significant increase in affected domains count, whereas multiple attacks on smaller service providers do not have a major contribution to the domain count. For example, Neustar ULTRADNS - A popular DNS service provider based in the US, was targeted in September 2020, which accounted for about 11M of the 16M registered domain names that we saw under attack, while reporting only 5 of the 67 attack instances in that month.

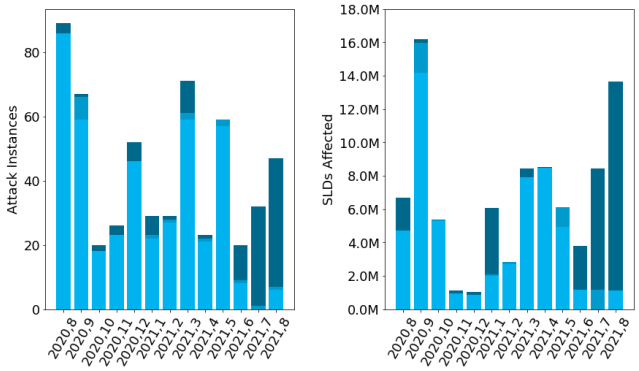

Figure 4: Attack instances and # Registered Domain Names affected - An overview of the attacks targeted towards the DNS infrastructure

Table 2, gives an average monthly account of attacks during our measurement period. The first column shows the average "metric" per month. The second, third and fourth columns represent these statistics in the context of Low, Moderate and High intensity attacks, respectively.

avg. / month	Low	Moderate	High
# of attacks	33	1	9
# of SLD's	4.1M	334k	2.2M
# of distinct ASNs	20	1	6
attack duration	26.2 Hr	9.1 min	9.4 Hr
# of attacker IPs	109K	432K	2.57M
packet rate$_{max}$	2.2K	15K	20K
traffic volume$_{max}$	143 Mbps	975 Mbps	1.3 Gbps
impact factor	1.15	0.71	1.77

Table 2: Characteristics of Attacks on the DNS infrastructure. Impact factor is the highest for High Intensity Attacks

The metric of packet rate here signifies the maximum backscatter packets per minute received on the network telescope for a given attack instance. As discussed in Section 3.1, this is only the observed backscatter traffic (in PPM) that is sent to the network telescope from the victim. The network telescope covers 1/341 IPv4 address space, i.e. it receives 1 backscatter packet for every 341 uniformly spoofed attack packets. We use this fact and assume 1500 byte-sized packets, to infer the attack traffic volume on the victim depicted by "traffic volume$_{max}$" in Table 2.

Another significant aspect of Table 2 is that for High Intensity (HI) attacks, impact is also the highest. Whereas the impact factor is lowest for Moderate Intensity (MI) attacks. Our measurement system can switch between name servers across several days of measurement, which can lead to false positives and an underestimation of resolution times for domains. This limitation is reduced if the number of registered domains measured for name servers under attack on a particular day is at least 1M. Therefore, it is particularly evident in the case of MI attacks, due to the significantly lower number of domains measured as compared to LI and HI attacks. For example, during January 2021, LI & HI attacks on name servers accounted for the measurement of 2M & 3.9M of its registered domains respectively, whereas the total number of domains measured by MI attacks on name servers was just 72K. We will discuss this limitation further in Section 6. Additionally, since the impact factor is calculated based on the average RTT increase on the attack day compared to the previous day, the visible degradation of resolution time shown in Table 2, provides a conservative lower bound of the impact on domains. We validate this further by performing a case study of two major attack instances on a large DNS service provider in the next section.

5.2 Case Study - DoS attack on a large DNS service provider

In this section, we discuss two instances of DoS attacks on a large European DNS service provider, henceforth referred to as "Victim X". We discuss the general characteristics of each of the two attack instances. Using this case study, we provide our reasoning for adopting the approach of impact factor calculations for the entire day rather than just within the duration of the attack. Finally, we conclude with a description of the impact of each of these attack instances.

In Table 3, we show the observed packet rate on the telescope (in PPM), the number of attacker IPs and the inferred attack traffic volume on "Victim X" during the attacks in December 2020 and March 2021. As discussed in the previous section, we consider the address space coverage of the network telescope and 1500 byte-sized packets to infer the attack traffic volume. "Victim X" has three name servers, with "A" being the target of a HI attack, while the other two name servers experienced low-intensity attacks. Furthermore, the instance on March 2021 was accompanied by HI attacks on all three name

servers. A comparison of the volume of attack traffic on both these instances revealed that the attack on March 2021, was far more severe, with traffic rate peaking at 8 Gbps and attack coming from 7M unique attacker IPs . We investigate and compare the severity of the impact of both these attack instances next.

Target Name Server		A	B	C
December 2020 Attack	Observed Packer Rate (PPM)	21.8K	3.8K	2.9K
	Inferred Traffic Volume	1.4 Gbps	247 Mbps	188 Mbps
	Attacker IP Count	5.79M	1.57M	1.33M
March 2021 Attack	Observed Packer Rate (PPM)	125K	123K	13K
	Inferred Traffic Volume	8 Gbps	7.8 Gbps	845 Mbps
	Attacker IP Count	7M	6.19M	823K

Table 3: A comparison of attack metrics for December 2020 attack vs March 2021 attack on Victim X. During December 2020, only name server A was targeted with an HI attack, whereas during March 2021, all the name servers were victims to HI attacks.

Tables 4 and 5 describe the effects of each of the name server attacks, on its registered domains for December 2020 and March 2021 respectively. The first column in both the tables shows the name server under attack. The second and third columns present the number of registered domain names measured and the impact factor respectively, when the measurement times of domains fell within the duration of the attack as observed on the network telescope. A similar analysis is showcased in the last two columns, but with measurements of domains from the entire day rather than just during the attack duration. For the December 2020 attack, we see the attack starting on the network telescope at 10 PM on November 30 and lasting till 12:30 AM on December 1[7]. A limitation of our research is that the measurements in OpenINTEL are not triggered by an attack, therefore, we do not have active DNS measurements on the 30^{th} November between 10 PM till 11:59 PM. We will discuss this limitation further in Section6. In this situation, we calculate the impact factor for the registered domains of name servers under attack between 12:00 AM and 12:30 AM on December 1 and assume the same effect on resolution time holds true for the entire duration of the attack. However, if we had performed reactive measurements, i.e., we start measuring domains right after an attack is observed on its name servers, we alleviate this assumption. This will be discussed further in Section 5.5.

