Evaluating DNS Resiliency with Truncation, Fragmentation and DoTCP Fallback

Pratyush Dikshit*, Mike Kosek†, Nils Faulhaber†, Jayasree Sengupta*, and Vaibhav Bajpai*
*CISPA Helmholtz Center for Information Security, Germany
[pratyush.dikshit | jayasree.sengupta | bajpai]@cispa.de
†Technical University of Munich, Germany
[kosek | nils.faulhaber]@in.tum.de

***Abstract*—Since its introduction in 1987, the DNS has become one of the Core components of the Internet. While it was designed to work with both TCP and UDP, DNS-over-UDP (DoUDP) has become the default option due to its low overhead. As new Resource Records were introduced, the sizes of DNS responses increased considerably. This expansion of message body has led to truncation and IP fragmentation more often in recent years where large UDP responses make DNS an easy vector for amplifying denial-of-service attacks which can reduce the resiliency of DNS services. This paper investigates the resiliency and usage of DoTCP and DoUDP over IPv4 and IPv6 for 10 widely used public DNS resolvers. In three experiments, these aspects are investigated from the Edge and from the Core of the Internet to represent the communication of the resolvers with DNS clients and authoritative name server. Overall, more than 14M individual measurements performed from 2500 RIPE Atlas Probes have been analyzed, highlighting that most resolvers show similar resiliency for both DoTCP and DoUDP. Yet, 3 out of 10 resolvers mainly announce very large EDNS(0) buffer sizes both from the Edge as well as from the Core, which potentially causes fragmentation. In reaction to large response sizes from authoritative name servers, we find that resolvers do not fall back to the usage of DoTCP in many cases, bearing the risk of fragmented responses. As the message sizes in the DNS are expected to grow further, this problem will become more urgent in the future.**

***Index Terms*—EDNS(0), DNS-over-TCP, Resiliency**

I. Introduction

The Domain Name System (DNS) which is responsible for the resolution of hostnames to IP addresses, has become one of the most widely used components on the Internet. Hostnames (domain names) are organized in a tree structure that is hierarchically separated into *zones*. The resolution of domain names is realized by different components such as stub resolvers, recursive resolvers, and authoritative Name Servers (NSes). While authoritative NSes are responsible for the authoritative mapping of domains in a zone to their IP addresses, stub, and recursive resolvers cache and deliver such information from the NSes to the clients via a DNS request (RFC 1034). DNS communication supports both major transport protocols on the Internet, namely the Transmission Control Protocol (TCP) (RFC 793) and the User Datagram Protocol (UDP) (RFC 768). Due to its comparably low overhead, UDP has become the default transport protocol for DNS. The UDP message body is restricted to 512 bytes (RFC 1035). However, the increase in deployment of DNS Security (DNSSEC) and IPv6 (RFC 7766) has resulted in larger message sizes, thereby leading to two important developments in the protocol. Firstly, DNS over TCP (DoTCP) was declared to be mandatory for hosts (RFC 5966) as it enables larger message body by default. Secondly, Extension Mechanisms for DNS (EDNS) were introduced to augment the capabilities of the DNS protocol in terms of message size expansion (RFC 2671). However, using too large UDP buffer sizes can cause IP fragmentation in certain networks, thereby reducing resiliency in DNS communication [1]. To avoid fragmentation, the DNS Flag Day, 2020[1], an association of DNS software maintainers and service providers, recommended using a default buffer size of 1232 bytes. DoTCP is a useful measure against fragmentation and can increase DNS resiliency by allowing fallback in such scenarios. Resolvers should also avoid fragmentation by using the recommended default EDNS(0) buffer size of 1232 bytes. To this end, our paper puts forward two **goals:** *a)* to evaluate DoTCP support (both over IPv4 and IPv6) and its usage across several DNS resolvers, and *b)* to investigate which buffer sizes are currently used in DNS traffic around the globe.

In pursuit of these goals, we evaluate the behavior of the resolvers from two different vantage points. Firstly, DoTCP adoption and EDNS(0) configuration are analyzed from the Edge where the interaction between recursive resolvers and DNS clients running on the RIPE Atlas probes is measured. To scope DNS requests to the *Edge* of the network, we perform DNS queries for a domain that is likely cached by all resolvers, unlike in previous studies [2]. Secondly, the interaction of recursive resolvers with authoritative NSes is further studied. To allow DNS requests to leave the Edge and move into the *Core* of the network, we provision dedicated NSes for a custom-crafted domain whose resolution is requested from the DNS resolvers. Using this methodology (see §III), we study failure rates and usage of DoTCP and DoUDP, as well as the usage of EDNS(0) buffer sizes both from the Edge and the Core that gives detailed insights into the potential resiliency of DNS communication on the Internet. We perform measurements over both IPv4 and IPv6 [3]. Our main **findings** (see §IV) are –

***Resiliency from the Edge**:* We observe that DoTCP (4.01%) tends to fail less often than DoUDP (6.3%) requests over IPv4.

[1]http://www.dnsflagday.net/2020/

ISBN 978-3-903176-57-7

Contrarily, in case of IPv6, we find higher failure rate over both the transport protocols (DoTCP 10%, DoUDP 9.61%). We also observe that several public DNS resolvers still lack adoption ($< 3.5\%$) of 1232B from the DNS Flag Day recommendation.

Resiliency from the Core: We find that DoTCP requests over IPv4 exhibit failure rates of 9.09% on public resolvers against higher failure rate of 11.53% over IPv6. Moreover, communication between resolvers and the authoritative NSes utilize EDNS(0) buffer size of 512 bytes less preferably (IPv4 0.24%, IPv6 0.13%) compared to the buffer sizes advertised to the RIPE Atlas probes (IPv4 27.41%, IPv6 26.04%). All DNS resolvers use EDNS(0) in most of the cases ($>$ 99.84%). We also see other DNS options such as Cookie (4.80% IPv4, 7.91% IPv6) and EDNS Client Subnet (ECS) (1.81% IPv4, 1.49% IPv6) advertised by the public resolvers, while Google mostly uses ECS (14.24% IPv4, 12.53% IPv6).

DoTCP Usage Rates: We observe that when 2KB responses are received from the NSes, all resolvers that mainly use canonical (see §IV scenarios, use TCP in their last request for >95% of the cases. In situations where 4KB responses are received, we observe that almost all resolvers use TCP in the vast majority of measurements over both IP versions (>98%).

We detail limitations and future work in §V, and conclude the paper in §VI.

II. Related Work

A. Fragmentation

With the increased message sizes, DNS queries can exceed the MTU of many networks. Giovane *et al.* in [4] analyzes the fragmentation rates of DNS queries to the *.nl* top-level domain showing that less than 10k of 2.2B observed DNS responses by authoritative NSes are fragmented. Although fragmentation is in general fairly rare in DNS communication, the consequences can have negative effects on resiliency and connectivity of Internet-applications (RFC 8900). Herzberg and Shulman in [5] presented an attack allowing to spoof Resource Records (RRs) in fragmented DNS responses by which attackers can hijack domains or nameservers under certain conditions. Following a similar procedure, Shulman and Waidner in [6] showed the opportunity to predict the source port used by the client. This potentially exposes the DNS user to several other types of attacks. Koolhaas *et al.* in [1] analyzed the behavior of different EDNS(0) buffer sizes in 2020. It was shown that the likelihood of a failing DNS query increases with growing buffer sizes. For a size of 1500 bytes, the default MTU of Ethernet which causes fragmentation of most of the DNS messages of the DNS queries to stub resolvers failed for 18.92% over IPv4, with 26.16% over IPv6 (RFC 2464).

