SIDN Labs

https://sidnlabs.nl

Jun 25, 2020

Peer-reviewed Publication

Title: Into the DDoS maelstrom: a longitudinal study of a scrubbing service

Authors: Giovane C. M. Moura, Cristian Hesselman, Gerald Schaapman, Nick Boerman, Octavia de Weerdt

Venue: In: 5th International Workshop on Traffic Measurements for Cybersecurity (WTMC 2020), co-located with 2020 5th IEEE European Symposium on Security and Privacy Workshops (EuroSPW)

DOI: TBD
Conference dates: September 7th, 2020

Citation:

- Giovane C. M. Moura, Cristian Hesselman, Gerald Schaapman, Nick Boerman, Octavia de Weerdt . Into the DDoS maelstrom: a longitudinal study of a scrubbing service . In : 5th International Workshop on Traffic Measurements for Cybersecurity (WTMC 2020), Genova, Italy.
- Bibtex:

```
@inproceedings{Moura20a,
  author =  Moura, Giovane C. M. and Hesselman, Cristian
  and Schaapman, Gerald and Boerman, Nick and
  de Weerdt, Octavia},
  title = { Into the DDoS maelstrom: a longitudinal
  study of a scrubbing service.}},
  booktitle = {5th International Workshop on Traffic
  Measurements for Cybersecurity (WTMC 2020)},
  year = {2020},
  address = {Genova, Italy},
}
```

Into the DDoS maelstrom: a longitudinal study of a scrubbing service

Giovane C. M. Moura
SIDN Labs
SIDN
Arnhem, The Netherlands
giovane.moura@sidn.nl

Cristian Hesselman
SIDN Labs and University of Twente
SIDN
Arnhem, The Netherlands
cristian.hesselman@sidn.nl

Gerald Schaapman
NBIP
Ede, The Netherlands
gerald@nbip.nl

Nick Boerman
SIDN
Arnhem, The Netherlands
nick.boerman@sidn.nl

Octavia de Weerdt
NBIP
Ede, The Netherlands
octavia@nbip.nl

***Abstract*—Distributed denial-of-service (DDoS) attacks are nowadays easy and cheap to carry out, and have become bigger and more frequent over the last years. Cloud-based scrubbers have emerged as a service which victims can hire on demand to fend off attacks. There are many industry players, but not much insights into their operations. This work unravels for the first time the inner workings of a DDoS scrubber — NaWas— a non-profit scrubber in the Netherlands. We analyze 1800+ DDoS attacks spanning over a period of 22 months, and show that while most attacks are not very large, they are still large enough to disrupt services and likely to disturb links. We estimate the collateral damage incurred by DDoS attacks, and demonstrate that the number of victims of is *at least* quadratically larger (IP^2) than the targeted addresses. Last, by correlating attacks metadata with authoritative DNS traffic, we show that DDoS attacks leave fingerprints on DNS traffic, which, in turn can be used to detect DDoS attacks at early stages, even if attackers attempt to deceive DNS based detection.**

1. Introduction

C1 Denial-of-service (DoS) attacks aim at overwhelming a machine or network resources with bogus traffic to disrupt its operations. Once overwhelmed, the service is susceptible to extortion [1], which can drive clients away. In some cases, attacks can last for weeks [2] or even months [3].

C1 DoS and Distributed DoS (DDoS) have become more frequent, bigger and cheaper over the last years. Anyone with a few euros can hire a DDoS attack [4], [5] from the so-called "booter" websites [6], [7], which offer DDoS-as-a-service. The emergence of large numbers of Internet-of-Things (IoT) devices has inflated the attacks firepower, surpassing the Tbps scale [8], [9], [10], [11]. A recent (Oct 2019) DDoS attack against Amazon's DNS anycast service [12] caused significant outages in various places — demonstrating that not even hypergiant cloud providers are safe from such attacks.

To cope with such increase in power and frequency of DDoS attacks, we have seen the development of the DDoS protection industry, which became a multi-million D1 dollar enterprise. In this industry, vendors either sell dedicated hardware or "cloud-based" solutions — the latter is commonly known as traffic *scrubbing services*, which can be activated on-demand by clients under attack, by redirecting traffic to the scrubbing service using typically DNS [13] or Border Gateway Protocol (BGP) protocol [14], [15]. Scrubbing services aim at discarding bogus traffic while forwarding only the legitimate part of the traffic, so services can remain active and operational during an attack.

Even though some scrubbing services providers release yearly reports (*e.g.,*[16], [17]), little is actually known about the frequency, victims profile, collateral damage, and the specifics of long term trends DDoS attacks that are filtered by such scrubbing services.

To fill this void, this works unravels for the first time the inner workings of a scrubbing provider, by analyzing DDoS attacks filtered at NaWas [18], a non-profit scrubber based in the Netherlands. Just like any scrubbing service provider, NaWas owns and operates various dedicated DDoS filtering hardware and provides on-demand filtering to its members (Section 2). The main difference between NaWas and other commercial services is its business model: NaWas uses a cooperative model in which members share the costs, and which the main goal is to provide stable service instead of generating profits.

We present a longitudinal analysis of DDoS attacks filtered at NaWas spanning over 22 months (July 2017 – May 2019). In total, we characterize more than C1 1800 DoS/DDoS attacks (Section 3) and show there is a large variation in both size and duration of attacks: we find that most attacks are of medium size (average 3.9 Gbps), and last less than two hours in average, *after* the scrubber service is activated. Even if they seem small compared to current Tbps attacks [10], they are enough to overwhelm many webservers as well as inter-domain links [7]. Moreover, we estimate the collateral damage caused by these attacks (Section 4) by looking at second-level domain

names (2LD, such as example.org and example.co.jp) hosted under the attacked IP addresses. We find that the attacked domain name space is at least *quadratically* larger than the number of attacked IP addresses (IP^2), given the same IP may host hundreds or thousands of domains (a practice known as shared hosting). Last, we find evidence of these DDoS attacks at higher levels of DNS hierarchy (Section 5) and classify a subset of the DDoS attacks with regards to their DNS traffic. The idea is that the fingerprints left on DNS by DDoS attacks can be used to improve DDoS detection mechanisms.

2. National Scrubbing Service (NaWas)

NaWas can be seen as association from which members can request DDoS scrubbing services on demand. It is a non-profit scrubber service originally designed to be a national scrubber of the Netherlands, and it is operated by NBIP [19] (*Nationale Beheersorganisatie Internet Providers*), the Internet Service Provider Management Association of the Netherlands. It has been active since 2014 and since 2018 it has extended its services to other European countries. It currently has 151 members (Jan. 2020), covering large and small ISPs and other businesses from four European countries.