In both tables, the slight difference in SLD count for the name servers is accounted for the case when some domains reference two of the three name servers of "Victim X". As a result, we see an impact factor variation when the measurements are done within the attack window. Comparing both the tables reveals that, the name servers of X were subject to higher intensity attacks on March 2021, than in December 2021. However, we notice a greater impact factor for December 2020 attack (day window) compared to the March 2021 attack. The attack in March was only for 20 minutes, while the attack in December was 2 hours, so the average RTT over the longer duration yields a greater value. Additionally, the degradation of resolution time for domains is far greater, when the window of measurement fell within the attack duration.

Target	Window - Attack		Window - Day	
	SLDs	Impact	SLDs	Impact
A	2720	25.43	143806	13.65
B	1150	14.57	144244	
C	1683	12.51	144253	

Table 4: December 2020 attack on Victim X, with impact factor comparisons for our two approaches. Impact factor is more significant in the first approach, but less number of domains were measured

Target	Window - Attack		Window - Day	
	SLDs	Impact	SLDs	Impact
A	11887	28.06	139746	3.27
B	14188	27.12	139224	
C	12299	12.51	139246	

Table 5: March 2021 attack on Victim X, with impact factor comparisons for our two approaches. Attack window - impact factor is generally larger in March 2021 than December 2020. Day window provides less visibility to the severity of impact, but nevertheless we see an impact that is not restricted by other limitations of our measurement [6].

This severity is also evident when the analysis window was the entire measurement day, although less visible than the previous approach. However, the number of domains measured in the first approach is far lesser than the latter approach. This makes the first approach prone to an overestimation of the effects of an attack due to possible intermittent network outages or latency caused in the intermediate path between our vantage point and the name servers. Adopting the latter approach, therefore, reduces this limitation, whilst still maintaining the evidence of the severity of impact, providing us with a conservative lower bound of the visible effect on resolution time for domains. Therefore, we perform all our calculations by comparing the RTT of domains measured on the attack day, with the RTT of these domains measured on the previous day to infer adverse effects of the attack on the operations of DNS.

Figure 5 describes the RTT variations calculated on an

[7]Timestamps are in the local timezone w.r.t Victim X

hourly basis for both the attack instances. The red bars shows the attacks as observed on the network telescope. From the figure, we can see that the resolution time for domains rose significantly to about 0.7 seconds during the attack on December 2020, and hovered in this increased state for the next 7 hours, even though, as per the network telescope - the attack had subsided. There can be two reasons for this. First, this can mean that the attack was very effective and the server could no longer send backscatter traffic to the network telescope. Second, this can also suggest the existence of several composite attacks along with DoS attacks on the name server that hinder its functionality. These attacks, although not visible on the network telescope, affect the name servers' performance. Therefore, we see the increased state of resolution time for domains persist longer. We will discuss these limitations further in Section 6. Afterwards, the RTT value then gradually starts to fade for about 3 hours, before it normalises again. In the case of March 2021, the telescope recorded two separate attack instances. For both these attack instances, the RTT increase (by a factor 25 at its peak) was more significant compared to the attack on December. This result suggests that a higher intensity attack on the name server causes a more severe impact on its registered domains. We evaluate the severity of the impact of DoS attacks on the DNS infrastructure in the next section.

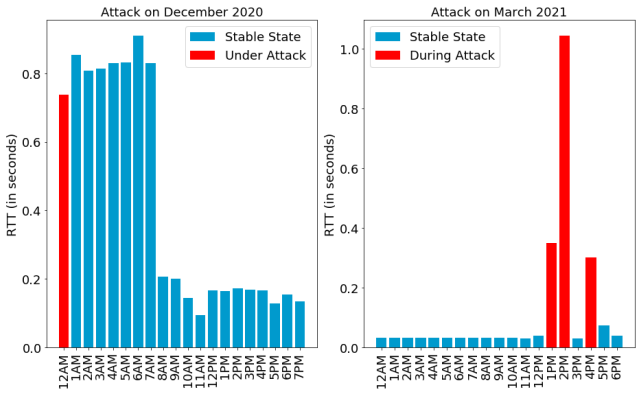

Figure 5: RTT Variations for December and March attacks on Victim X. During the attack on December, effects sustained long after the attack was detected to be over on the telescope. Additionally, we see a higher RTT increase for the domains using Victim X's name servers for March compared to December.

5.3 Effects of DoS attacks

In this section, we examine four aspects of the effects of DoS attacks on the DNS infrastructure. First, we investigate frequently targeted services on name servers by identifying protocol and port numbers from backscatter packets and tracking them using IANA assigned services name port numbers[8]. Second, based on services identified, we analyse the existence of a correlation between the intensity of attack and the severity of its impact. We then discuss the name server infrastructure of affected domains[9] and its relation to the severity of impact. Finally, we explore the visibility of the impact of an attack, on other services hosted on the same address block as that of the target name server.

5.3.1 Targeted Services

We analyse frequently targeted services on name servers, by tracking the target protocol and target port from our attack flows dataset. We match these instances, with zmap probes from Rapid7 during our measurement period to determine if the port was open or blocked. Figure 6 gives an account for the commonly targeted services on the DNS infrastructure.

As depicted in the figure, TCP is the most commonly targeted protocol, followed by UDP and ICMP. About $2/3^{rd}$ of the attacks on TCP, were on port 53 followed by port 80 and 443. DNS usually operates on UDP port 53, so we suspect the attacker's motive behind choosing TCP port 53 as the target, is resource exhaustion on the victim name server. Multiple erroneous TCP connections overwhelm the victim's CPU, thereby rendering it incapable of serving legitimate clients.

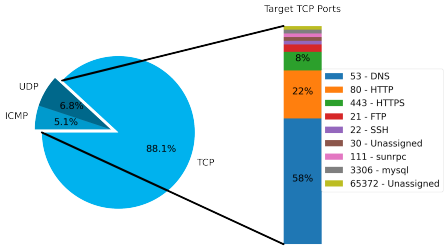

Figure 6: Commonly Targeted Services. About 90% of attacks are targeted on TCP, with $2/3^{rd}$ of them towards port 53

Moreover, referencing our dataset with zmap probes, revealed that TCP port 53 was open during all but one attack instance. DNS uses TCP as a transport mechanism only when TCP fallback is supported and followed-up on by the recursive resolver. Such requests/responses may include the transfer of a large number of RRs, IPv6 responses, zone transfers or DNSSEC responses. With TCP port 53 left in an "accept all" state, attackers target the application running on name servers by leveraging multiple TCP connections and exhausting its resources.