As countermeasures, in 2017, Cao *et al.* in [7] presented an "Encoding scheme against DNS Cache Poisoning Attacks". Berger *et al.* in [8] presented a way of detecting DNS cache poisoning attacks in 2019. Herzberg and Shulman in [6] recommend keeping the indicated buffer size less or equal to 1500 bytes. As a consequence, Weaver et al. summarize a list of recommendations to stakeholders in the DNS ecosystem.

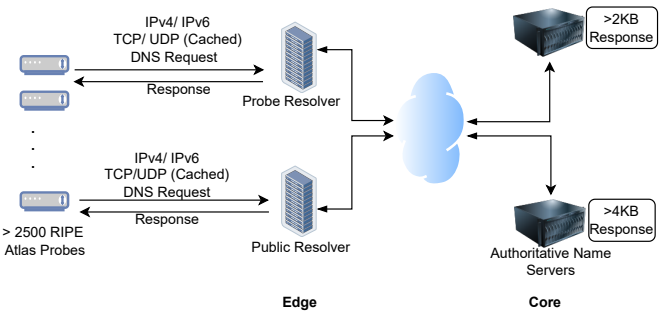

Fig. 1: *2527 RIPE Atlas Probes communicate the DNS requests with the Edge (Probe and Public Resolvers) and with the Core (authoritative NSes) using IPv4 and IPv6. Cached DNS responses are sent by the Edge, while uncached DNS responses (2KB and 4KB) are sent by the Core, represented by the Authoritative NSes.*

The recommendations were adopted at the DNS Flag Day, 2020 claiming that "defaults in the DNS software should reflect the minimum safe size which is 1232 bytes".

B. DNS Resiliency and EDNS(0) Adoption

Kosek *et al.* in [2] conducted the first study comparing DNS failure rates based on the underlying transport protocols using RIPE Atlas for ten public resolvers and probe resolvers using uncached DNS queries. 8% of the queries over both UDP and TCP failed, with a very high DoTCP failure rate of 75.0% for probe resolvers. To take it forward, and to get a preferably broad and unbiased comparison of the different DNS resolvers over the particular underlying protocols, we perform DNS queries to cached domain on each resolver (google.com).

Van den Broek *et al.* in [9] analyzed more than 8 million DNS queries to an authoritative NS of which 75% used EDNS(0). Additionally, it was observed that 36% announced a buffer size higher than 1232 bytes possibly causing fragmentation. Based on the analysis of 164 billion queries to authoritative NSes, Moura *et al.* in [4] stated that many resolvers "announce either small (512 bytes) or large (4096 bytes) EDNS(0) buffer sizes, both leading to more truncation, and increasing the chances of fragmentation/packets being lost on the network". This paper continues the measurements of Moura *et al.* in [4] to observe changes in the advertised buffer sizes of resolvers. Additionally, maintaining an own authoritative NS gives the opportunity to further analyze the EDNS(0) configuration of the resolvers communicating with NSes. Besides observing the buffer sizes that are used in the communication with our NS, the usage of other EDNS(0) options, like TCP usage, is evaluated in this paper.

III. Methodology

In A, we detail the RIPE Atlas probe selection process, followed by an analysis process of different DNS resolvers from the edge in B. Finally, in C, we explain the method to analyze the behavior of the resolvers while communicating with the authoritative name servers, that is, with the core.

A. Probe Selection

The measurements in this paper are conducted using the RIPE Atlas measurement network. To avoid potential load issues occurring in the first two probe versions [10], we choose only probes of version 3 or 4 that are hosted with a hometag [11]. The probes need to be capable of using IPv4, IPv6, or both protocol versions. To conduct as many individual measurements as possible, we are interested in starting them from all RIPE Atlas probes that have the given attributes. A scan of all RIPE Atlas probes on 20th December 2021 shows the availability of 2527 probes with the desired attributes. All of them are IPv4 capable, 1137 can be used for measurements over IPv6. The density of probes can be observed as- 70% in Europe, 18% in North America, 6% in Asia, 3% in Oceania, 1% in Africa and 1% in South America, distributed over 671 different Autonomous Systems (ASs). Before running the actual measurement series consisting of the analysis of DNS resolver from the Edge and from the Core as shown in Figure 1, DoTCP usage, EDNS(0) configuration and DoTCP fallback, we evaluated 4343 probe resolvers found for the 2443 probes participating in the measurements yielding an average of 1.78 resolvers per probe.

B. From the Edge

To analyze the behavior of the different DNS resolvers from the Edge, RIPE Atlas measurements targeting the public DNS resolvers are configured programmatically. The DNS measurements are carried out over IPv4 and IPv6 alongside both the transport protocols TCP and UDP. A record for the most frequently used domain on the Internet, i.e. *google.com* is requested which is most likely cached by all resolvers. This should avoid any recursive resolution of the requested domain or unexpected errors (e.g., due to an unknown domain) and scopes during communication between client programs and recursive resolvers. The RIPE Atlas probes participating in the measurement audit and other information such as error messages, UDP buffer sizes are all advertised by each resolver.

C. From the Core

This evaluation allows us to analyze the behavior of the resolvers while interacting with authoritative NS (see Figure1). The circumstances for this experiment are slightly different from the one observed from the Edge, as uncached domain names are used here and the requested domain is maintained by authoritative NS under our control. The DNS configuration used by the resolvers is further analyzed where two custom authoritative NSes are deployed to retrieve and store information about incoming DNS requests.

Our customized authoritative NS encodes incoming DNS requests alongside additional information, such as the transport protocol and the IP address of the requester. For later analysis, the encoded requests can be sent back to the client or stored locally on the servers. Observing the EDNS section of the requests that reach the authoritative NS provides an unfiltered view of the resolvers' EDNS configuration. This also includes the potential usage of options such as *Cookie* or *Client Subnet*. The NSes are developed based on COREDNS[2], a flexible DNS server that allows to chain custom plugins. For the observations from the Core, we developed two custom plugins tailored for the different experiments. Both plugins are mainly designed to receive and process incoming requests so that they can be analyzed later. Using the first plugin[3], the request-data is encoded by the servers and returned as a TXT record. The second plugin[4] developed for our experiment constructs large responses and writes incoming DNS requests to a *.csv-file* along with the domain queried and its timestamp.

DoTCP Usage and EDNS(0) Configuration: To evaluate which transport protocols and EDNS configuration the DNS resolvers use in the Core, measurements are performed over the RIPE Atlas network. The target domain is one that is managed by our NSes. To avoid cached responses by the resolvers, the domains are made unique by prepending the probe id and timestamp to each DNS request. The IP addresses of the resolvers provide information about the distribution of the respective public resolver in terms of physical (by continent) and network location (by AS). The information on the DoTCP usage in the Core can also be obtained by observing the transport protocols used.