Economic incentive: By being non-profit, NaWas is able to offer very competitive prices to its members. Most importantly, it offers a *flat-fee* price model, meaning that the costs *are not tied to the size of the attack*. This is a major difference with current commercial services (*e.g.,*[20]), in which the DDoS protection is metered, and, as such, the services costs can quickly escalate with the duration and intensity of the attack. This non-metered based model is the main economical incentives for members to join NaWas. There are currently five membership plans, which prices vary according to the number of protected network prefixes.

Operations: NaWas operates a 24/7 C2 infrastructure that include set of diverse DDoS filtering hardware from multiple vendors located in several data centers. Whenever a client is under attack, it can request the scrubbing services by using BGP to redirect traffic to NaWas filtering locations. One requirement is that the client is present in at least one of the locations that NaWas is present, so they can establish a direct peering session. To redirect traffic, members make more specific BGP announcements of the prefixes under attack. D3 They employ a private VLAN from the upstream providers to deliver traffic to NaWas locations. D8 Upon receiving traffic, scrubbers determine which IPs addresses within the prefixes are under attack, by analyzing the destination of bogus traffic.

The incoming traffic is then scrubbed, and the legitimate traffic is then forwarded to its destination (outgoing traffic from clients *is not routed via NaWas*, but instead using the client's regular upstream provider).

DDoS metadata collection: NaWas collects metadata associated with the DDoS attacks it scrubs for research purposes only [21]. In this research, we analyze *only metadata* associated with each attack, which includes the IP address under attack, attack peak (Gbps), and attack duration (C3,C5which does not included peak packets per second, so we focus on the volume of attacks). Contractual and privacy obligations restrain us from collecting and sharing information about DDoS and targets (Section 6).

Attacks	1826
Targets (/32)	576
Prefixes Attacked	180
Autonomous Systems	65
Pool IPs (April 2019)	3,962,368

TABLE 1. DDOS ATTACKS FILTERED AT NAWAS– JULY 2017 – MAY 2019)

	Duration (min)	*Peak (Gbps)*
25%ile	10	0.4
Median	20	1.4
Average	57	3.9
90%ile	98	13.1
Max	7560	79.0

TABLE 2. NAWAS ATTACKS DURATION AND PEAK GPBS

3. View from the DDoS maelstrom

We start by presenting an overview of the attacks scrubbed at NaWas. Our dataset contains metadata associated with 22 months worth of data (July 2017 to May 2019), as shown in Table 1. In this period, there were 1826 attacks, an average 2.73 attacks/day. All attacks were conducted against IPv4 addresses. Even though NaWas supports IPv6, no attacks were observed in this period on its filtering, D7 as no clients requested IPv6 prefixes to be filtered.

These attacks targeted 576 distinct IPv4 addresses (some addresses were attacked multiple times), which were announced by 65 distinct Autonomous Systems (ASes) in 180 prefixes on the BGP table. Altogether, these 180 prefixes encompass ∼3.9 million IPv4 addresses, increasing the chances of shared infrastructure (*e.g.,*routers, lines) [22], and, as such, also become victims of collateral damage (Section 4), which is known to happen during DDoS [23].

Figure 1 shows all attacks we analyzed in this period. In this figure, each point denotes an individual attack, and x axis shows when it took place, and the y axis is peak traffic (Gbps) measured while filtering the attacks. We see that larger attacks ($>$ 30Gbps) occur more often from Sept. 2018. Figure 2 shows the attack size distribution: most attacks are under 5Gbps, and 90% are under 13.1 Gbps.

That may seem rather small compared to the terabit scale attacks covered by media outlets [10]. These attack sizes we show also corroborate the sizes measured at an Internet Exchange Point (IXP) and different ISPs by Kopp *et al.* [7], in which the authors average attacked peaked at 2.4Gbps, slightly lower than what we observe at NaWas.

In regards to duration, we also see in Table 2 that most attacks did not last very long, having an average of 57 minutes. Moreover, 90% of attacks lasted less than two hours. This duration, however, corresponds to how long the *scrubbing service was active*, and, as we show in Section 5, the attack may have been active for hours before the scrubbing service was requested (operational experience from engineers at NaWas confirms that attacks typically stop shortly after the scrubber is active,

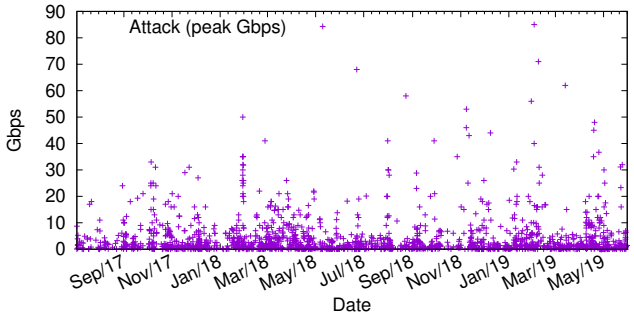

Figure 1. Attacks and Peak

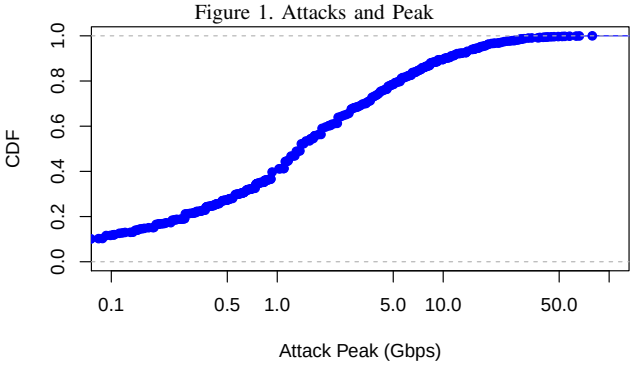

Figure 2. CDF of Attack Rates (peak)

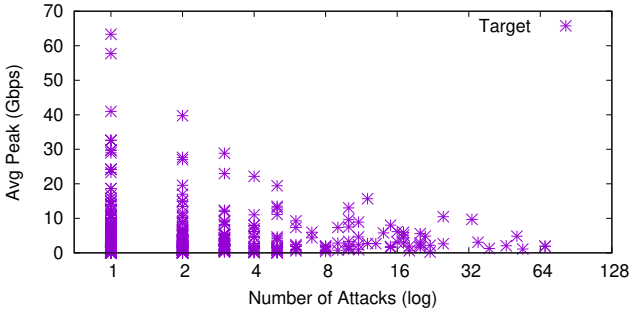

Figure 3. Attacks per IPv4 target and Mean Peak (Gpbs)