Based on our suspicion about the attackers' motive being resource exhaustion, we investigate in the next section, whether an increase in our attack intensity metrics of packet rate or the number of attacker IPs, increases the severity of impact on the resolution time of the target name servers' registered domains.

[8]https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml
[9]Registered Domains of a Name Server under attack

5.3.2 Impact Factor Correlation

An increase in the number of attacker IPs or packet rate on the network telescope signifies a higher intensity attack on the victim name server. Backscatter on the network telescope is detected only when the victim sends packets in response to the attacker's spoofed packets. An increase in the attacker IP count is observed, when the victim sends multiple IP flows (5-tuple) to the telescope. This can also be accompanied by a huge number of packets, either distributed across several flows or concentrated on one flow. From the victim's perspective, this signifies multiple erroneous connections being established that consume its internal resources. Resource exhaustion on the victim (name server), eventually leads to a degradation in name resolution for all of its registered domains.

However, a similar logic is applicable for packet rate as well. An increase in the packet rate is observed, when the victim sends a huge amount of response packets to the telescope, irrespective of the number of IP flows. Just like the previous case, this burst traffic, as seen on the telescope, may or may not be accompanied by multiple attacker IPs. From the victim's perspective, the enormous packet rate chokes up the network, thereby rendering it incapable of serving legitimate DNS requests. Therefore, an increase in impact factor, in either case, indicates an increase in the number of attacker IPs or an increase in packet rate, or both. We investigate the above correlation in our measurement data, between attack intensity metrics and impact intensity by using Pearson Correlation Coefficient. Its value ranges from -1 to 1, with 1 signifying a strong positive linear relationship and -1 being a strong negative linear relationship.

Our DNS measurements record the round trip times for the resolution of registered domains on a daily basis. Therefore, we also calculate Pearson's correlations with a daily granularity. Pearson's Correlation does not require any time-based bifurcations in the data. However, our daily windows allow us to verify that for a given day if an increase in the attack intensity metrics also results in an increase in the impact factor for registered domains of name servers under attack. Additionally, later in the section, we provide the correlation coefficients calculated on our entire dataset without the daily windows.

The variation of correlation between the impact factor and each of the attack intensity metrics - the number of attacker IPs and packet rate respectively, calculated on a daily basis during our measurement period is presented in Figure 7 and Figure 8. Lighter shades signify lower-intensity attacks, while darker shades show higher intensity attacks. As can be seen from Figure 7, a majority of attack instances have a high positive correlation between "number of attacker IPs" and the impact factor. A majority of attack instances - specifically MI and HI attacks lie on the positive side of y-axis. This positive correlation, although present, is comparatively less visible between packet rate and impact factor as can be seen from Figure 8. Attack instances are almost equally distributed between the positive and negative y-axis.

We also calculate the correlation coefficient between attack intensity metrics and impact factor for all the days of our measurement data, without taking into account the daily windows. The correlation coefficient between packet rate and impact factor is 0.033 with a P-Value of 0.809. The low value of correlation and high P-value, signifies that there is no correlation between packet rate and impact factor in our dataset. The correlation coefficient between the number of attacker IPs and the impact factor is 0.617 with a P-Value of 0.036. A value of 0.617 points to a positive correlation between these variables and a low P-Value signifies that there is only a 3% chance that our results occurred due to chance.

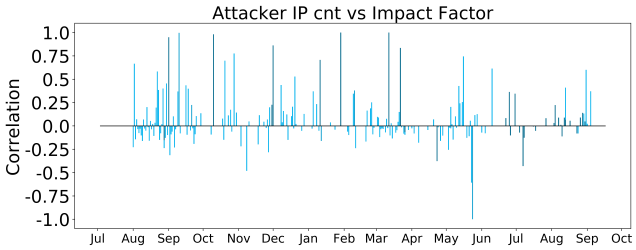

Figure 7: #Attacker IP Vs Impact Correlation. We see a strong positive correlation for a majority of days during our measurement period

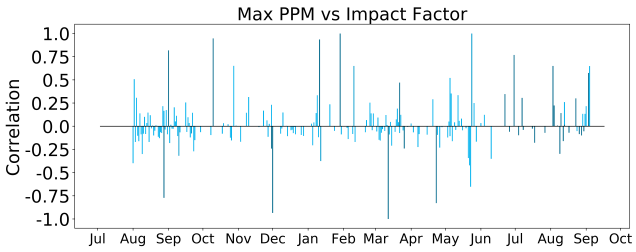

Figure 8: Packet Rate Vs Impact Correlation. A weaker correlation between impact factor and packet rate as compared to the number of attacker IPs

Based on our findings from Section 5.3.1, and an increased number of observations for a high correlation between the number of attacker IPs and impact factor, we advocate that an attack instance with a higher number of attacker IPs on the name server infrastructure, generates a larger effect on the resolution time for its registered domains. We investigate this effect on resolution time for domains in the next section. As we discuss further in Section 6, in order to overcome the limitation of intermittent network latency between our vantage point and the measured name server under attack, we take into account - the most prominent 24 days of measurement data during which we were able to report the maximum num-

ber of domains affected. These days are spread evenly, accounting for at least one day of every month within our measurement period. We base our next results, on these 24 days of attacks comprising more than 42M domains.

5.3.3 Effects on Registered Domains of name servers

As discussed in Section 4.3, the attack percentage for a domain is defined as the proportion of its name server under attack to the total number of name servers. Figure 9 shows the total number of affected domains on the y-axis versus different attack percentages on the x-axis during the days of our analysis. The colors represent the name server infrastructure of these domains. For example, the bars at 50% and 66% respectively show the domains, for which "1 out of 2" and "2 out of 3" name servers were under attack. Furthermore, the bar at 50% is mostly turquoise with a small minority as orange, accounting for a majority of domains with 2 name servers, with a minority of domains having 6 name servers. Additionally, we divide the x-axis as follows - Category A - 0% - 20%, Category B - 20% - 50% and Category C- 50% - 100%. This categorisation is based on the fact that Category A contains all LI attack instances, Category B contains a combination of LI and HI attack instances and Category C contains a mixture of all three LI,MI & HI attack instances.

Category A & B - As depicted in Figure 9, all domains in category A have 3 or more authoritative name servers, but they comprise a minor proportion, specifically 7% of the total number of affected domains. Category B follows a similar trend comprising of 12% of the total affected domains and having 3 or more name servers. Category C - 81% of the affected domains fall under this category. Moreover, most of the domains here, are served by only one or two authoritative name servers.