DoTCP Fallback: This measurement from the Core aims at observing the DoTCP fallback behavior of the public DNS resolvers. For this, the authoritative NSes return large responses, containing 72 AAAA records (resulting in a ¿2KB response) for one server, and 145 AAAA record (resulting in a ¿4KB response) for the other one. AAAA records are used as they have larger size than A records and thereby help in assembling large responses. Note that we cover different RR-types, A, AAAA, and TXT in each of the three measurements. As observed from the previous experiment that resolvers request both NSes with equal frequency, it is expected that roughly 50% of the requests receive responses of 4KB and 2KB, respectively. This way, we can investigate the reaction to both response sizes within the same experiment. As mentioned earlier, such large response body cannot be handled by UDP due to fragmentation-related issues. Hence, resolvers are expected to fallback to the usage of DoTCP. To investigate the reaction of the resolvers to large responses, we observe which of them still utilize UDP and whether they switch to DoTCP at all. This could help to evaluate the potential resiliency risks. The resolvers are expected to send more than one request to the authoritative NS (e.g., one over UDP followed by the fallback request over TCP). All incoming requests at the servers need to be taken into consideration to understand their behavior in detail. To allow a clear mapping of the incoming requests and the RIPE Atlas measurements, the domains queried from each probe are made unique using the aforementioned prepending technique.

IV. Results

We evaluate the results of the measurement from the Edge concerning failure rates and EDNS(0) buffer sizes. Afterwards,

[2]https://github.com/coredns/coredns

[3]https://github.com/nilsfaulhaber/echo-plugin-for-coredns

[4]https://github.com/nilsfaulhaber/fallbackmonitor-plugin-for-coredns

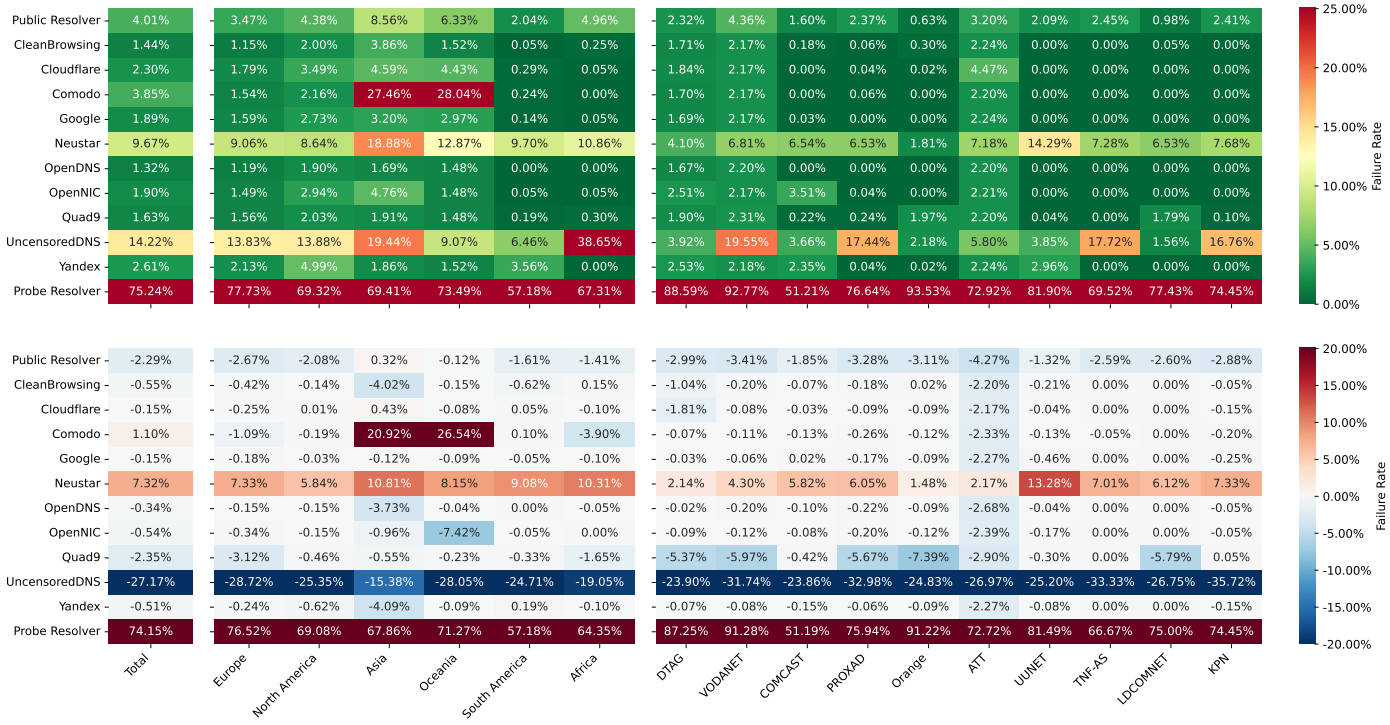

Fig. 2: *Failure rates observed from the Edge over IPv4. The upper part represents the DoTCP failure rates of all resolvers in total and per continent and AS. The lower part reflects the difference between the DoTCP and the DoUDP failure rates for a particular pairing (a negative value hence indicates a higher DoUDP failure rate). 'Public Resolver' summarizes the observations of all resolvers that are not probe resolvers.*

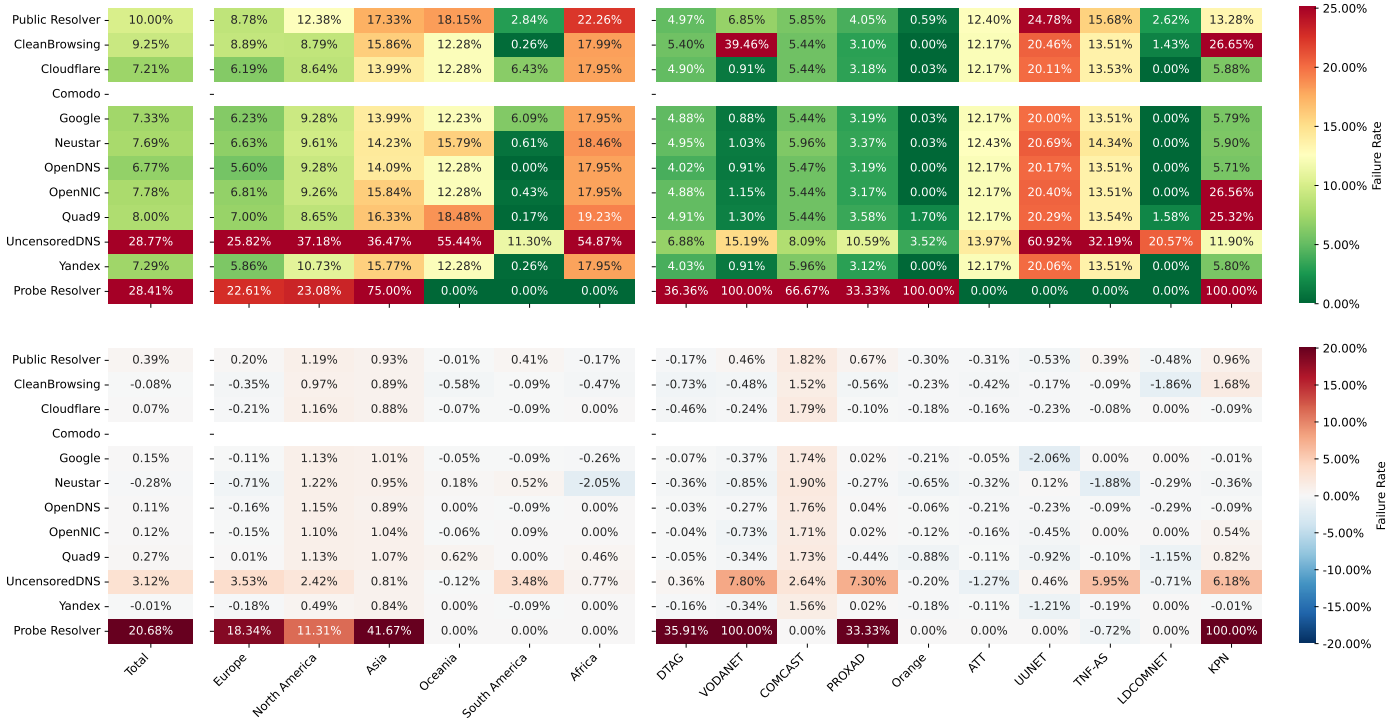

Fig. 3: *Failure rates observed from the Edge over IPv6. The upper part presents the failure rates over DoTCP, the lower one the difference between DoTCP and DoUDP failure rates. White cells indicate that there is no data for the given pairing.*