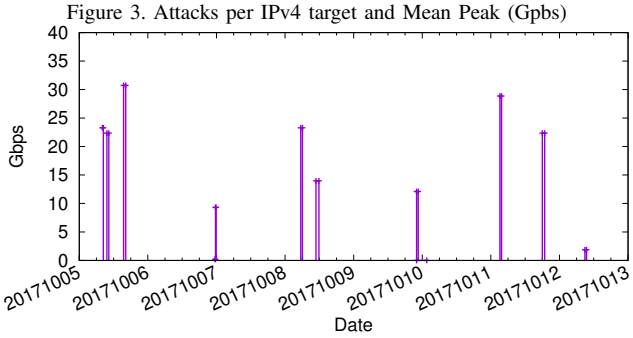

Figure 4. Twelve DDoS attacks against one IPv4 address in a week time

supposedly because their attacks become less effective, undermining the attacker's goals).

With relation to target distribution, we see that the same IP address, on average, received 3.2 DDoS attacks over the 22 month period (multiple attacks against the same IP are not that unusual — the Root DNS servers, for example, suffered two attacks 20h apart on Nov/Dec 2015 [24]). Figure 3 shows the number of attacks per targeted IPv4 address scrubbed at NaWas, and the average peak traffic over all combined attacks per target. We see that single attacks range from few up to more than 60Gpbs, while more often targeted IPs have a lower average peak in terms of attack size.

Figure 4 shows 12 DDoS attacks that were conducted against a single IP address over one week. These attacks ranged from 2 to 31 Gbps in size, and each of them lasted between ten minutes to one hour.

Altogether, these findings demonstrate a large diversity in terms of attack firepower and target distribution observed a NaWas, where most attacks are shorted lived *after the scrubbing service starts* and are under few Gbps – enough to disrupt many websites. We can speculate that other scrubber services should also have a similarly diverse attack distribution.

4. Collateral Damage

There have been several infamous cases of collateral damage caused by DDoS attacks. In 2015, when the secure email provider ProtonMail was attacked, it knocked down a datacenter, hurting all the other remaining clients [25]. In 2015, on an attack against the Root DNS servers, some of the Netherlands' `.nl` ccTLD authoritative servers experienced reachability issues due to the Root's attack [23]. In Oct. 2016, a larger attack was directed at Dyn, a provider of DNS service for many second-level domains [9], and there were reports of intermittent failure of prominent websites including "Twitter, Netflix, Spotify, Airbnb, Reddit, Etsy, SoundCloud and The New York Times" [26]. Attacks against Amazon's DNS network in 2019 disrupted part of its storage services [12].

Collateral damage occurs because services share infrastructure with others, such as links, servers, and routers [27], [22]. When one service is attacked, services that share parts of the same infrastructure are likely to suffer, amplifying the impact of DDoS attacks.

In this section, we estimate the shared infrastructure of the 576 targeted IP addresses shown in Table 1 by looking into what domain names were hosted under the same IP addresses *during* their respective attack time frame. By being hosted on the same addresses, it is more likely they share similar routers/links and servers.

To determine domains that shared the same IP addresses with the 576, we look up all 2LDs that had a matching DNS A record [28] on *on each attack date*, for each IP address. We look these domains in various top-level domains (TLDs) using OpenIntel [29], a research project that crawls series of DNS zones daily.

Table 3 shows the number of 2LD under different zones that were hosted on the 576 addresses which traffic was scrubbed by NaWas. We see that, altogether, the *affected* domain name space is *quadratically* larger than number of IP addresses itself: 330k domains where affected, given they shared the 576 IP addresses under attack — IP^2. These domains, even though many of them were not the primary target, are likely to suffer from these attacks altogether, given they are more likely to share parts of the infrastructure.

We also see that most domain names are on the `.nl` TLD, which can be explained by the fact that most of the NaWas members are from the the Netherlands. For the `.nl` zone, we see that 242k unique domains (4.1% of `.nl` zone [30]) were affected by these attacks.

DNS Zone	Second-level domains	IPs
.nl	242,355	226
.com	72,180	178
.net	5,220	100
.org	5,314	94
Others	6,541	98
Total	331,610	576

TABLE 3. COLLATERAL DAMAGE TO SECOND-LEVEL DOMAINS ON IPS UNDER ATTACK. OTHERS ARE: .AT, .CA, .DK, .FI, .RU, .SE, .US, XN--P1AI COMBINED.

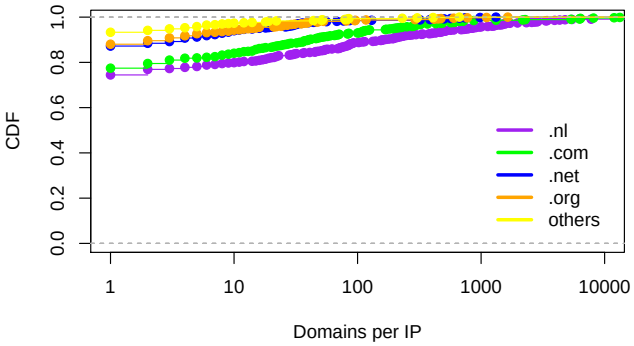

Figure 5. CDF of domains per IP address per zone

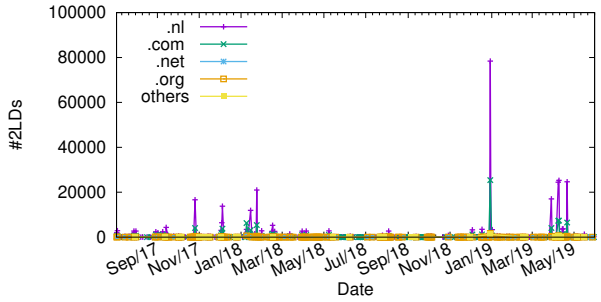

Figure 6. Time series of affected second-level domains per TLD

Figure 5 shows the CDF of domain names per IP address for each zone we evaluated. We see that most IPs host only one domain for all zones, but some IP addresses host almost 100,000 domain names (shared hosting [31]). Figure 6 shows a timeseries of 2LDs hosted on the 576 IP addresses targeted by the DDoS attacks. Comparing it to Figure 1, we see that the spikes do not necessarily overlap. The reason for that is that some shared hosting IP addresses may host thousands of domains, and a single DDoS attack may impact the reachability of all the websites hosted under these IPs.