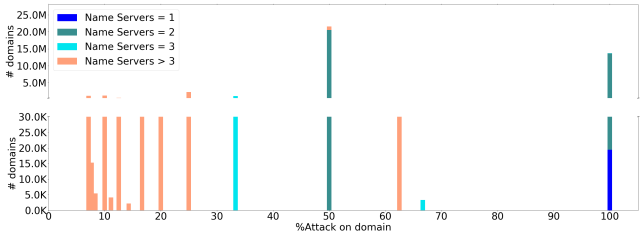

Figure 9: Registered Domains distribution over attack percentages. 81% of affected domains have 1-2 authoritative name servers

The distribution has two major implications. First, when DoS attacks are targeted towards the name server infrastructure, a majority of domains that are affected, either have just one or no secondary servers to carry on with DNS resolutions. Second, DNS specifications recommend having at least three name servers for organisational level zones [33], therefore a majority of domains in the .com TLD do not follow this recommendation. Without secondary servers to aid name resolution in the event of an attack, we hypothesize that the performance degradation of such domains magnifies with higher intensities of attacks.

To validate the above hypothesis, we reference Table 2 and take the case of HI attacks from our dataset because here we see the maximum increase in impact factor. We calculate the impact factor with the granularity of a day because of the measurement time mismatch between our DNS probes and the attack observed on the telescope. This will be discussed further in Section 6. Additionally, in order to accommodate for the limitations of network outages between our vantage point and the name servers under attack, we take into account, the attack instances on name servers for which at least 10K registered domains were measured. Therefore, our results on the impact factor increase for different categories of attack (A,B or C) are conservative lower bounds, and the actual effect on resolution time degradation may be higher. Figure 10 shows the variation of impact factor for different categories of attack as defined earlier.

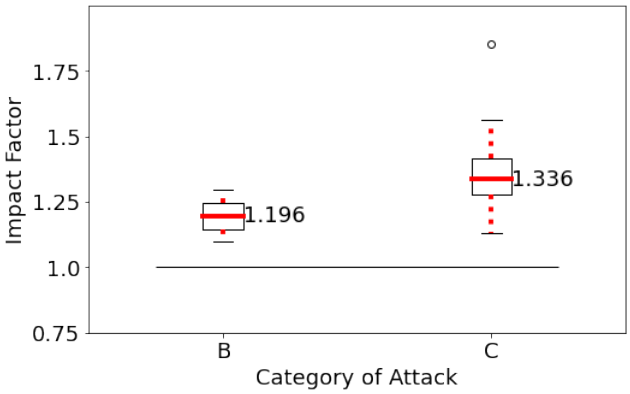

Figure 10: Impact factor variation for High Intensity Attacks. An increase in attack percentage causes an increase in impact factor

From Figure 10, we only have categories B and C , because category A is dominated by LI attacks. As can be seen from the figure, the median impact factor for category B (dominated by domains with 3 or more name servers) is 1.196, which is lower than the median impact factor for category C (dominated by domains with 1 or 2 name servers) - 1.336. As discussed previously in Section 4.3, any deviation from the baseline impact factor of 1 demonstrates a deterioration in resolution performance for domains. The whiskers and box are constrained for category B as compared to category C, suggesting a smaller variation of impact factor among domains affected. In addition, the maximum impact factor for category B is in-line with the 25^{th} percentile of category C, signifying a more severe impact on name resolutions for three-quarters of domains belonging to category C. Therefore, during HI attacks, domains using more than 3 name servers experience a lesser degrada-

tion of resolution time compared to domains that use 1 or 2 name servers.

5.3.4 Effects on same IP Prefix

After analysing the impact on domains in the previous section, we investigate the visibility of the effects of an attack, on other services hosted in the same address block as that of the target name server. Figure 11 shows the variation of impact factor for domains using name servers (other than the target name server) in the /24 prefix, compared to the impact factor variation for domains using the name server under attack. We demonstrate that the impact factor variation observed for domains using other name servers in the same address block, applies to other services as well, such as web hosting, email, etc. within this prefix.

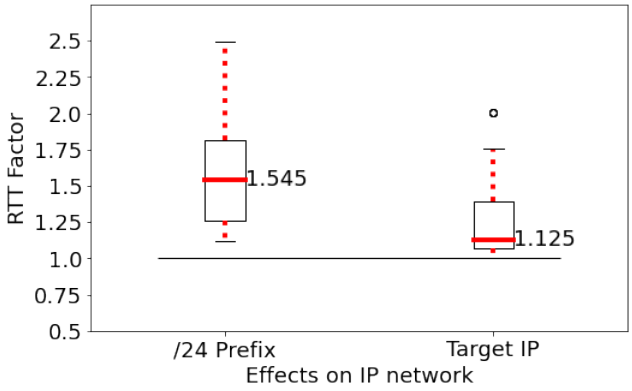

Figure 11: Impact factor variation for services in the same address block vs target name server. Effects of an attack are visible on other services of the network.

We measure the visible effects of an attack using the impact factor, which is calculated with the granularity of a single day rather than the duration of the attack, due to which the impact on the /24 address block described below is a lower bound of the effects due to DoS attacks. We take the case of only HI attacks because of the maximum visibility of the degradation in resolution time. We then plot attack instances in which the impact factor is greater than 1 for all domains of target name servers, which is why y=1 line touches the minimum value for the boxplot of "Target IP". This allows us to track, if a similar performance degradation is experienced by other services in the /24 address block. As can be seen from Figure 11, boxplots for the Target IP and the /24 prefix are proportional in terms of the structure of the box and length of whiskers. This suggests that the variation of impact factor in both the cases is similar. Although the boxplots are proportional, the median and interquartile values for /24 prefix are significantly higher than that of the target IP. This suggests that not only are the effects of an attack visible on other services in the same address block, but more often than not, the impact is higher.

From Section 5.3.2, we can see that there are several days of attack instances when a negative correlation exists between attack intensity metrics and impact intensity, suggesting the presence of resilience techniques within DNS to mitigate DoS attacks. Moreover, from Section 5.3.3, we also see that adopting defenses such as multiple authoritative name servers, instead of just one or two, can reduce the severity of impact on domains. We investigate the efficacy of these resilience techniques in the next section.