		512	1232	4096	none	other
CleanBrowsing	*IPv4*	**97.04%**	0.63%	1.46%	0.57%	0.30%
	IPv6	**99.41%**	0.11%	0.48%	0.01%	0.00%
Cloudflare	*IPv4*	0.20%	**97.43%**	1.45%	0.53%	0.40%
	IPv6	0.11%	**99.44%**	0.44%	0.01%	0.00%
Comodo	*IPv4*	0.18%	0.64%	**98.30%**	0.57%	0.30%
	IPv6	-	-	-	-	-
Google	*IPv4*	**96.82%**	0.78%	1.47%	0.58%	0.34%
	IPv6	**99.22%**	0.10%	0.67%	0.00%	0.01%
Neustar	*IPv4*	0.18%	0.64%	**98.32%**	0.56%	0.30%
	IPv6	0.10%	0.10%	**99.79%**	0.00%	0.00%
OpenDNS	*IPv4*	0.18%	0.63%	**98.20%**	0.57%	0.43%
	IPv6	0.10%	0.11%	**99.79%**	0.00%	0.00%
OpenNIC	*IPv4*	0.18%	**97.53%**	1.42%	0.56%	0.30%
	IPv6	0.11%	**99.43%**	0.47%	0.00%	0.00%
Quad9	*IPv4*	**19.15%**	**55.47%**	1.48%	**23.55%**	0.35%
	IPv6	**20.98%**	**62.09%**	0.47%	**16.46%**	0.00%
UncensoredDNS	*IPv4*	0.30%	**95.87%**	2.39%	0.96%	0.49%
	IPv6	0.13%	**99.29%**	0.57%	0.01%	0.00%
Yandex	*IPv4*	0.19%	0.64%	**98.04%**	0.75%	0.39%
	IPv6	0.11%	0.10%	**99.79%**	0.00%	0.00%
Overall	*IPv4*	24.97%	36.12%	35.30%	3.24%	0.36%
	IPv6	24.86%	38.81%	34.46%	1.87%	0.00%

TABLE I: *EDNS(0) Buffer Sizes announced to the RIPE probes by the resolvers observed from the Edge. Buffer sizes which are not equal to 512, 1232 or 4096 bytes are summarized in the column other. If EDNS is not used at all this is reflected in the column none.*

we analyse two experiments from the Core. The first measurement focuses on the EDNS(0) configuration of the resolvers, the second one analyzes DoTCP fallback.

A. From the Edge

This section analyzes the failure rates observed from the Edge to investigate the resiliency of DNS with respect to the protocol and IP version used. To evaluate how the individual resolvers adopt the DNS Flag Day recommendation of 1232 bytes, the EDNS(0) buffer sizes are analyzed.

According to Kosek *et al.* [2], measurements with no DNS response at the probe are defined as failed [2]. In IPv4, for public resolvers, DoTCP (4.01%) requests tend to fail less often than DoUDP requests (6.3%) indicating a slightly higher resiliency of DoTCP, see Figure 2. The failure rates of the requests to probe resolvers show a very different picture as DoTCP exceeds DoUDP by 74.15%. The failures of DoUDP are exclusively caused by Timeouts (5000ms). DoTCP requests to public resolvers as well mainly fail due to Timeouts (42.75%). However, other failure reasons such as READ-ERROR (33.91%), CONNECT-ERROR (23.24%) and TCP-READ (0.09%) also occur. Failures of DoTCP requests to probe resolvers are predominantly caused by bad address (99.17%). Comparing the results, probe resolvers exhibit very high DoTCP failure rates over all continents.

In case of IPv6, for public resolvers, we find lower resiliency over both transport protocols (DoTCP 10%, DoUDP 9.61%). Most public resolvers exhibit failure rates between 6.77% and 9.25%, see Figure 3. Uncensored DNS shows by far the worst DoTCP and DoUDP resiliency where few exceptions shows very similar failure rates between the individual resolvers when comparing them by continent and AS. This would best explain the very similar (and fairly high) failure rates of probes from the same AS for different resolvers and a small difference between DoTCP and DoUDP failure rates.

To analyze the adoption of the DNS Flag Day 2020 recommendations by the public resolvers from the Edge, we evaluate the EDNS(0) buffer sizes which the individual resolvers announce to the RIPE Atlas probes. Table I summarizes the buffer sizes that have been observed in the UDP measurements. As for all resolvers except Quad9 the difference in the percentages of the announced buffer sizes between IPv4 and IPv6 are fairly low ($\leq$3.5%). The buffer sizes advertised by Cloudflare, OpenNIC, UncensoredDNS (highlighted in orange), and Quad9 (55.47% IPv4, 62.09% IPv6) conform to the DNS Flag Day 2020 recommendation of a default buffer size of 1232B in most cases. Neustar, Comodo, OpenDNS and Yandex (blue) mainly use 4096 bytes. In 23.55% of the Quad9 DNS responses over IPv4, EDNS(0) is not used at all leaving clients to the default DoUDP message size limit of 512 bytes (IPv6 16.46%). This first view from the Edge shows that several public DNS resolvers still lack adoption to the DNS Flag Day 2020 recommendations. To see whether this also holds for the communication with authoritative NSes, we conducted another experiment from the Core.

B. From the Core

Figure 4 shows that DoTCP requests over IPv4 from the Core exhibits higher failure rates for Public resolvers (9.09%) than in the previous measurement series (4.01%). Over IPv4, 59.91% of the failures are caused by Timeouts and fewer CONNECT-ERRORS are observed (10.63% from the Core). In general, CleanBrowsing, Cloudflare, Google, OpenDNS, OpenNIC, and Yandex show fairly high resiliency (DoTCP failure rates 1.34%-2.62%, DoUDP 1.84%-3.42%). The failure rates by continent and AS are characterized by high variations between the continent/resolver and AS/resolver pairings. E.g., the failure rates of DoTCP requests to Comodo from Asia (34.44%) and the ASs DTAG (26.25%), VODANET (28.28%), and PROXAD (29.04%) are considerably higher than from other AS. The same holds for Quad9 from North America (24.72%), Asia (41.60%), Oceania (40.32%), Africa (20.02%), and from COMCAST (27.40%) and AT&T (18.77%).

