Frag-DNS

(IP FRAGMENTATION AND MEASURES AGAINST DNS-CACHE-POISONING)

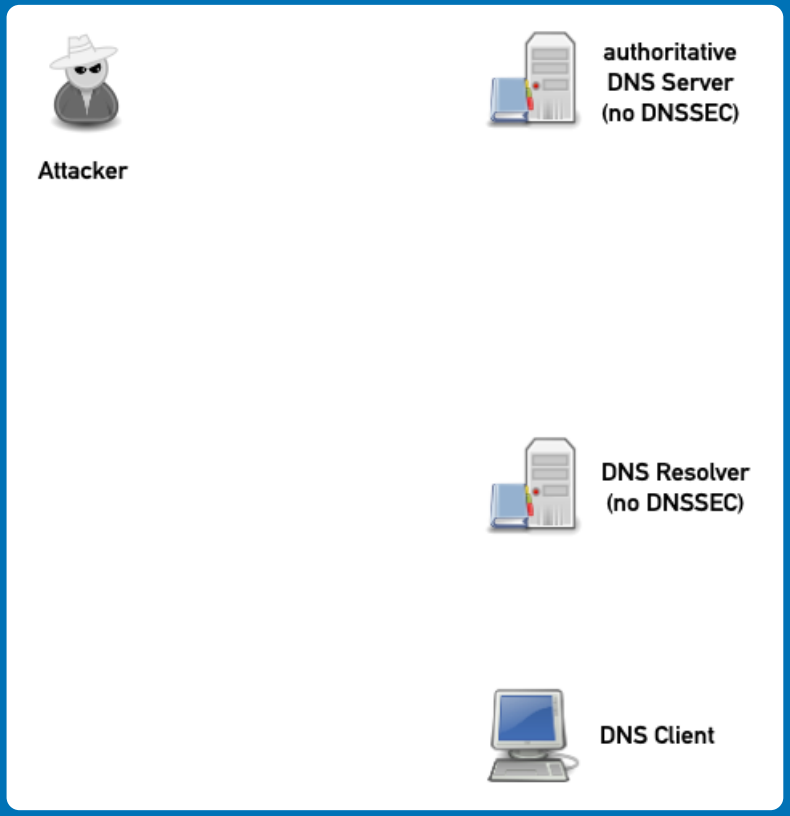
R. v. Rijswijk-Deij (NLnetLabs), P. Koetter (sys4), C. Strotmann (sys4), M. DeBrün (BSI), A. Kölligan (BSI)

About the project

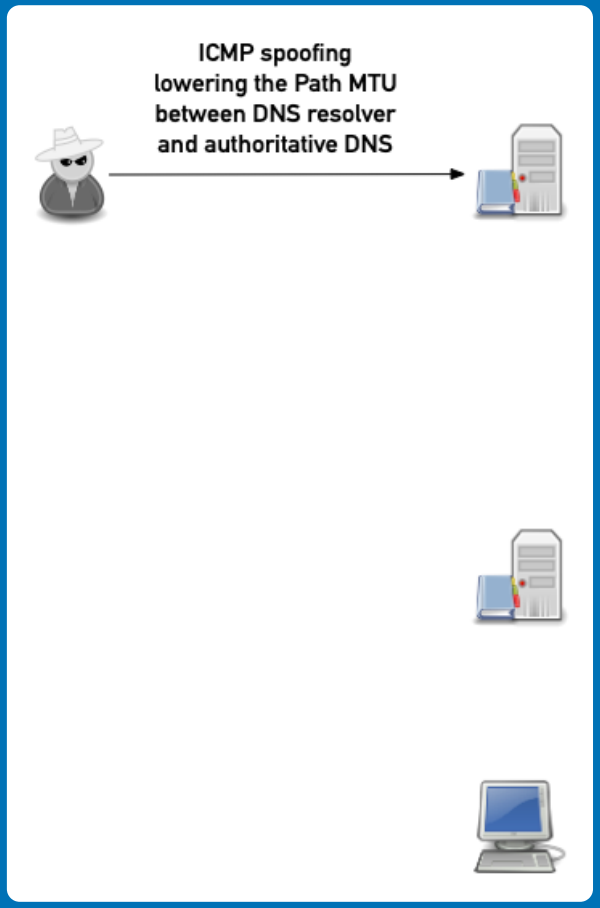
- Study under contract from BSI (German Federal Cyber Security Authority) between December 2019 and September 2021
- Roland van Rijswijk-Deij (NLnetLabs), Patrick Koetter (sys4), Carsten Strotmann (sys4), Markus DeBrün (BSI), Anders Kölligan (BSI)
- Questions:
  - Do DNS cache poisoning attacks via fragmentation impose a real threat?
  - Is it possible to mitigate such attacks?
  - How effective are these mitigations?