Altogether, we see that the number of *victims* is, *at least*, actually the square of scrubbed addresses by NaWas. The potential figures could be higher if considering domains from other zones and deeper in the DNS hierarchy (third, fourth level domains), other DNS zones we did not have access to, as well as other parts of the shared infrastructure, as links and upstream providers. Still, our results present a conservative estimate of potential collateral damage, demonstrating that DDoSes nowadays rarely leaves a single victim.

5. Does DDoS leave its footprints on DNS?

Section 4 shows that ∼ 330k 2LDs were hosted in the 576 IP addresses that had traffic scrubbed by NaWas. In this section, we use authoritative DNS traffic to determine (i) if DDoS attacks leave traces on DNS traffic and, if so, if (ii) we can determine which domain names they actually targeted, among the pool of domains in the shared IP address, and (iii) the attackers strategy when carrying out DDoS attacks.

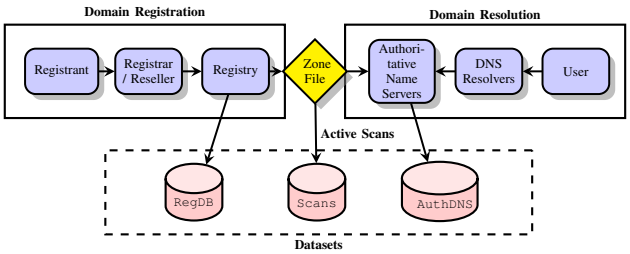

Figure 7. TLD operations: registration (left), domain resolution (right), and datasets.

To do that, we focus on the data available at SIDN, the registry of the .nl domain name, which, as shown in Table 3, contains most of domains that were hosted in these attacked IP addresses. Next, we summarize a DNS registry operation and then proceed to analyze its authoritative DNS traffic.

5.1. DNS registry operation and name resolution

(We describe the process below for .nl, but most TLDs have a similar process.) To register a domain name under .nl, typically involves three entities: a *registrant*, *registrar* (or reseller), and *registry*. The registrant (a user) requests a registrar to register an available domain name at the registry (SIDN). The registrar (*e.g.,*GoDaddy) then executes this request on behalf of the user, once payment and other checks are executed (left part of Figure 7).

Domains are registered for a period of one year. After being registered, domains name are inserted into the *Zone File* (Figure 7) that contains the list of all domains under .nl, and their respective DNS records. These Zone Files are used as input on the *authoritative name servers*, which are type of DNS server that know the "contents of the zone from local knowledge" [32] to answer queries about .nl domain names.

Domain name *resolution*, in turn, it is executed by another type of DNS server — a DNS *resolvers* — which, on behalf of users (left part of Figure 7), attempts to resolve a domain name into its IP address or other specific types of DNS records. These IP addresses are published as A/AAAA records [28] in DNS. Note, however, that not all queries from clients reach the authoritative server: caching on DNS resolvers [33], [34] is used to eliminate frequently issued queries, improving response times and reducing overall traffic from resolvers to authoritative servers.

5.2. Attack Detection using DNS traffic

Do DDoS attacks leave footprints on DNS traffic? To determine that, we focus on D9 passive DNS traffic collected at the *authoritative* sever side (AuthDNS dataset in Figure 7) from the .nl zone, the ccTLD of the Netherlands. Given their position in the DNS hierarchy, authoritative server traffic provides a centralized view to

the DNS traffic to the DNS zone they are authoritative for, given they are the ones that *answer* DNS resolvers.

Our hypothesis is that DDoS attacks change temporal query patterns for domains under attack, by inflating both the number of queries and number of resolvers given its possible large number of attacking IP addresses (we have used the same heuristic do detect phishing attacks in a previous study [35]). For example, if a certain domain is attacked by a large botnet, this domain is more likely to receive a large set of queries from typically unusual locations, given botnets tend to be globally spread.

This hypothesis is far from fail proof: (i) DDoS attacks from few sources cannot be detected this way, and (ii) caching at resolver's limits the number of queries we see arriving at authoritative servers [34]. D11 Moreover, (iii) DDoS attacks that target IP addresses directly without the use of domain names cannot be detected by this method and (iv), in our case, we have access to traffic from only one TLD (`.nl`).

To investigate DDoS leaves traces on DNS, we proceed with the following steps. first, we create tuples $[IP\text{-}Date]$, which consists of the target IP addresses scrubbed by NaWas (Table 1) at each respective attack date. Then, for each tuple, we create a list of all `.nl` domain names hosted at the same IP address ($[IP\text{-}Date] = [d_1, d_2 ... d_n]$), on the same date, using historical records from OpenIntel (as in Section 4). Then, for each individual domain name d_n in each tuple, we extract the following metrics obtained from the `.nl` authoritative DNS traffic:

- $Q_{attackDay}$: number of queries on the day of the attack.
- $Q_{weekBefore}$: average daily queries on the week before the attack day.
- `Ratio` = $Q_{attackDay}/Q_{weekBefore}$.

In this metrics, `Ratio` shows the increase in query volume on the attack date, compared to the average daily queries on the week before the attack, for a domain d. We choose one week value given Internet traffic typically follows diurnal, weekly patterns [36]. To obtain these metrics, we use ENTRADA [37], our open-source data streaming warehouse for authoritative DNS traffic that stores `.nl` authoritative traffic and has been also in use by several other TLDs.

These steps led to 549 $[IP\text{-}Date]$ tuples, which, in turn, covered the 242,355 unique `.nl` domain names, as shown in Table 4. Figure 8 shows a scatter plot of the number of queries on the attack day ($Q_{attackDay}$, y axis) and the average on the week before ($Q_{weekBefore}$, x axis). In this figure, each point corresponds to a domain name d.

As can be seen, the majority of domains *did not* experience a significant increase in number of DNS queries. There are two main reasons for that. First, maybe they were not attacked (and other domains from other zones hosted in the same IP were). A second possibility is that they may have been victims of attacks that do not require DNS resolution of these domains, such as amplification/reflection attacks. For example, a large botnet can send queries with IP address of the targeted domain to open resolvers, which will then answer queries to the IP address of the targeted domain.

Metric	Value
$[IP\text{-}Date]$ groups	549
normal	514
after filtering	35
Total domains	242355
normal	114561
after filtering	127794
suspicious	74

TABLE 4. RESULTS FROM `.NL` AUTHORITATIVE TRAFFIC ANALYSIS

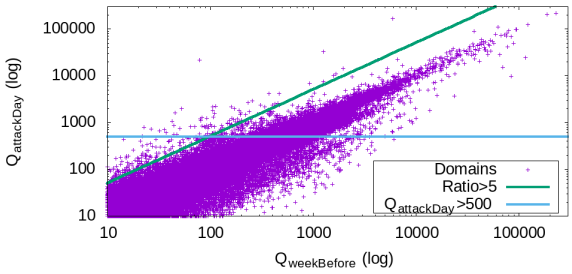

Figure 8. Domain names and number of queries on the attack day (y axis) and average daily week before it (x axis).