When IPv6 is used, the general failure rates observed for all public resolvers for this measurement series (11.53%) are slightly higher. More READ-errors (19.67%) and less Timeouts (29.99%) occur compared to the measurements from the Edge. DoTCP requests to Quad9 (15.10%) and UncensoredDNS (39.12%) fail with the highest percentage. Figure 5 shows that from all continents, the DoTCP failure rates for the public resolvers is surged in comparison to that of explained in Figure 2. However, more than 88% of the measurements to the public resolvers over both IP versions and transport protocols receive a valid DNS response. This shows that there are no significant problems with our NSes and yields enough data to reliably analyze the resolvers' DoTCP usage and EDNS(0).

To understand the usage of different EDNS(0) configurations by the resolvers communicating with the authoritative NSes for uncached domains, we firstly present an overview of the distribution of resolvers in Table II which shows their distribution over the 10 most widely used ASs in all successful

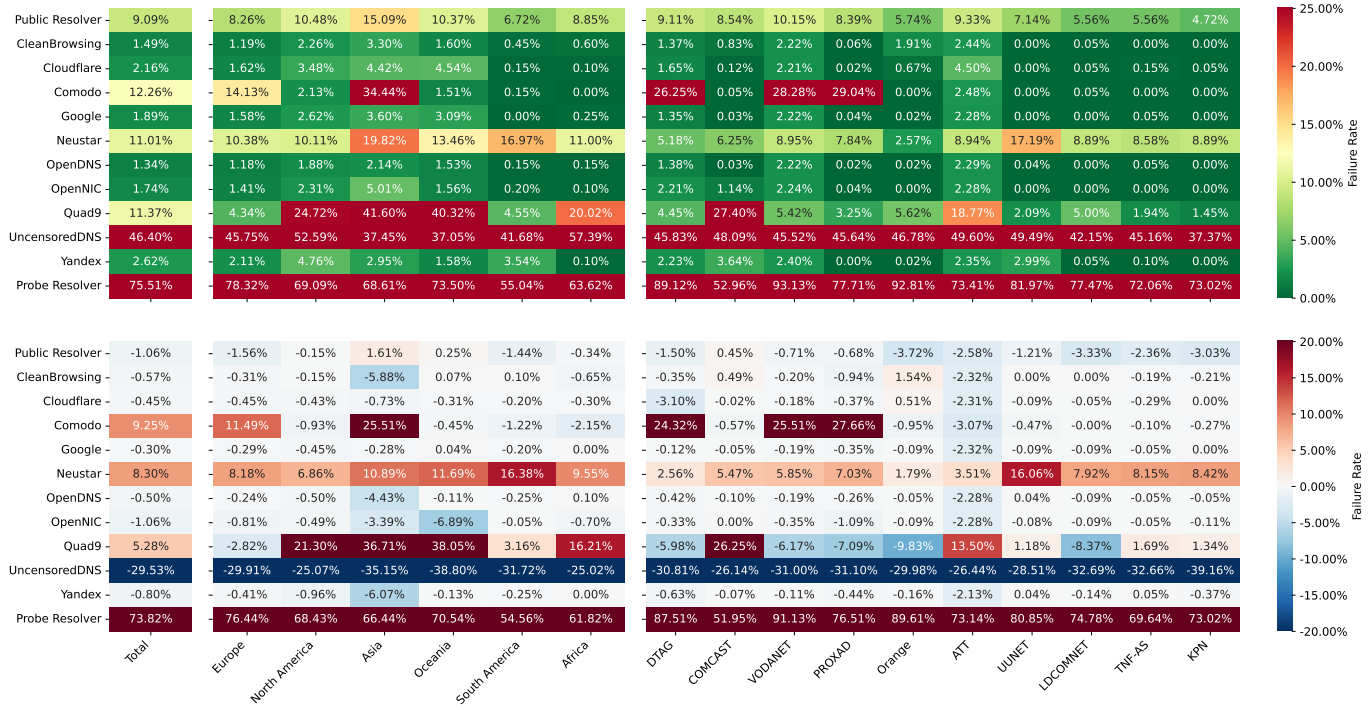

Fig. 4: *Failure rates observed from the Core over IPv4. The upper part represents the DoTCP failure rates of all resolvers in total and per continent and AS. The lower part reflects the difference between the DoTCP and the DoUDP failure rates for a particular pairing (a negative value hints a higher DoUDP failure rate). Public Resolver summarizes the observations of all resolvers that are not Probe resolvers.*

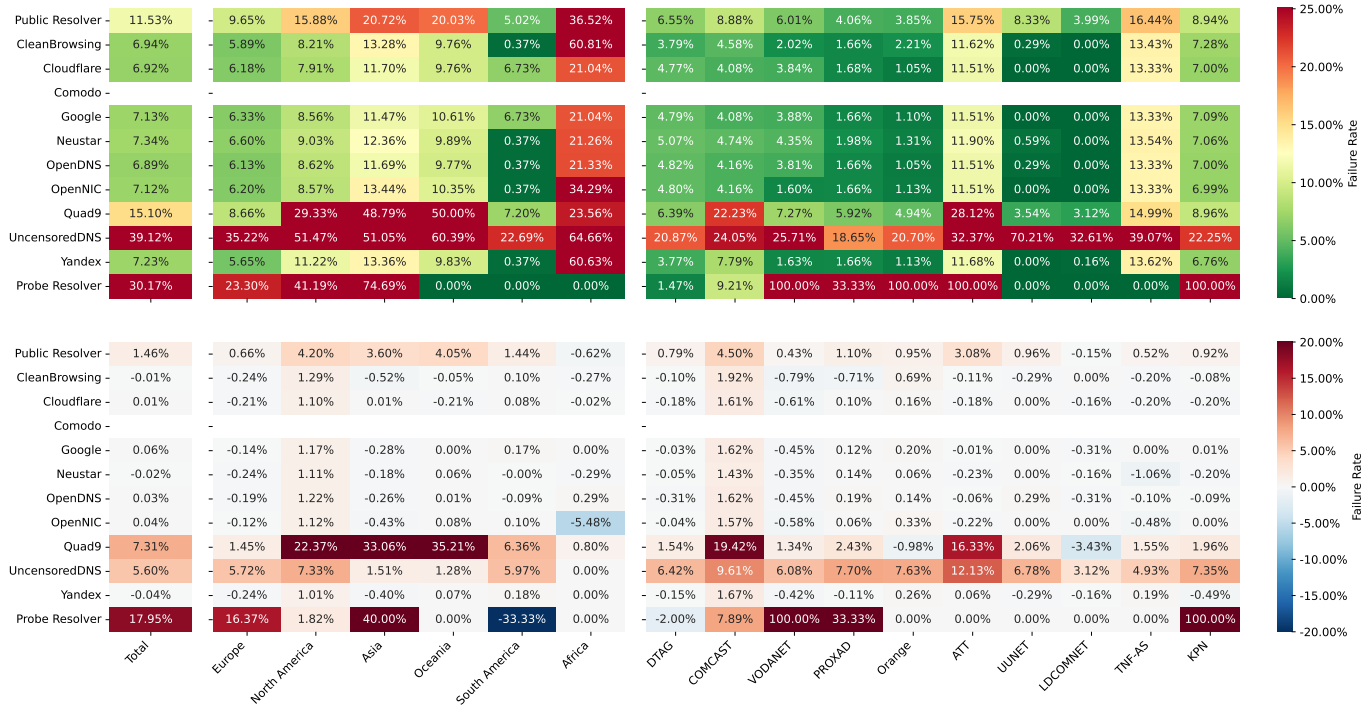

Fig. 5: *Failure rates observed from the Core over IPv6. The upper part presents the failure rates over DoTCP, the lower one is the difference between DoTCP and DoUDP failure rates. White cells indicate that there is no data for the given pairing.*