5.4 Resilience Techniques

In this section, we investigate two primary resilience techniques that can be deployed by registered domains, when DoS attacks are targeted on their name servers - AS level segmentation of name servers and the use of anycast enabled name servers. We investigate the efficacy of these resilience techniques next.

5.4.1 Use of Anycast Name Servers

In this section, we analyse, if the use of anycast enabled name servers is an efficient DoS resilience mechanism for domains. We classify domains into three categories based on the composition of their authoritative name servers as follows: **Full Anycast** - If all the name servers for a domain use anycast, **Partial Anycast** - If only a fraction of the name servers for a domain use anycast and the rest use unicast and **Full Unicast** - If all the name servers for a domain use only unicast. We then compare the average impact factor for domains belonging to each category, based on whether its' unicast or anycast name servers were under attack. As discussed previously and further outlined in Section 6, our measurement system can switch between name servers across several days of measurement. Therefore, in the case of domains that use heterogeneous name servers, responses to DNS queries might be given by the name server that is not under attack, leading to false positives during resolution time calculations. Additionally, for each pair **[Category,NS under attack]**, there is a huge variation in the number of domains measured. For example, the pair of [Full Anycast, Anycast], has 5M measured domain names, however, for the case of [Partial Anycast,Anycast] we have just 2200 domains measured. Intermittent network outages between our vantage point and the name server under attack, may skew our results in such a case where so few domains are measured. In order to circumvent these limitations, we first take the case of HI attacks by referencing Table 2 because we see the highest rise of impact factor in HI attacks. Next, we choose a threshold number of domains - "n", and cycle through each pair of [Category,NS under Attack], by randomly selecting "n" domains in each pair. This process is repeated "t" times to to calculate the average impact factor for each pair. However, the choice of "n" and "t", is independent of our results, as long as we do not over-sample domains from

a specific pair of [Category,NS under attack].

First, we take the value of n=1000 and t=3, i.e. 1000 domains at random from each pair of [Category,NS under attack] repeated 3 times, to derive our results below. Figure 12 shows the efficacy of using anycast name servers as a resilience technique. The categories defined above are shown along the x-axis and the type of name server under attack is listed on the y-axis. As discussed earlier, an impact factor of 1 or less, signifies no effect on resolution time for domains. An impact factor greater than 1 demonstrates a significant deterioration of resolution time. Therefore, the percentage increase from an impact factor of 1 is used as a radius while plotting the circles for each category, which signifies the RTT degradation. The impact factor for each category, is also explicitly mentioned in the figure.

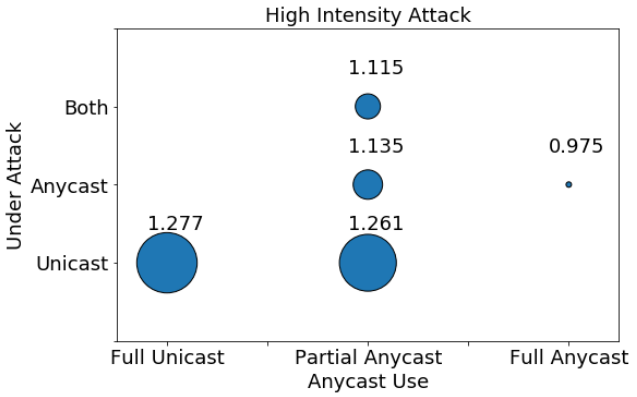

Figure 12: Efficacy of Anycast as a DoS resilience technique, for n=1000 & t=3. For HI attacks, using anycast enabled name servers reduces the severity of impact by 30%.

As we can see from the figure, the use of Full Anycast name server setup reduces the severity of impact by 30% as compared to the use of a Full Unicast setup. The case of partial anycast is particularly interesting. We see an increase in the magnitude of impact severity when the unicast part is under attack, as compared to when the anycast part is under attack. With consideration to the name server ambiguity discussed earlier, the impact factor measured during an attack on unicast name servers would be a conservative lower bound of resolution time degradation, whereas the impact factor would be a conservative upper bound of resolution time degradation when anycast name servers are under attack.

Consider the following example - a set of domains use both unicast and anycast name servers. As a worst-case scenario, we received all responses for a given domain from the anycast name server while the unicast name server was under attack. In this situation, the impact factor comes out to be x, i.e. the best possible round trip time. Conversely, we received all responses for a given domain from the unicast name server while the anycast name server was under attack. In this situation, the impact factor comes out to be y, i.e., the worst possible round trip time. Looking at our empirical evidence, we see y to be less than x, i.e., the worst performance of anycast name servers under attack is still better than the best performance of unicast name servers under attack.

Additionally, we investigate if our results change with varying values of n and t. Figure 13, shows the efficacy of anycast as a resilience technique for the case of n=10000 and t=5. As can be seen from Figures 13 and 12, an increase in the threshold value of n, eliminates [Category,NS Under Attack] pairs for which lesser domains were measured. Moreover, the impact factor increases for each pair (that is common to both figures) in this scenario, because of the average value calculations done over a larger set of domains. However, for a given value of n,t - we see the same trend of RTT variations, i.e., the impact factor is highest for the category of Full Unicast, followed by Partial Anycast and Full Anycast.

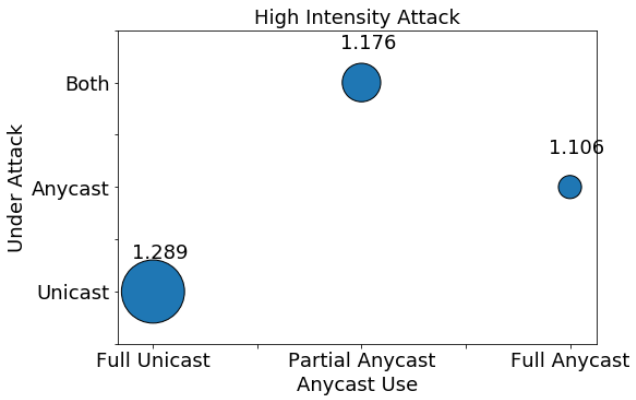

Figure 13: Efficacy of Anycast as a DoS resilience technique, for n=10000 & t=5. For HI attacks, using anycast enabled name servers reduces the severity of impact by 16%.

In this case, the use of Full Anycast name server setup reduces the severity of impact by 16% as compared to the use of a Full Unicast setup. Additionally, because of the limitations in our data as mentioned previously in this section, impact factor calculations without the n,t considerations provide us with a worst-case performance of a Full Anycast name server setup and the best-case performance of a Full Unicast setup. Even in this case, we see a Full Anycast Setup outperform the Full Unicast setup by 6%. Therefore, using a anycast name server setup for domains provides significant protection against DoS attacks targeted at the DNS infrastructure.