DNS cache poisoning via DNS fragmentation

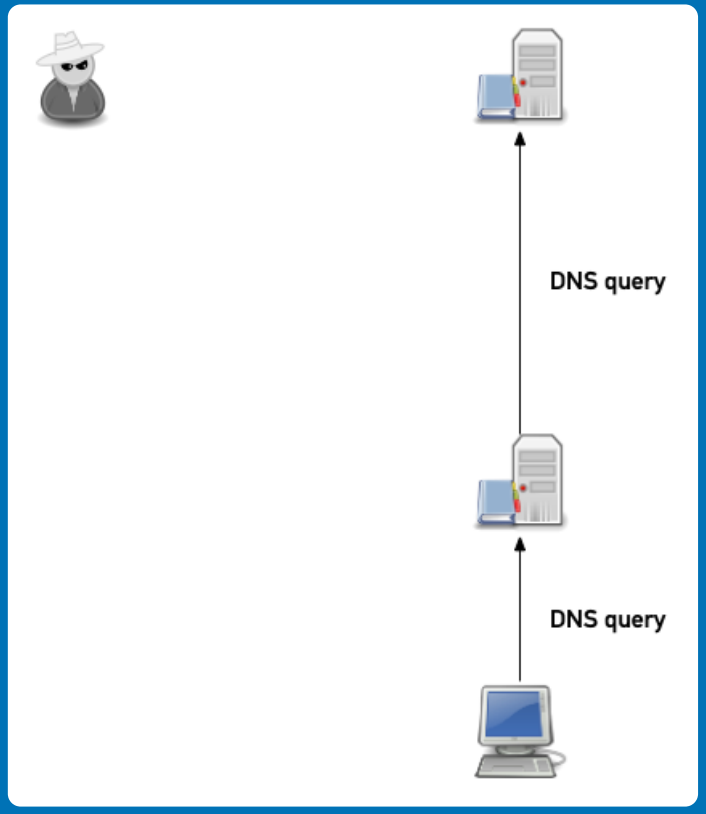
Example attack (simplified)