		Choopa	Cogent	Packet	Cloudfl.	Google	Neust.	O-DNS	Myth.	W-net-1	FSKNET	Yandex	Other
CleanBrowsing	*IPv4*	**66.58%**	**28.54%**	-	0.64%	0.45%	-	0.05%	-	0.05%	-	-	3.69%
	IPv6	**59.07%**	**14.89%**	**25.26%**	0.07%	0.21%	-	-	-	-	-	-	0.49%
Cloudflare	*IPv4*	0.06%	0.04%	-	**98.04%**	0.46%	-	0.04%	-	0.05%	-	-	1.32%
	IPv6	-	-	-	**99.38%**	0.21%	-	-	-	-	-	-	0.41%
Comodo	*IPv4*	**95.63%**	0.04%	-	0.67%	0.48%	-	0.05%	-	0.06%	-	-	3.07%
	IPv6	-	-	-	-	-	-	-	-	-	-	-	-
Google	*IPv4*	0.06%	0.04%	-	0.64%	**97.86%**	-	0.05%	-	0.09%	-	-	1.28%
	IPv6	-	-	-	0.06%	**99.41%**	-	-	-	-	-	-	0.52%
Neustar	*IPv4*	0.06%	0.04%	-	0.63%	0.48%	**94.79%**	0.05%	-	0.06%	-	-	3.90%
	IPv6	-	-	-	0.06%	0.23%	**99.23%**	-	-	-	-	-	0.48%
OpenDNS	*IPv4*	0.06%	0.04%	-	0.59%	0.53%	-	**97.68%**	-	0.01%	-	-	1.10%
	IPv6	-	-	-	0.07%	0.22%	-	**99.30%**	-	-	-	-	0.41%
OpenNIC	*IPv4*	**87.92%**	0.03%	-	0.59%	0.45%	-	0.05%	**9.77%**	0.05%	-	-	1.12%
	IPv6	**65.21%**	-	-	0.06%	0.23%	-	-	**34.09%**	-	-	-	0.41%
Quad9	*IPv4*	0.06%	0.04%	-	0.70%	0.51%	-	0.05%	-	**82.34%**	-	-	**16.30%**
	IPv6	-	-	-	0.08%	0.22%	-	-	-	**75.57%**	-	-	**24.13%**
UncensoredDNS	*IPv4*	0.15%	0.09%	-	1.49%	1.14%	-	0.12%	-	0.14%	**94.07%**	-	2.81%
	IPv6	-	-	-	0.08%	0.34%	-	-	-	-	**98.98%**	-	0.60%
Yandex	*IPv4*	0.06%	0.04%	-	0.69%	0.56%	-	0.04%	-	0.05%	-	**97.43%**	1.13%
	IPv6	-	-	-	0.06%	0.22%	-	-	-	-	-	**99.30%**	0.41%
Overall	*IPv4*	18.87%	3.52%	-	12.35%	12.22%	10.98%	11.83%	1.19%	9.19%	4.54%	11.75%	3.57%
	IPv6	14.49%	1.72%	2.95%	11.50%	11.70%	11.44%	11.51%	3.97%	8.33%	7.85%	11.46%	3.09%

TABLE II: *Distribution of the resolvers communicating with the authoritative name servers for uncached domains used by the public resolvers over AS. The ten ASs that were used in most requests over each IP version are presented individually, all others are summarized in the column 'Other'.*

		512.0	1232.0	1400.0	1410.0	1452.0	4096.0	other
CB	*IPv4*	0.11	**98.24**	0.45	0.05	0.64	0.36	0.16
	IPv6	0.01	**99.47**	0.21	0.00	0.07	0.05	0.19
Cloudflare	*IPv4*	0.36	0.65	0.46	0.04	**98.04**	0.30	0.16
	IPv6	0.01	0.26	0.21	0.00	**99.38**	0.04	0.10
Comodo	*IPv4*	0.11	0.70	0.48	0.05	0.67	**95.21**	2.78
	IPv6	-	-	-	-	-	-	-
Google	*IPv4*	0.22	0.78	**97.86**	0.05	0.64	0.27	0.19
	IPv6	0.02	0.26	**99.41**	0.00	0.06	0.14	0.10
Neustar	*IPv4*	0.04	0.70	0.48	0.05	0.63	**97.45**	0.66
	IPv6	0.02	0.31	0.23	0.00	0.06	**98.79**	0.60
OpenDNS	*IPv4*	0.08	0.61	0.53	**97.68**	0.59	0.32	0.19
	IPv6	0.01	0.26	0.22	**99.30**	0.07	0.04	0.10
OpenNIC	*IPv4*	0.06	**98.29**	0.45	0.05	0.59	0.37	0.18
	IPv6	0.01	**99.56**	0.23	0.00	0.06	0.05	0.10
Quad9	*IPv4*	0.07	**98.05**	0.51	0.05	0.70	0.40	0.21
	IPv6	0.01	**99.54**	0.22	0.00	0.08	0.04	0.11
U-DNS	*IPv4*	2.68	**93.15**	1.14	0.12	1.49	0.97	0.45
	IPv6	1.46	**97.91**	0.34	0.00	0.08	0.07	0.15
Yandex	*IPv4*	0.03	0.65	0.56	0.04	0.69	**92.86**	5.16
	IPv6	0.00	0.26	0.22	0.00	0.06	**94.31**	5.14
Overall	*IPv4*	0.24	39.74	12.22	11.83	12.34	22.78	0.85
	IPv6	0.13	42.09	11.70	11.51	11.49	22.33	0.75

TABLE III: *EDNS(0) buffer sizes announced to the authoritative NSes. Other buffer sizes and cases in EDNS that is not used are summarized in the column "other". NOTE: CB= CleanBrowsing;U-DNS= UncensoredDNS. All the values are in percentage (%).*

		EDNS	Cookie	ECS
CleanBrowsing	*IPv4*	99.93%	0.22%	0.10%
	IPv6	99.91%	0.05%	0.04%
Cloudflare	*IPv4*	99.94%	0.32%	0.10%
	IPv6	100.00%	0.05%	0.05%
Comodo	*IPv4*	98.10%	0.33%	0.11%
	IPv6	-	-	-
Google	*IPv4*	99.93%	0.31%	**14.23%**
	IPv6	100.00%	0.16%	**12.53%**
Neustar	*IPv4*	99.93%	0.23%	0.10%
	IPv6	99.93%	0.05%	0.04%
OpenDNS	*IPv4*	99.94%	0.22%	0.10%
	IPv6	100.00%	0.05%	0.04%
OpenNIC	*IPv4*	99.93%	0.22%	0.11%
	IPv6	100.00%	0.05%	0.05%
Quad9	*IPv4*	99.93%	0.24%	0.13%
	IPv6	100.00%	0.06%	0.03%
UncensoredDNS	*IPv4*	99.84%	**94.62%**	0.24%
	IPv6	100.00%	**99.06%**	0.06%
Yandex	*IPv4*	99.93%	0.22%	0.11%
	IPv6	100.00%	0.05%	0.04%
Overall	*IPv4*	99.93%	4.80%	1.81%
	IPv6	99.98%	7.91%	1.49%

TABLE IV: *EDNS options announced to the authoritative name servers*

requests for IPv4 and IPv6. As expected, most resolvers exhibit one preferred AS through which the vast majority of requests is resolved. Cloudflare, Google, Neustar, OpenDNS, and Yandex DNS use their own ASs in more than 94% of their domain name resolutions.