However, we can see in the same figure that *some* domains indeed have a large number of DNS queries. To single out these domains among the 330k, we apply a empirically chosen threshold: we consider a domain d as "suspicious" if it had a `Ratio`>5 (horizontal line in Figure 8) and $Q_{attackDay}$ >500 (vertical line). This step led to 74 domains being classified as suspicious (shown above both lines in Figure 8). They belong to 35 $[IP\text{-}Date]$ tuples, which collectively hosted 127794 unique `.nl` domains.

We proceed to further analyze the 74 domain names and its respective 35 $[IP-Date]$ tuples. We classify these tuples into four categories, as shown in Table 5. First, we classify if the IP address under attack was *dedicated* or *shared* hosting. Dedicated hosting is the case in which one IP hosts only one domain. Shared hosting, in turn, is when one IP address host multiple domains. As can be seen, we see only one dedicated hosting attack.

We distinguish shared hosting attacks in terms of number of targeted domains, which we estimate based on the number of domains with significant traffic increase. We see that 18 attacks against different IP addresses can be mapped into attacks that targeted a single domain name, while 15 of these attacks targeted *more* than one domain in the pool — which we derive from the number of domains in the pool with significant query volume increase. Next we show examples of DDoS attacks from each category.

DDoSes	$[IP-Date]$	Susp. Domains	Collateral
Total	35	74	127794
Dedicated hosting	1	1	0
Shared hosting	34	73	127794
one target	18	18	53447
multiple targets	15	57	74347

TABLE 5. DDOS ATTACKS CLASSIFICATION BASED ON `.NL` DNS DATA

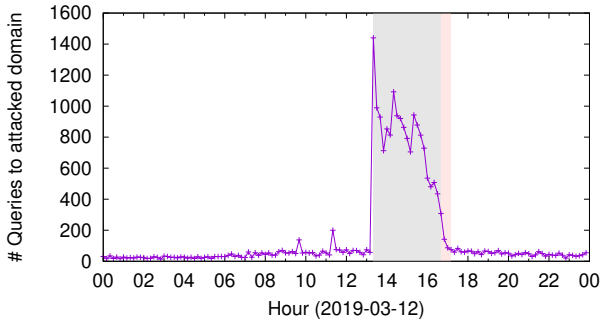

Figure 9. Timeseries of incoming queries to `.nl` for dedicated hosting domain. Pink area show area which scrubbing service was active; gray area shows when attacked started being noticed on DNS

	Shared-1 target		Dedicated	
	$Q_{attackDay}$	$Q_{weekBefore}$	$Q_{attackDay}$	$Q_{weekBefore}$
queries	163312	5955.28	22849	4225.28
resolvers	23774	1642.14	8960	1967.28
ASes	3396	272.71	1396	288.42
countries	150	58.28	101	49.14

TABLE 6. DDoS EVIDENCE FROM DNS

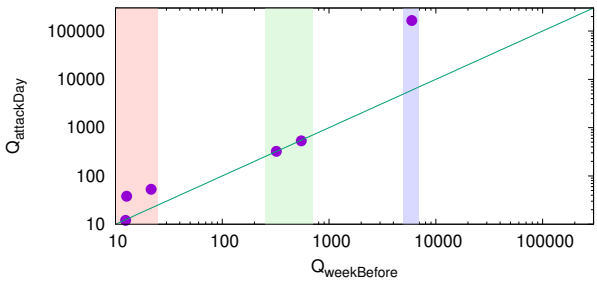

Figure 10. $[IP\text{-}Date]$ shared hosting 6 `.nl` domains and number of queries with one targeted domain

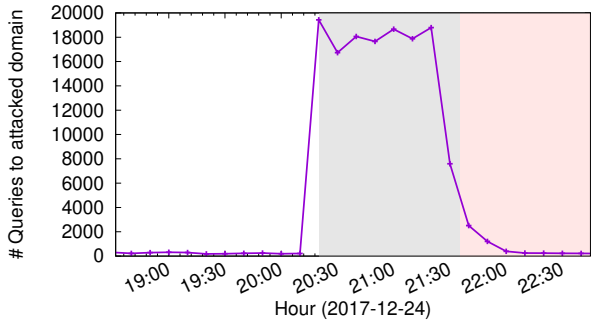

Figure 11. Timeseries of queries to `.nl` for single targeted domain of share hosting from Figure 10.

5.3. Classifying attacks based on DNS

We start with the only *dedicated hosting* attack we observed in Table 5. This attack took place on 2019-03-12, and peaked at 5.18Gbps. Figure 9 shows a time series of queries to this domain at the `.nl` authoritative DNS servers, on an hourly basis. Analyzing this figure, we see the number of queries experiencing a significant increase, from less than 50 queries per hour to more than 1000 per hour – an unusual DNS traffic pattern from this domain.

Accordingly to metadata obtained from NaWas, this attacked lasted for 30 minutes. This, however, refers to the time in which the scrubber service was active, which we overlay as a shaded pink area in Figure 9. We see, however, that the queries increase lasted *for almost four hours*, and possibly the attack itself. Why such a difference the DNS queries timeseries traffic and the scrubbing period from NaWas?

The reasons for that is that in the case of NaWas, members decide if and *when* to use the scrubber services — and it may take members some time between the start of the attack, detection and requesting traffic to be scrubbed — shown as the gray area in Figure 9. The deployment of the scrubbing service can be done an automatic or manual way. In the manual way, a NaWas member detects and manually announces the attack prefix(es) to NaWas using BGP (Section 2). In the automatic way, DDoS detection systems (such as based on IPFIX/Netflow [38]) are employed to detect attacks.D4 If they move beyond predetermined thresholds, they trigger, automatically, more specific route announcements to NaWas. A2 For the case of Figure 9, this filtering was *manually* activated – explaining the longer gray area in the figure.

The increase in number of queries is also followed by an increase in terms of their origins, as can be seen in the *Dedicated* column in Table 6. We see more resolvers, ASes, and countries being active on the attack day, compared to the average of the week before. We hypothesize that such growth can be due to use of bots that are distributed across the globe to carry out such attacks and need to reach the `.nl` authoritative servers to resolve the targeted domain name.