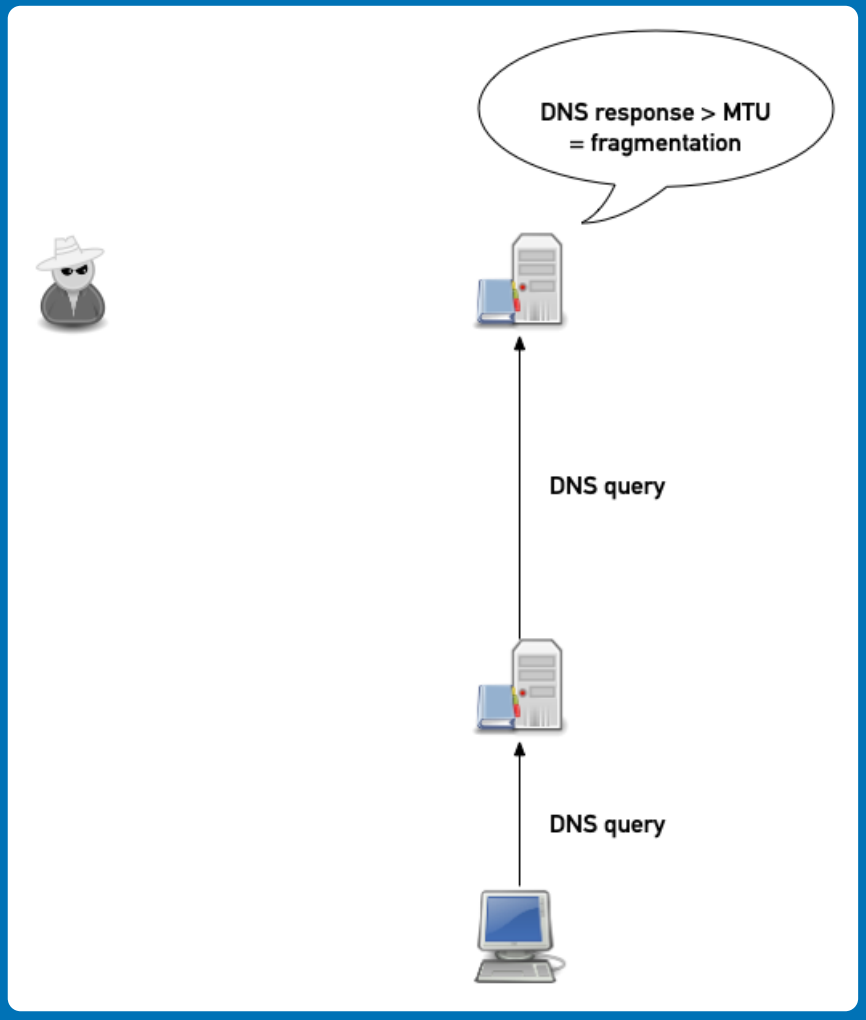
Example attack (simplified)



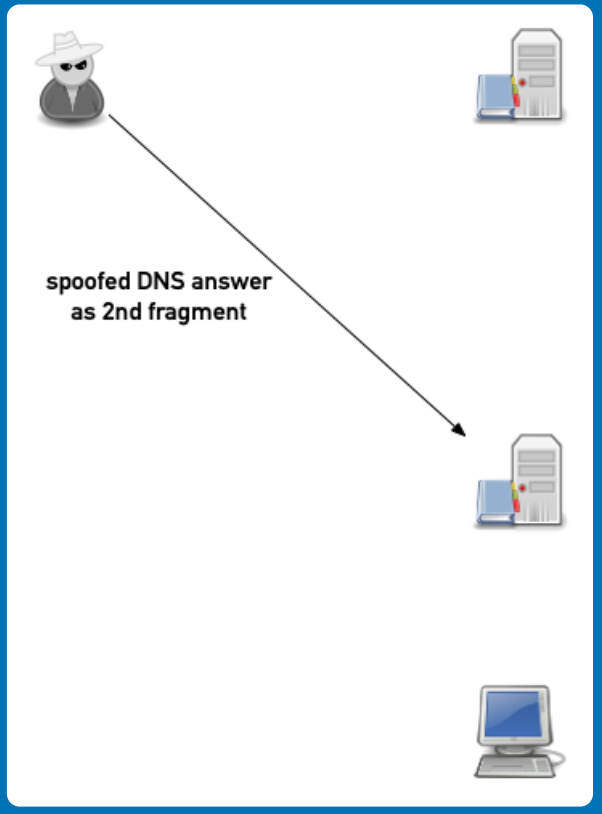
Example attack (simplified)