5.4.2 AS-level Resilience

In this section, we analyse if AS level segmentation of name servers, for domains provides resilience against DoS attacks. Just like in the previous section, we sample domains (n) for each pair [Total number of AS, AS under attack], and repeat the analysis t times to derive our results below. The choice of “n” and “t” here also is inde-

pendent of our results, as long as we do not over-sample domains from a specific pair of [Total number of AS, AS under attack].

First, we take the value of n=1000 and t=4, i.e., 1000 domains sampled at random from each pair of [Total number of AS, AS under attack] repeated 4 times, to derive our results below. We refer to Figure 14, where for a given domain, the AS-level composition of name servers in shown in the x-axis and the number of ASNs attacked is shown on the y-axis. The percentage increase from an impact factor of 1 is used as radius while plotting the circles for each pair, signifying the RTT degradation. The impact factor for each category, is also explicitly mentioned in the figure. As can be seen from the figure, moving from left to right - the magnitude of impact severity decreases with each additional AS by 10-15%. Varying the values for n and t revealed that the impact factor decreases with each additional AS by 5-20%.

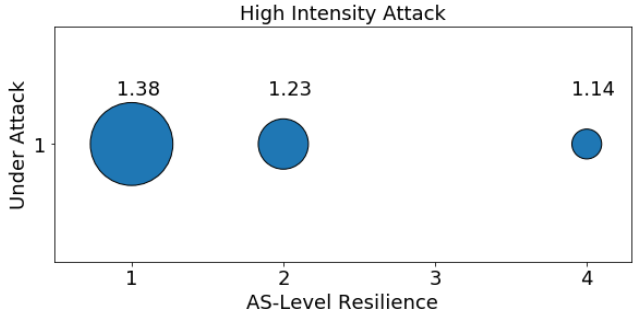

Figure 14: Efficacy of AS segmentation of name servers as a DoS resilience technique, for n=1000 & t=4. For HI attacks, each additional AS reduces the severity of impact by 10-15%.

Using name servers that belong to separate ASNs provides topological diversity to domains. Therefore there isn't one single point of failure when name servers of one ASN get attacked, as the domain has other name servers in disjoint ASNs to carry on with name resolutions. Therefore, for a given domain, having AS level diversity among authoritative name servers is an efficient mechanism to mitigate resolution time degradation due to DoS attacks. Based on empirical evidence in our data, we argue that domains using resilience techniques such as anycast enabled name servers or AS level diversity in authoritative name servers, are efficient resilience techniques when DoS attacks are targeted on their name server infrastructure.

However, one major limitation of our data is the misalignment of measurement times from active DNS measurements and the network telescope, as previously discussed in Section 5.2. Our analysis accommodates these limitations and provides conservative lower bound results through Sections 5.1-5.4. However, to completely bypass this limitation and validate our methodology, we use "reactive-measurements", which is discussed next.

5.5 Reactive Measurements

In this section, we describe a first effort in reactively evaluating the visible effect on resolution time degradation of domains, immediately after attacks are observed on their name servers. To this effect, we adopt the following procedure. When attacks are observed on name servers, we send out a fixed set of DNS queries for a randomly selected set of its 50 registered domain names. We chose the threshold to be small, i.e., 50 domains, so that our measurements do not put any extra stress on name servers already under attack. Additionally, we measure this fixed set of domains every 5 minutes for a 24 hour period from the start of the attack. This allows us to evaluate the degradation of resolution time for domains during the attack as well as some time after the attack has ended.

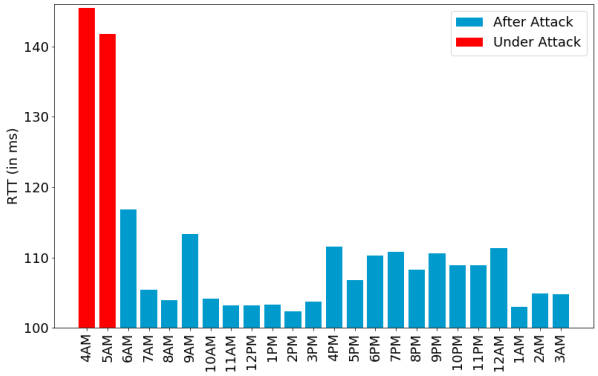

Figure 15: Avg. RTT variation for registered domains during an attack on Target Zero's name servers. RTT gradually recedes from a factor of 1.44, 12 hours after the attack

As a first step to analyse reactive measurements, we look at an attack on - "Target Zero" in the month of November 2021. The attack on Target Zero was a High Intensity attack with traffic volume peaking at 1 Gbps. Target Zero had two name servers that were under attack, but they were configured to use the same IP address.

Figure 15 shows the average RTT variation every hour for a set of 50 registered domains of the name servers of Target Zero. The average impact factor for domains, in this case, was 1.44. As depicted in the figure, the RTT gradually goes down from 148ms during the first half-hour of the attack to 103ms, 12 hours later. The anomalous peaks in the figure are attributed to usual variations in resolution times for domains. Since these measurements were done reactively, the gradual RTT decrease, as can be seen from the figure, validates our methodology of using impact factor to infer the adverse effects of attacks on DNS operations. As future work, we plan to analyse reactive measurements across a broader set of attack instances over multiple months of measurement data.

6 Limitations

Due to the architecture of our measurement system, there are certain limitations to our analysis that impact our results, which are described as follows. First, OpenINTEL performs active DNS measurements from a single vantage point in the Netherlands. Therefore, RTT for domains that use authoritative name servers in the European region have a regional bias compared to the RTT of domains registered to name servers located elsewhere. Moreover, the RTT bias also arises for domains that use DNS-based load balancing, such that our vantage point receives answers to queries for these domains from regional name servers. However, our results are based on the observable improvement/deterioration of RTT for domains, consequent to DoS attacks on the DNS infrastructure, rather than the actual values of RTT. Therefore, evaluating the magnitude of RTT increase/decrease reduces the impact of this limitation on our results.

Second, OpenINTEL performs forward DNS measurements on any given domain once every 24 hours. Registered domains for a given name server are measured at different times of the day, i.e. an attack on name servers does not trigger DNS measurements to evaluate the resolution of its registered domains. Therefore, a limited number of domains are measured in the duration of an attack on its name servers, leading to an underestimation of the visible effect of attacks on resolution times. Our methodology described in Section 4.3 provides a conservative lower bound approach to account for this limitation as well as observe the effect on domain resolution time. Furthermore, we also perform reactive measurements, i.e. DNS measurements for domains triggered when attacks on name servers are observed, to alleviate this limitation as well as to validate our prior methodology.