Attacks on shared-hosting: next, we show attacks on shared hosting domains. We first show an example with a single targeted domain (Table 6). We show the query patterns of an attack performed on 2017-12-24, which peaked at 7.7 Gbps. There were 6 `.nl` domain names hosted at this IP address on the attack day, and we show their query volumes before and during the attack day in Figure 10. The three left-most domains had few queries either before or during the attack day (< 100 daily), so we disregard it, given these number likely does not qualify as a DDoS attack. The second group, in the middle, had essentially the same number of queries on the week previously to the attack (avg) and the attack day — so they lie on the $y = x$ line and show no anomaly. The last group, right most, consists of one domain only, which went from average $\sim 6k$ daily queries on the week before the attack to ~165k queries on the attack day. This domain, therefore, is likely to be the one targeted during the DDoS. This particular domain hosts an online store of specialized equipment.

Figure 11 shows a timeseries with the number of queries for this single targeted domain. Similarly to Figure 9, we see a significant increase in term of queries, and one order of magnitude increase in terms of resolvers and ASes, as can also be seen in Table 6 (Shared-1 target column). This attack is also likely to have started at least 1 hour *before* the scrubbing service was activated (gray area in Figure 11).

Given this attack is on a shared-host IP, it showcases a likely case in which domains that shared this IP address suffered collateral damage.

Shared hosting with multiple targets: the last category of DDoS attacks is the one in which shared hosting

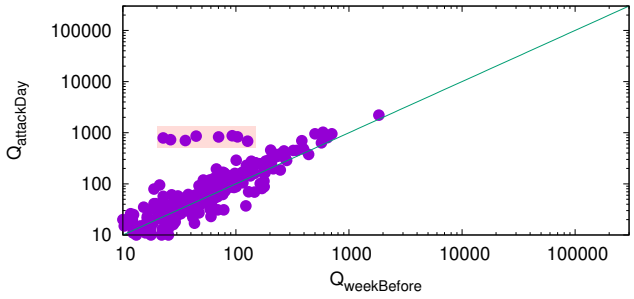

Figure 12. $[IP\text{-}Date]$ group hosting 8 among `.nl` domains 319 suspicious domains

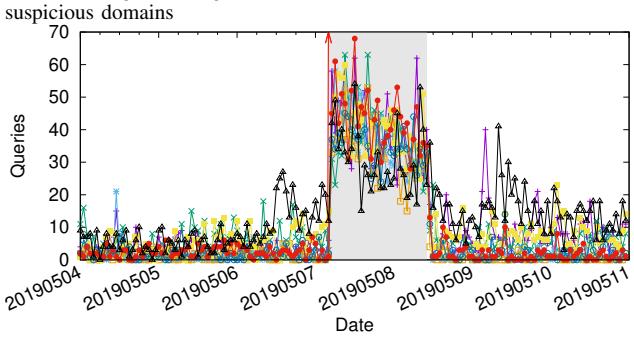

Figure 13. Timeseries of queries to 8 targeted `.nl` domains

services attacked, but multiple domain names are attacked instead of one.

We illustrate this category of attacks with a DDoS that took place on 2019-05-07, peaking at 1.4 Gbps. On this date, this IP hosted 319 `.nl` domain names. By analyzing their query patterns, we see that 8 of them experienced a similar increase in traffic on the attack time, as can be seen in Figure 12 — the domains in the small shaded box.

We wonder if this is part of the diversion strategy, in which random domains known to the attackers that are host on the same IP address are in fact attacked — while the intended targeted domain name suffers collateral damage. To investigate that, we look into the query time-series per hour of these 8 domains, which can be seen in Figure 13. We see that the 8 domains have synchronized time series spikes, even though they are from completely different businesses. Moreover, analyzed which resolvers queried these domains during the attack period (gray area) and it turned out that 63 resolvers from 10 ASes queried all eight domains, which is very unlikely.

If these 8 attacked domains were to diverge attention from the real target, which one is then the actual target? To identify that, we look at each domain from this shared pool (Figure 12). The most queried domain — the point at the far right — in is in fact a gaming website, and gaming websites are frequent victims of DDoS attacks (*e.g.,* [39]).

In this attack, however, the scrubbing services were used for one minute (arrow up), which coincides exactly when the DNS traffic started to increase, as shown in Figure 13. This is different from previous results – showing there has been little delay between the attack start and filtering. The reasons for that is that this particular member has an automated detection and scrubbing service request.

Takeway: Out of 549 $[IP - Date]$ groups analyzed, we found that 35 (6.3%) of them left fingerprints on DNS traffic on one ccTLD (`.nl`), which in turn had 74 suspicious domains. Even though the number is not that large, it showcases the potential use of analysis of DNS traffic at upper levels to detect DDoS attacks. Diversion may be an strategy used by attackers for carefully deceive defenses. Finally, once the scrubber service is activated, attacks tend to fade away quickly, which we can see also on DNS.

6. Privacy considerations

This work triangulates datasets from three different sources, which, each of them, have their own privacy and ethical considerations. The metadata datasets provided by NaWas follows an agreement between NaWas and its members, which conforms to both EU and the Netherlands legislation, as well as the EU General Data Protection Regulation (GDPR) [40]. NaWas makes its Privacy Statement publicly available [41].

For the authoritative traffic datasets from `.nl` used in Section 5, SIDN has developed a publicly available data privacy framework [42] that conforms to both EU and Dutch [42], [43] legislation. This framework has been implemented, including a privacy board that oversees SIDN Labs' research. For the OpenIntel project and datasets, we refer the reader to their primary publication [29].

7. Related Work

DDoS attacks have been under strong research interest over the last years. To our knowledge, this is the first work to analyze attacks on a scrubbing service for a period of almost two years, and that estimated the collateral damage of such attacks by analyzing historical domain name data. Moreover, we show how DDoS leaves footprints that on higher levels of DNS hierarchy, by triangulating attacks metadata with authoritative DNS traffic.

With regards to scrubbing services, Jonker *et al.* [13] have evaluated the adoption of DNS-based DDoS protection services for domain names under `.com`, `.org` and `.net`, by analyzing their respective DNS records. In a follow-up study, Jonker *et al.* [44] analyze two years of data collected at a large network telescope and from a amplification honeypot network (we focus on a scrubber data instead). They map the targeted addresses using DNS with OpenIntel, as we do in Section 4. They find that only 9% of their attacked IP addresses hosting a website. Our data, however, shows quite the opposite behavior: the quadratic number of websites with regards to the number of target of IP addresses. This, however, can be due to the nature of the datasets. Kopp *et al.* [7], in turn, have analyzed ISP and IXP traffic and show similar figures with regards DDoS attack sizes as we show in this paper.