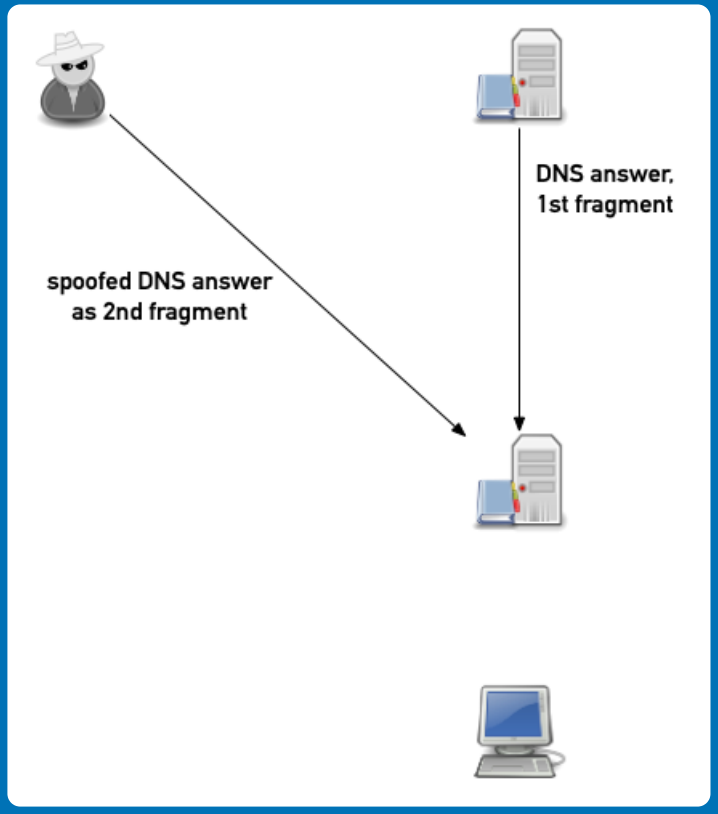
Example attack (simplified)



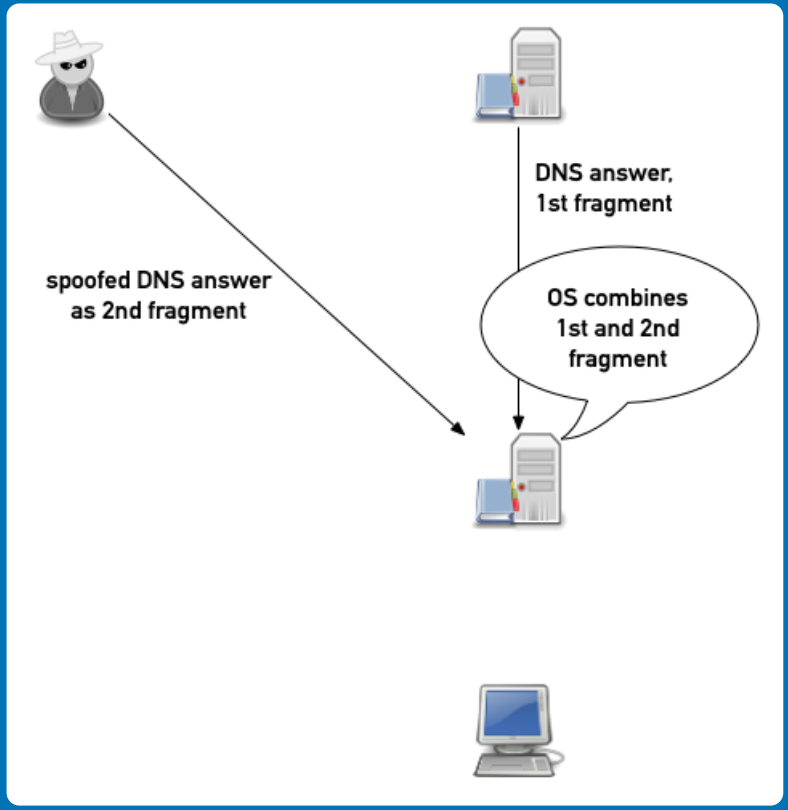
Example attack (simplified)