Third, as outlined in Section 3.2, instead of using the glue records (A records associated with the NS records) in zone files, we use our measurement system to learn the name servers for domains via NS queries and subsequent A queries during active resolution. We generally trust that the authoritative NS responses received during active resolution, are the same that are defined in the parent zone. About 10% of domains do not conform to this, i.e. there are name server inconsistencies between parent and child zones [34].

Fourth, we quantify the effect on domains by defining the impact factor as the domain's resolution RTT calculated on the attack day compared to the RTT on the previous day. Comparing the RTT for a single day with the attack day, may lead to false negatives, as resolution RTT of a domain may already be in an elevated state on the previous day. This can be due to a variety of reasons, such as a change in name-server infrastructure for the domain or intermittent network outages between our vantage point and the measured name server. As part of our future work, we plan to extract a stable state resolution RTT of domains using measurement data of multiple days rather than just a single day.

Additionally, intermittent disruptions and latency in between our vantage point and authoritative name servers being measured, can lead to an overestimation of resolution times for its registered domains. Moreover, measurement of domains with heterogeneous authoritative name servers (for example, unicast and anycast) can lead to an underestimation of resolution times consequent to DoS attacks on the DNS infrastructure. This is because our measurement system can switch between name servers across several days of measurement for a given domain name. This might result in false positives when assessing the impact, or the effectiveness of deployed defenses. After verification across several days of attack instances, we discovered that on any given day, measuring at least 1M domains, or for any given name server, measuring at least 10K of its registered domains, provides us a conservative norm of the effect on resolution times while accounting for both the above limitations. Therefore, measuring the RTT improvement/deterioration across a significant number of domains, reduces the chances of network delay or false positives as contributing factors to our results.

Finally, the end time of an attack cannot be universally inferred from the network telescope. A victim is observed to be under a DoS attack on the network telescope, only when it receives backscatter traffic from it. However, an effective attack on the victim would cause it completely shut down, thereby sending out no backscatter traffic to the telescope. In such a case, the telescope records the attack to have ended, while in actuality, the attack endured on the victim for a longer period of time.

7 Ethical Considerations

Our goal is to measure the visible effect on resolution times and resilience of the DNS infrastructure consequent to DoS attacks. However, during our analysis, we found attack instances on several major organisations, with some organisations being the target of frequent attacks. Additionally, our analysis also reveals the lack of resilience in several domains. We find it unacceptable that our paper can be used as a basis to launch further DoS attacks on these vulnerable organisations and domains. Therefore, we do not reveal their identities throughout the paper, and instead use pseudonyms wherever necessary.

Additionally, our reactive measurements are triggered when the name server infrastructure is observed to be under attack. The stress of our measurements should not be a burden to the name servers already under attack. To this effect, we issue our queries only for a small set of domains. We also limit our requests, issuing queries

to these name servers in 5-minute intervals. The small scope and rate-limited nature of these measurements allows us to gather as much information as possible from the name server infrastructure under duress, without incurring further stress on it.

8 Conclusions and Future Work

Our paper presents a first look on the operations of the global DNS infrastructure and the efficacy of its defenses in the face of DoS attacks detected in the wild. In our study, we have classified DoS attacks on the DNS infrastructure based on their intensity, and found a majority of the attacks were low-intensity attacks. Our investigations also revealed TCP port 53 as the main target for attacks on name servers. Our analysis across several days of measurement also suggests the presence of a positive correlation between the impact of an attack and the number of attacker IPs related to the attack. Furthermore, we developed a method using the resolution time of domains to infer the impact of attacks on the operations of DNS. With this, we determined domains having at least one "available" name server, experience a lesser degradation in resolution time as compared to domains with no "available" name servers. Additionally, we found that the impact of an attack on the name server infrastructure is also visible to other services hosted on the same address block as that of the target IP. Based on our analysis, we advocate that the use of resilience techniques by domains such as anycast enabled name servers, and multiple (AS-level) name servers, reduces the impact of attacks by 15-30% and 5-20%, respectively. Finally, we take a first step in analysing reactive measurements for a name server under attack and determine that the resolution time improves gradually following the attack duration.

Our work focuses on the *.com* TLD because it is the largest zone covering a significant portion of the namespace [35], therefore our analysis is representative of the global DNS infrastructure. As future work, we will be extending our analysis to other major TLDs as well, such as *.org* and *.net* to see if there are any inter-zone differences. Our methodology determines the impact of an attack based on the degradation of resolution time for domains. As future work, we plan to integrate failure of domain resolution, along with the resolution time as metrics to infer the impact of attacks. Finally, our work presents a first look at analysing reactive measurements. As future work, we plan to analyse reactive measurements across a broader set of attack instances over multiple months of measurement data.

References

[1] "Domain names - concepts and facilities," RFC 1034, Nov. 1987. [Online]. Available: https://www.rfc-editor.org/info/rfc1034

[2] T. H. Kim and D. Reeves, "A survey of domain name system vulnerabilities and attacks," *Journal of Surveillance, Security and Safety*, vol. 1, no. 1, pp. 34–60, 2020. [Online]. Available: http://dx.doi.org/10.20517/jsss.2020.14

[3] S. Dietrich, N. Long, and D. Dittrich, "Analyzing distributed denial of service tools: The shaft case," in *Proceedings of the 14th USENIX Conference on System Administration*, ser. LISA '00. USA: USENIX Association, 2000, p. 329–340.

[4] M. Antonakakis, T. April, M. Bailey, M. Bernhard, A. Arbor, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, A. Arbor, L. Invernizzi, M. Kallitsis, M. Network, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, Y. Zhou, M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the Mirai Botnet This paper is included in the Proceedings of the Understanding the Mirai Botnet," *USENIX Security*, pp. 1093–1110, 2017. [Online]. Available: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

[5] "Eight-hour ddos attack struck aws customers," https://www.darkreading.com/cloud/eight-hour-ddos-attack-struck-aws-customers, (Accessed on 01/28/2022).

[6] N. Perlroth, "Hackers used new weapons to disrupt major websites across u.s. - the new york times," https://www.nytimes.com/2016/10/22/business/internet-problems-attack.html, 2016.