The fact that the public resolvers use DoUDP independent of the transport protocol used by the clients further emphasizes the usage of proper EDNS(0) buffer sizes in the Core where we observe that the sizes of 1400, 1410, and 1452 bytes are used more often. We therefore extend the number of buffer sizes particularly displayed in Table III. All resolvers exhibit one preferred buffer size which is advertised to the NSes in more than 90% of the cases.

Table IV exhaustively lists all options that the DNS resolvers use in the communication with the name servers. Again, there are small differences between the usage rates over IPv4 and IPv6. All DNS resolvers use EDNS(0) in most of the cases ($\geq$ 99.84%). Cookie (4.80% IPv4, 7.91% IPv6) and EDNS Client Subnet (ECS) (1.81% IPv4, 1.49% IPv6) are furthermore the only options that are advertised by the public resolvers where Google mostly uses ECS (14.24% IPv4, 12.53% IPv6). The other ones send client subnet information in $<$ 0.24% of their requests. RFC 7871 [12] states that NSes must include an ECS with matching parameters in their response. Google furthermore claims that "if name servers do not implement

		Canonical	Non-Canonical	Single UDP
CleanBrowsing	*IPv4*	**99.42%**	0.47%	0.11%
	IPv6	**99.56%**	0.24%	0.21%
Cloudflare	*IPv4*	**95.36%**	**4.45%**	0.19%
	IPv6	**91.96%**	**7.70%**	0.34%
Comodo	*IPv4*	1.19%	**97.83%**	0.98%
	IPv6	-	-	-
Google	*IPv4*	**99.24%**	0.67%	0.10%
	IPv6	**99.44%**	0.26%	0.30%
Neustar	*IPv4*	1.31%	**97.61%**	1.08%
	IPv6	0.25%	**98.67%**	1.07%
OpenDNS	*IPv4*	**97.71%**	2.08%	0.21%
	IPv6	**99.13%**	0.62%	0.25%
OpenNIC	*IPv4*	**55.68%**	**35.60%**	**8.72%**
	IPv6	**35.20%**	**26.89%**	**37.91%**
Quad9	*IPv4*	**46.14%**	**53.67%**	0.19%
	IPv6	**46.28%**	**53.47%**	0.25%
UncensoredDNS	*IPv4*	**94.11%**	4.70%	1.19%
	IPv6	**99.52%**	0.20%	0.29%
Yandex	*IPv4*	1.19%	**97.79%**	1.02%
	IPv6	0.22%	**98.94%**	0.84%
All	*IPv4*	59.69%	39.82%	0.48%
	IPv6	69.05%	30.50%	0.45%

TABLE V: *Classification of the incoming sequences of DNS queries at the 2KB name server for each resolver.*

		Canonical	Non-Canonical	Single UDP
CleanBrowsing	*IPv4*	**98.65%**	1.27%	0.08%
	IPv6	**98.83%**	0.25%	0.92%
Cloudflare	*IPv4*	**91.76%**	**8.22%**	0.00%
	IPv6	**85.24%**	**14.37%**	0.36%
Comodo	*IPv4*	8.91%	**91.07%**	0.02%
	IPv6	-	-	-
Google	*IPv4*	**98.67%**	0.83%	0.51%
	IPv6	**98.82%**	0.19%	0.99%
Neustar	*IPv4*	3.13%	**96.86%**	0.01%
	IPv6	0.73%	**98.25%**	1.02%
OpenDNS	*IPv4*	**97.55%**	2.37%	0.08%
	IPv6	**98.38%**	0.79%	0.83%
OpenNIC	*IPv4*	**64.36%**	**35.42%**	**0.22%**
	IPv6	**38.34%**	**7.06%**	**54.60%**
Quad9	*IPv4*	**26.37%**	**73.60%**	0.03%
	IPv6	**31.23%**	**68.21%**	0.56%
UncensoredDNS	*IPv4*	**94.52%**	5.43%	0.05%
	IPv6	**99.48%**	0.11%	0.41%
Yandex	*IPv4*	3.34%	**96.56%**	0.10%
	IPv6	0.83%	**97.86%**	1.31%
All	*IPv4*	70.91%	28.95%	0.13%
	IPv6	75.43%	23.79%	0.78%

TABLE VI: *Classification of the incoming sequences of DNS queries at the 4KB name server for each resolver.*

ECS, Google public DNS may not send ECS[5] queries to it". This leads to the assumption that the ECS usage rates of Google could be much higher if the servers would correctly handle the requests. Nevertheless, we can observe that several Google's resolvers in the Core, uniquely identifiable by their IP addresses, send subnet information to our servers multiple times. The question of how Google's usage rate of the ECS option in the communication with NSes that correctly answer the requests remains open for research.

Overall 4,576,757 individual measurements are conducted with distinct domains, 3,241,545 over IPv4, 1,335,212 over IPv6. Combining RIPE Atlas measurements and the incoming requests at the NSes yields 12,328,029 individual results. Filtering these out based on unique domain names leaves us with 11,637,539 results. We furthermore observe that the NS returning 2KB responses receives more requests (5,642,439) than the one returning 4KB (2,395,455). Moreover, some domains are requested on both NSes (2,733,540 results). We analyse the classification of incoming requests as canonical and non-canonical according to Mao et. al. [13] followed by the analysis of DoTCP usage rates of the resolvers. This is to evaluate how the resolvers react to the large buffer sizes. Additionally, the scenario of a single incoming UDP request is introduced (this would not happen in the experiment by Mao *et al.*[13] as the TC-bit is set manually). As the focus of further analyses lies particularly on the reaction of the DNS resolvers to either of the response sizes- 2KB or 4KB, we decide to only evaluate the results that can directly be matched to one of the servers and thus one of the response sizes. We furthermore distinguish between the requests that were answered by the 2KB and by the 4KB NS. Table V shows the resolvers' usage of the different scenarios in communication with the 2KB NS. CleanBrowsing, Cloudflare, Google, OpenDNS, and UncensoredDNS send a UDP message followed by a TCP message in most of the cases. Referring to Table III, resolvers have advertised the usage of EDNS(0) buffer sizes $\leq$1452B, showing the expected fallback behavior.

The usage of the different scenarios observed from the NS returning 4KB responses is presented in Table VI. We see that the number of non-canonical scenarios used by Cloudflare is significantly higher. Quad9 shows more non-canonical sequences in reaction to the 4KB than to the 2KB responses. Note that we have seen in the previous experiment (Table II) that OpenNIC uses resolvers communicating with the Core from the AS Mythic more often IPv6 than over IPv4. This can be one reason for the resolvers' different behavior over the two protocol versions.