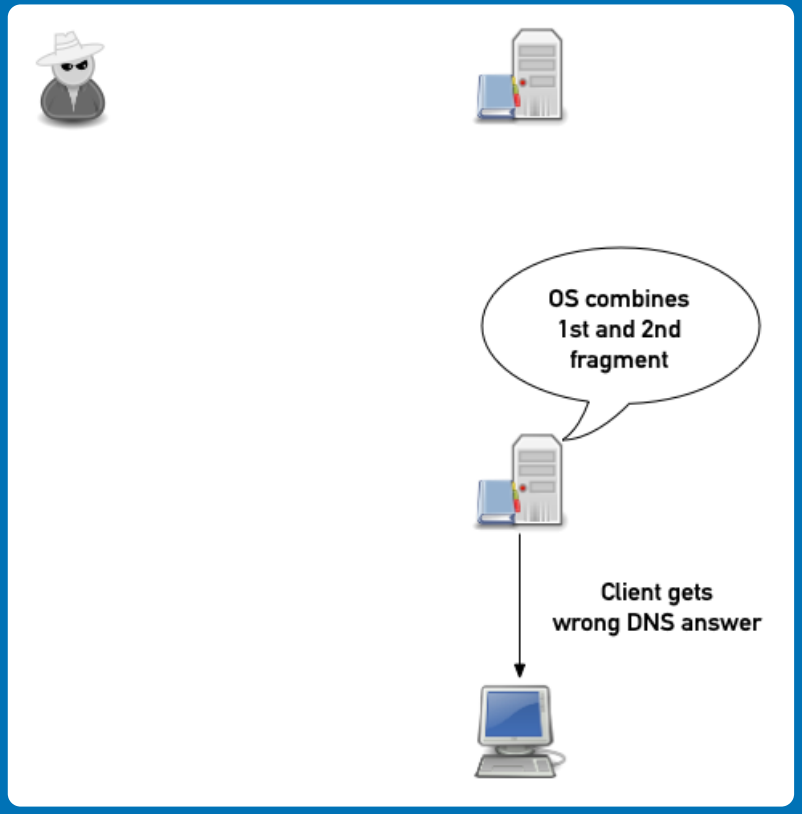
Example attack (simplified)



Example attack (simplified)



Example attack (simplified)