BGP blackholing has also been explored by two main studies. Jonker *et al.* [45] found that, in the wild, 85% of the BGP blackholing events they observed happen within 10min from the start of the attacks. Similar to us, they also map which domain names were hosted under DNS zones hosted by OpenIntel, and found that there were 228 million. Differently from ours, their DDoS data is from BGP communities propagated by blackholing collected at RIPE RIS [46]. Ours, however, comes from NaWas, and does not rely upon public BGP communities. Given attacks filtered by NaWas rely upon private peering sessions, they are not visible in public BGP route collectors

such as RIPE RIS. Last, Nawrocki *et al.* [47], in turn, has analyzed BGP blackholing practices at a large European IXP (so do not rely upon public BGP route collectors), and found similar results as [45]. Given they have this centralized vantage point, they can characterize the attacks in much more detail.

In relation to collateral damage, Moura *et al.* [23] show evidence of collateral damage due to DDoS. While some of the Root server letters under attack, the `.nl` ccTLD experienced reachability issues on sites that were nearby anycast sites the attacked letters.

8. Conclusion

Cheap, easy, and relatively popular: that is how DDoS are seeing from the point of view of attackers. To fend off such attacks, an entire industry has emerged, and scrubbing services are among the most popular solutions. To shed light into the operations of scrubbers, we present the first longitudinal study of one non-commercial scrubbing service provider (NaWas) that has been operating over the last several years.

We show that most attacks do not last very long once they start to be scrubbed, likely due to making the attack less effective and, in this way, disrupting the goals of the attacker. With regards to attack sizes, we see that they not as big as the terabit scales witnessed over the last few years. Still, even gigabit-level DDoS are likely to overwhelm webservers and disturb some inter-domain links.

We triangulate the DDoS attacks metadata with two other datasets, which allows to estimate the collateral damage incurred by the attacks. For the analyzed datasets, we show that the number of victim second-level domains is exponential to the number of attacked IP addresses — a measure of the level of collateral damage of such attacks.

Last, we show that some DDoS attacks leave footprints DNS authoritative traffic: some domain names experience an increase in queries *during* the DDoS attack. We show that most attacks target shared-host domains, and some target multiple domain names hosted on the same IP address, likely to make it more filter such DDoS based on DNS traffic. The findings of this paper can be used to improve DDoS detection systems.

References

[1] N. Perlroth, "Tally of cyber extortion attacks on tech companies grows," New York Times Bits Blog, http://bits.blogs.nytimes.com/2014/06/19/tally-of-cyber-extortion-attacks-on-tech-companies-grows/, Jun. 2016.

[2] R. Chirgwin, "Linode: Back at last after ten days of hell," The Register, http://www.theregister.co.uk/2016/01/04/linode_back_at_last_after_ten_days_of_hell/, Jan. 2016. [Online]. Available: http://www.theregister.co.uk/2016/01/04/linode_back_at_last_after_ten_days_of_hell/

[3] Arbor Networks, "Rio Olympics Take the Gold for 540gb/sec Sustained DDoS Attacks! https://www.arbornetworks.com/blog/asert/rio-olympics-take-gold-540gbsec-sustained-ddos-attacks/," Aug. 2016.

[4] M. Karami and D. McCoy, "Understanding the Emerging Threat of DDoS-as-a-Service," in *Presented as part of the 6th USENIX Workshop on Large-Scale Exploits and Emergent Threats*. Washington, D.C.: USENIX, 2013.

[5] M. Karami, Y. Park, and D. McCoy, "Stress Testing the Booters: Understanding and Undermining the Business of DDoS Services," in *25th International Conference on World Wide Web (WWW)*, 2016, pp. 1033–1043.

[6] J. J. Santanna, R. van Rijswijk-Deij, R. Hofstede, A. Sperotto, M. Wierbosch, L. Zambenedetti Granville, and A. Pras, "Booters-An analysis of DDoS-as-a-service attacks," in *IFIP/IEEE Intl. Symposium on Integrated Network Management (IM)*. IEEE, May 2015, pp. 243–251.

[7] D. Kopp, M. Wichtlhuber, I. Poese, J. Santanna, O. Hohlfeld, and C. Dietzel, "DDoS Hide & Seek: On the Effectiveness of a Booter Services Takedown," in *Proceedings of the Internet Measurement Conference*, ser. IMC '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 65–72.

[8] 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," in *Proceedings of the 26th USENIX Security Symposium*. Vancouver, BC, Canada: USENIX, Aug. 2017, pp. 1093–1110. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-antonakakis.pdf

[9] S. Hilton, "Dyn analysis summary of Friday October 21 attack," Dyn blog https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/, Oct. 2016.

[10] S. Kottler, "February 28th DDoS Incident Report — Github Engineering," Mar. 2018, . https://githubengineering.com/ddos-incident-report/.

[11] MIT Technology Review, "The first DDoS attack was 20 years ago. This is what we've learned since." https://www.technologyreview.com/s/613331/the-first-ddos-attack-was-20-years-ago-this-is-what-weve-learned-since/, Apr. 2019.

[12] K. McCarthy, "Amazon is saying nothing about the DDoS attack that took down AWS, but others are," https://www.theregister.co.uk/2019/10/28/amazon_ddos_attack/, Oct. 2019.

[13] M. Jonker, A. Sperotto, R. van Rijswijk-Deij, R. Sadre, and A. Pras, "Measuring the Adoption of DDoS Protection Services," in *Proceedings of the 2016 Internet Measurement Conference*, ser. IMC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 279–285. [Online]. Available: https://doi.org/10.1145/2987443.2987487

[14] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," IETF, RFC 4271, Jan. 2006. [Online]. Available: http://tools.ietf.org/rfc/rfc4271.txt

[15] Neustar, "Ddos prevention & protection faqs," https://www.home.neustar/resources/faqs/ddos-faqs, 2019.

[16] Akamai, "The State of the Internet," Akamai. Available online at: http://www.akamai.com/stateoftheinternet/, Tech. Rep., 2019.