[7] "Ovh hosting hit by 1tbps ddos attack, the largest one ever seensecurity affairs," https://securityaffairs.co/wordpress/51640/cyber-crime/tbps-ddos-attack.html, (Accessed on 01/28/2022).

[8] R. Sommese, G. Akiwate, M. Jonker, G. Moura, M. Davids, R. van Rijswijk - Deij, G. Voelker, S. Savage, K. Claffy, and A. Sperotto, "Characterization of anycast adoption in the dns authoritative infrastructure," in *TMA Conference 2021*. IFIP, Sep. 2021, 5th Network Traffic Measurement and Analysis Conference, TMA 2021, TMA 2021 ; Conference date: 14-09-2021 Through 15-09-2021.

[9] G. C. Moura, R. d. O. Schmidt, J. Heidemann, W. B. de Vries, M. Muller, L. Wei, and C. Hesselman, "Anycast vs. DDoS," pp. 255–270, 2016.

[10] "Root server operations attack - 2015, events of 2015-11-30," https://root-servers.org/media/news/events-of-20151130.txt.

[11] G. C. M. Moura, J. Heidemann, M. Müller, R. D. O. Schmidt, and M. Davids, "When the Dike Breaks : Dissecting DNS Defenses During DDoS," *Proceedings of the Internet Measurement Conference 2018 (IMC '18)*, pp. 8–21, 2018.

[12] L. Kröhnke, J. Jansen, and H. Vranken, "Resilience of the Domain Name System: A case study of the.nl-domain," *Computer Networks*, vol. 139, pp. 136–150, 2018.

[13] J. Yuan and K. Mills, "Monitoring the macroscopic effect of DDoS flooding attacks," *IEEE Transactions on Dependable and Secure Computing*, vol. 2, no. 4, pp. 324–335, 2005.

[14] M. Anagnostopoulos, G. Kambourakis, P. Kopanos, G. Louloudakis, and S. Gritzalis, "DNS amplification attack revisited," *Computers and Security*, vol. 39, no. PART B, pp. 475–485, 2013.

[15] J. Park, A. Khormali, M. Mohaisen, and A. Mohaisen, "Where Are You Taking Me? Behavioral Analysis of Open DNS Resolvers," *Proceedings - 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2019*, pp. 493–504, 2019.

[16] A. Abhishta, R. Van Rijswijk-Deij, and L. J. Nieuwenhuis, "Measuring the impact of a successful DDoS attack on the customer behaviour of managed DNS service providers," *Computer Communication Review*, vol. 48, no. 5, pp. 70–76, 2018.

[17] G. Kambourakis, T. Moschos, D. Geneiatakis, and S. Gritzalis, "Detecting DNS amplification attacks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5141 LNCS, no. October 2002, pp. 185–196, 2008.

[18] G. C. Moura, J. Heidemann, R. d. O. Schmidt, and W. Hardaker, "Cache me if you can: Effects of DNS time-to-live," *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, no. 1, pp. 101–115, 2019.

[19] C. L. Schuba, I. V. Krsul, M. G. Kuhn, E. H. Spafford, A. Sundaram, and D. Zamboni, "Analysis of a denial of service attack on TCP," *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, no. June, pp. 208–223, 1997.

[20] G. Kambourakis, T. Moschos, D. Geneiatakis, and S. Gritzalis, "Detecting DNS amplification attacks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5141 LNCS, no. October 2002, pp. 185–196, 2008.

[21] "Autonomous system number (asn) - afrinic - regional internet registry for africa," https://afrinic.net/asn.

[22] D. Moore, G. M. Voelker, and S. Savage, "Inferring internet denial-of-service activity," *Proceedings of the 10th USENIX Security Symposium*, vol. 24, no. 2, pp. 115–139, 2001.

[23] C. Fachkha, E. Bou-Harb, and M. Debbabi, "Fingerprinting internet DNS amplification DDoS Activities," *2014 6th International Conference on New Technologies, Mobility and Security - Proceedings of NTMS 2014 Conference and Workshops*, pp. 1–5, 2014.

[24] L. Krämer, J. Krupp, D. Makita, T. Nishizoe, T. Koide, K. Yoshioka, and C. Rossow, "AmpPot: Monitoring and defending against amplification DDos attacks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9404, pp. 615–636, 2015.

[25] M. Bailey, E. Cooke, and D. Watson, "A hybrid honeypot architecture for scalable network monitoring," 2004.

[26] M. Jonker, A. Sperotto, R. Van Rijswijk-Deij, R. Sadre, and A. Pras, "Measuring the adoption of DDoS protection services," *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, vol. 14-16-November-2016, pp. 279–285, 2016.

[27] M. Jonker, A. Pras, A. Dainotti, and A. Sperotto, "A first joint look at DoS attacks and BGP blackholing in the wild," *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, pp. 457–463, 2018.

[28] R. Sommese, L. Bertholdo, G. Akiwate, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. Claffy, and A. Sperotto, "Manycast2: Using anycast to measure anycast," in *Proceedings of the ACM Internet Measurement Conference*, ser. IMC '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 456–463. [Online]. Available: https://doi.org/10.1145/3419394.3423646

[29] "Ucsd network telescope daily randomly and uniformly spoofed denial-of-service (rsdos) attack metadata - caida," https://www.caida.org/catalog/datasets/telescope-daily-rsdos/, (Dates used - 06/22/2021 - 02/17/2022).

[30] R. van Rijswijk, M. Jonker, A. Sperotto, and A. Pras, "A high-performance, scalable infrastructure for large-scale active dns measurements," *IEEE*

journal on selected areas in communications, vol. 34, no. 6, pp. 1877–1888, Jun. 2016.

[31] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.

[32] M. Allman, "Comments on DNS robustness," *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, pp. 84–90, 2018.

[33] M. A. Patton, S. O. Bradner, R. Elz, and R. Bush, "Selection and Operation of Secondary DNS Servers," RFC 2182, Jul. 1997. [Online]. Available: https://www.rfc-editor.org/info/rfc2182

[34] R. Sommese, G. C. M. Moura, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. Claffy, and A. Sperotto, "When parents and children disagree: Diving into DNS delegation inconsistency." in *Proceedings of the Passive and Active Measurement Workshop*, Eugene, OR, USA, 2020.

[35] "Global domain report - sidn," https://www.sidn.nl/downloads/59TDj5jUtLH8FeC7XOptG3/1795894a2bb7ff5e19782d454de0900c/Global_Domain_Report_2020_Sedo.pdf.