DNS fragmentation on an ISP DNS resolver

DNS fragmentation as perceived on an ISP DNS resolver

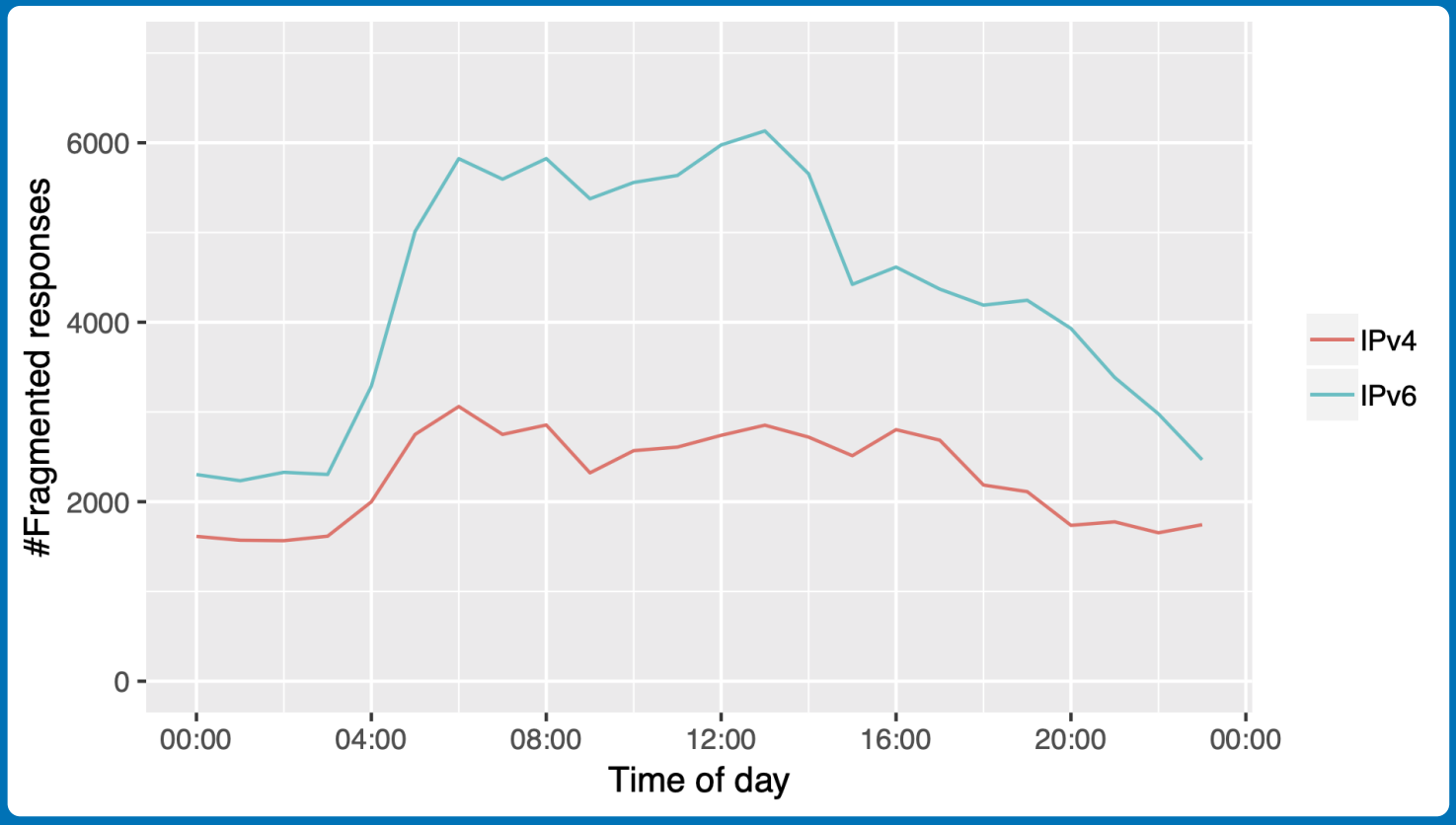
- Measurement of number, size and source of fragmented DNS responses on a DNS resolver
- Conducted in July 2020 at a large German ISP with about 4 million home and business Internet access customers

DNS fragmentation on an ISP DNS resolver

- IPv4: 55 064 DNS responses from a total of 54 023 478 have been fragmented (0.10 %)
- IPv6: 104 129 DNS responses from a total of 96 620 298 have been fragmented (0.11 %)
- DNSSEC: 93% (IPv6) and 97% (IPv4) of fragmented DNS answers came from DNSSEC signed zones