In order to evaluate the TCP usage, we investigated appearances of DoTCP requests in the sequence of queries reaching the NSes. In case of DoTCP usage rates of the resolvers when 2KB responses are received, all resolvers that mainly use canonical scenarios, use TCP in their last request. As per our findings, this also holds for Quad9 (99.69% IPv4, 99.70% IPv6). Neustar and OpenNIC use DoTCP for the resolution of significantly fewer domains. Neustar thereby shows high differences between the usage rates of TCP in general (73.52% IPv4, 72.17% IPv6). OpenNIC on the other hand again shows high varieties between IPv4 and IPv6. Yandex and Comodo rarely use DoTCP when they deal with responses of 2KB (1.58%-7.94% in general, 0.94%-3.36% in the last request). In case when 4KB responses are received, we observe that almost all resolvers use TCP in the vast majority of measurements over both IP versions ($\geq$98.67%). We can still see a non-negligible number of measurements for which TCP is not used at all by many resolvers (up to 1.33%). Fragmentation on this path from NS to the resolver is very likely in these scenarios. Over IPv6, OpenNIC uses TCP in less than half of

[5]https://developers.google.com/speed/public-dns/docs/ecs

the sequences that reach the name server (45.09%). Cloudflare (95.76% IPv4, 92.63% IPv6) and OpenNIC (94.72% IPv4, 41.72% IPv6) furthermore tend to use DoTCP in their last requests less often than the other resolvers.

V. Limitations and Future Work

Around 88% of the probes that participated in our measurements are located in North America and Europe. As such, the observations on DNS resiliency should not be seen as an estimation for DNS traffic everywhere in the world. When documenting measurement failures, the RIPE Atlas measurement network applies its custom error codes for the classification of the origin of the errors. Due to lack of documentation to this end, we are unable to perform further causal analysis of these failure cases. The EDNS(0) options are studied for cases where the different resolvers communicate with our custom authoritative NSes. As such, the usage numbers do not generally reflect the capabilities of the respective resolvers and their EDNS(0) options. The NSes have observed many different non-canonical sequences that the DNS resolvers use in reaction to the large response sizes. A further investigation of these scenarios would be necessary to understand the behavior of the different resolvers. For instance, whether and in which scenarios the resolvers adjust their announced EDNS(0) buffer sizes while receiving large responses.

While our study focussed on the unencrypted DNS protocols DoUDP and DoTCP, the recently standardized encrypted DNS protocol DNS-over-QUIC (DoQ) (RFC 9250)[14][15] does inherently solve fragmentation by means of the QUIC protocol (RFC 9000)[16] while also supporting increased DNS message sizes. However, DoQ adoption currently is scarce [17]; yet, DNS over QUIC is a promising candidate to supersede both DoUDP and DoTCP in the future, thereby warranting a detailed investigation when DoQ adoption rises.

VI. Conclusion

We conducted three different sets of measurements analyzing DoTCP resiliency and deployment both from the Edge and from the Core over IPv4 and IPv6. Additionally, the EDNS(0) configurations of known public resolvers were studied. Issuing more than 14M individual DNS requests using around 2500 globally distributed RIPE Atlas probes, we performed multiple experiments focusing on observations from the Edge as well as from the Core. For the examination of the DoTCP usage and EDNS configurations from the Core, two dedicated NSes were deployed that was able to collect information about incoming requests. While we find that most resolvers show similar resiliency for both DoTCP and DoUDP. 3 out of 10 resolvers mainly announce very large EDNS(0) buffer sizes both from the Edge as well as from the Core, which potentially causes fragmentation. In reaction to large response sizes from authoritative name servers, we find that resolvers do not fall back to the usage of DoTCP in many cases, bearing the risk of fragmented responses. As the message sizes in the DNS are expected to grow further, this problem will become more urgent in the future.

Acknowledgment

This work was supported by the Volkswagenstiftung Niedersächsisches Vorab (Funding No. ZN3695).

References

[1] Axel Koolhaas and Tjeerd Slokkker. Defragmenting DNS: Determining the Optimal Maximum UDP Response Size for DNS, 2020. [Last Accessed: 19.April.2023]: https://bit.ly/3Ag6Mck.

[2] Mike Kosek, Trinh Viet Doan, Simon Huber, and Vaibhav Bajpai. Measuring DNS over TCP in the Era of increasing DNS Response Sizes: A View from the Edge. *Computer Communication Review*, 2022.

[3] Vaibhav Bajpai and Jürgen Schönwälder. A Longitudinal View of Dual-Stacked Websites - Failures, Latency and Happy Eyeballs. *IEEE/ACM Trans. Netw.*, 27(2):577–590, 2019.

[4] Giovane C. M. Moura, Moritz Müller, Marco Davids, Maarten Wullink, and Cristian Hesselman. Fragmentation, Truncation, and Timeouts: Are Large DNS Messages Falling to Bits? In *Passive and Active Measurement Conference (PAM)*, 2021.

[5] Amir Herzberg and Haya Shulman. Fragmentation considered poisonous, or: One-domain-to-rule-them-all.org. In *IEEE Conference on Communications and Network Security (CNS) 2013*. IEEE, 2013.

[6] Haya Shulman and Michael Waidner. Fragmentation Considered Leaking: Port Inference for DNS Poisoning. In *Applied Cryptography and Network Security Conference, (ACNS)*. Springer, 2014.

[7] Jin Cao, Maode Ma, Xilei Wang, and Haochen Liu. A selective re-query case sensitive encoding scheme against DNS cache poisoning attacks. *Wireless Personal Communications*, 94(3):1263–1279, 2017.

[8] Harel Berger, Amit Z. Dvir, and Moti Geva. A wrinkle in time: A case study in DNS poisoning. *International Journal of Information Security*, 20(3):313–329, 2021.

[9] Gijs Van Den Broek, Roland Van Rijswijk-Deij, Anna Sperotto, and Aiko Pras. DNSSEC meets real world: Dealing with unreachability caused by fragmentation. *IEEE Communications Magazine*, 2014.

[10] Vaibhav Bajpai, Steffie Jacob Eravuchira, and Jürgen Schönwälder. Lessons Learned From Using the RIPE Atlas Platform for Measurement Research. *Computer Communications Review*, 45(3):35–42, 2015.

[11] Vaibhav Bajpai et al. Vantage point selection for IPv6 measurements: Benefits and limitations of RIPE Atlas tags. In *IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, 2017.

[12] Carlo Contavalli, Wilmer van der Gaast, David C. Lawrence, and Warren Kumari. Client subnet in DNS queries. *RFC*, 7871:1–30, 2016.

[13] Jiarun Mao, Michael Rabinovich, and Kyle Schomp. Assessing Support for DNS-over-TCP in the Wild. In *Passive and Active Measurement Conference (PAM)*, pages 487–517. Springer, 2022.

[14] Christian Huitema, Sara Dickinson, and Allison Mankin. DNS over dedicated QUIC connections. *RFC*, 9250:1–27, 2022.

[15] Mike Kosek, Luca Schumann, Robin Marx, Trinh Viet Doan, and Vaibhav Bajpai. DNS Privacy with Speed?: Evaluating DNS over QUIC and its Impact on Web Performance. In *IMC*, pages 44–50. ACM, 2022.

[16] Jana Iyengar and Martin Thomson. QUIC: A UDP-Based Multiplexed and Secure Transport. *RFC*, 9000:1–151, 2021.

[17] Mike Kosek, Trinh Viet Doan, Malte Granderath, and Vaibhav Bajpai. One to Rule Them All? A First Look at DNS over QUIC. In *Passive and Active Measurement Conference*, 2022.