[17] Nationale Beheersorganisatie Internet Providers, "NBIP DDoS data report," Tech. Rep., 2019. [Online]. Available: https://www.nbip.nl/wp-content/uploads/2019/10/NBIP-DDoS-data-rapport-first-half-year-2019.pdf

[18] NaWas, "National Scrubbing Center against DDoS attacks," Feb 2019. [Online]. Available: https://www.nbip.nl/en/nawas/

[19] NBIP, "*Nationale Beheersorganisatie Internet Providers*," Feb 2019. [Online]. Available: https://www.nbip.nl/en/

[20] Microsoft, "Azure DDoS Protection," Feb 2019. [Online]. Available: https://azure.microsoft.com/en-us/pricing/details/ddos-protection/

[21] NBIP, "NBIP DDoS data report," July 2019. [Online]. Available: https://www.nbip.nl/wp-content/uploads/2019/10/NBIP-DDoS-data-rapport-first-half-year-2019.pdf

[22] M. Allman, "Comments on dns robustness," in *Proceedings of the Internet Measurement Conference 2018*, ser. IMC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 84–90.

[23] G. C. M. Moura, R. de O. Schmidt, J. Heidemann, W. B. de Vries, M. Müller, L. Wei, and C. Hesselman, "Anycast vs. DDoS: Evaluating the November 2015 root DNS event," in *Proceedings of the ACM Internet Measurement Conference*, Nov. 2016. [Online]. Available: https://www.isi.edu/%7ejohnh/PAPERS/Moura16b.html

[24] Root Server Operators, "Events of 2015-11-30," Nov. 2015, http://root-servers.org/news/events-of-20151130.txt.

[25] ProtonMail, "Guide to DDoS protection," https://protonmail.com/blog/ddos-protection-guide/, Dec. 2015.

[26] N. Perlroth, "Hackers used new weapons to disrupt major websites across U.S." *New York Times*, p. A1, Oct. 22 2016. [Online]. Available: http://www.nytimes.com/2016/10/22/business/internet-problems-attack.html

[27] J. Arkko, "Centralised Architectures in Internet Infrastructure," in *IETF Internet Draft*, Nov. 2019. [Online]. Available: https://tools.ietf.org/html/draft-arkko-arch-infrastructure-centralisation-00

[28] P. Mockapetris, "Domain names - concepts and facilities," IETF, RFC 1034, Nov. 1987. [Online]. Available: http://tools.ietf.org/rfc/rfc1034.txt

[29] R. van Rijswijk-Deij, 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, June 2016.

[30] SIDN Labs, ".nl stats and data," Jun. 2019, http://stats.sidnlabs.nl.

[31] S. Tajalizadehkhoob, T. Van Goethem, M. Korczyński, A. Noroozian, R. Böhme, T. Moore, W. Joosen, and M. van Eeten, "Herding Vulnerable Cats: A Statistical Approach to Disentangle Joint Responsibility for Web Security in Shared Hosting," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '17. New York, NY, USA: ACM, 2017, pp. 553–567. [Online]. Available: http://doi.acm.org/10.1145/3133956.3133971

[32] P. Hoffman, A. Sullivan, and K. Fujiwara, "DNS Terminology," IETF, RFC 8499, Nov. 2018. [Online]. Available: http://tools.ietf.org/rfc/rfc8499.txt

[33] G. C. M. Moura, J. Heidemann, M. Müller, R. de O. Schmidt, and M. Davids, "When the dike breaks: Dissecting DNS defenses during DDoS," in *Proceedings of the ACM Internet Measurement Conference*, Oct. 2018. [Online]. Available: https://www.isi.edu/%7ejohnh/PAPERS/Moura18b.html

[34] G. C. M. Moura, J. Heidemann, R. de O. Schmidt, and W. Hardaker, "Cache me if you can: Effects of DNS Time-to-Live," in *Proceedings of the ACM Internet Measurement Conference*, Oct. 2019.

[35] G. C. M. Moura, M. Müller, M. Wullink, and C. Hesselman, "ndews: A new domains early warning system for tlds," in *NOMS 2016 - 2016 IEEE/IFIP Network Operations and Management Symposium*, April 2016, pp. 1061–1066.

[36] L. Quan, J. Heidemann, and Y. Pradkin, "When the Internet Sleeps: Correlating Diurnal Networks with External Factors," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, ser. IMC '14. New York, NY, USA: ACM, 2014, pp. 87–100.

[37] M. Wullink, G. C. Moura, M. Müller, and C. Hesselman, "Entrada: A high-performance network traffic data streaming warehouse," in *Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP*. IEEE, Apr. 2016, pp. 913–918.

[38] B. Claise, B. Trammell, and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information," IETF, RFC 7011, Sep. 2013. [Online]. Available: http://tools.ietf.org/rfc/rfc7011.txt

[39] R. Harb, "Ubisoft sues handful of gamers for ddosing rainbow six: Siege," Jan. 20 2020. [Online]. Available: https://www.theregister.co.uk/2020/01/20/ubisoft_sues_gamers_rainbow_six_ddos_claim/

[40] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," April 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj

[41] Nationale Beheersorganisatie Internet Providers, "Privacy Statement NBIP," 2020. [Online]. Available: https://www.nbip.nl/privacy-statement/

[42] C. Hesselman, J. Jansen, M. Wullink, K. Vink, and M. Simon, "A privacy framework for DNS big data applications," Tech. Rep., 2014. [Online]. Available: https://www.sidnlabs.nl/downloads/yBW6hBoaSZe4m6GJc_0b7w/2211058ab6330c7f3788141ea19d3db7/SIDN_Labs_Privacyraamwerk_Position_Paper_V1.4_ENG.pdf

[43] C. Hesselman, G. C. M. Moura, R. d. O. Schmidt, and C. Toet, "Increasing DNS Security and Stability through a Control Plane for Top-Level Domain Operators," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 197–203, January 2017.

[44] M. Jonker, A. King, J. Krupp, C. Rossow, A. Sperotto, and A. Dainotti, "Millions of targets under attack: A macroscopic characterization of the dos ecosystem," in *Proceedings of the 2017 Internet Measurement Conference*, ser. IMC '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 100–113. [Online]. Available: https://doi.org/10.1145/3131365.3131383

[45] M. Jonker, A. Pras, A. Dainotti, and A. Sperotto, "A first joint look at dos attacks and bgp blackholing in the wild," in *Proceedings of the Internet Measurement Conference 2018*, ser. IMC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 457–463. [Online]. Available: https://doi.org/10.1145/3278532.3278571

[46] RIPE Network Coordination Centre, "RIPE - Routing Information Service (RIS)," https://https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris, 2020.

[47] M. Nawrocki, J. Blendin, C. Dietzel, T. C. Schmidt, and M. Wählisch, "Down the Black Hole: Dismantling Operational Practices of BGP Blackholing at IXPs," in *Proceedings of the Internet Measurement Conference*, ser. IMC '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 435–448. [Online]. Available: https://doi.org/10.1145/3355369.3355593