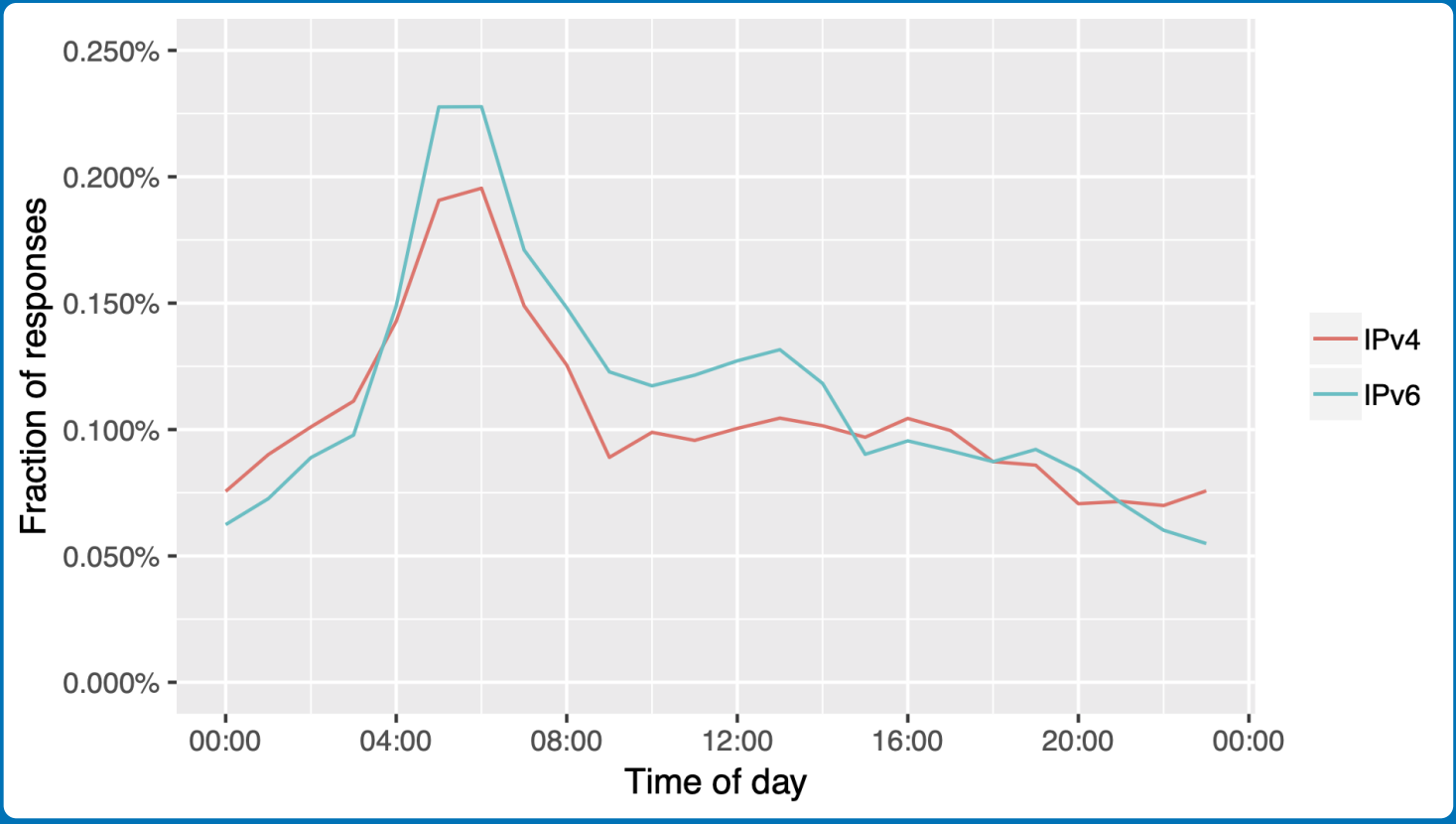
Fragmented DNS responses distribution per 24 hours

Number of fragmented DNS responses seen over 24 hours

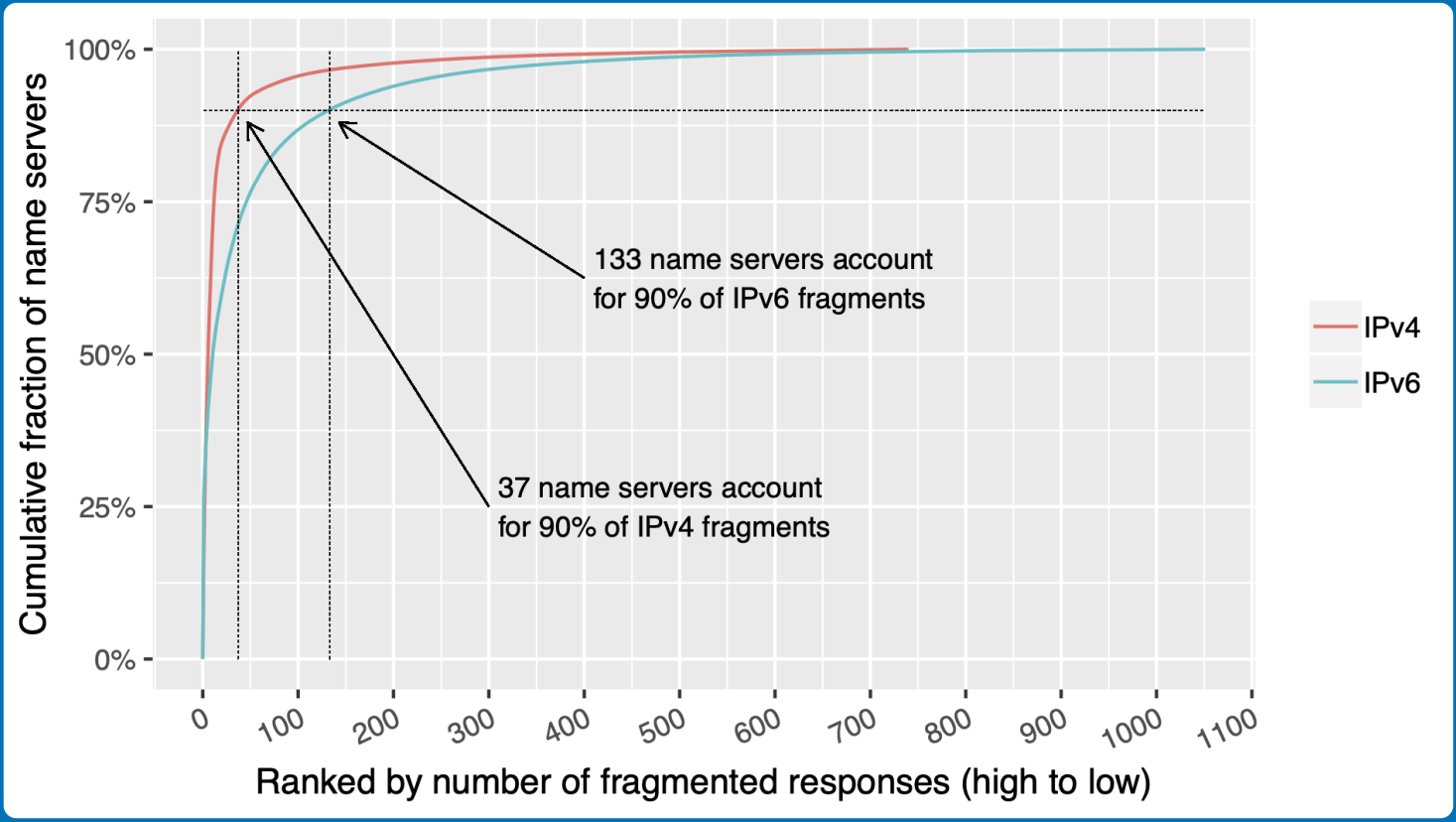


Fragmented DNS responses distribution per 24 hours

Percentage of fragmented DNS responses from the total number of responses over 24 hours



DNS Server sending fragmented DNS responses



Notable domains with fragmented DNS responses

- Domains from where fragmented DNS responses have been seen
  - office.com (Microsoft)
  - army.mil (US Army)
  - fnfis.com (Fidelity National Information Services)
  - ekom21.de (kommunales Gebietsrechenzentrum Hessen)
  - fraunhofer.de (Fraunhofer Gesellschaft)
  - rwe.de (RWE Aktiengesellschaft)
  - agilent.com (Agilent, Research)
  - checkpoint.com (Check Point Security - Firewall and VPN products)
  - salesforce.com and force.com (Salesforce.com, Inc - Cloud based customer relationship management solutions)
  - fedex.com (FexEx Corporation - multi national delivery services company)
  - gnome.org (Gnome Desktop Software - open source GUI desktop for Linux and Unix)

Fragmented DNS resonses sent from authoritative DNS server

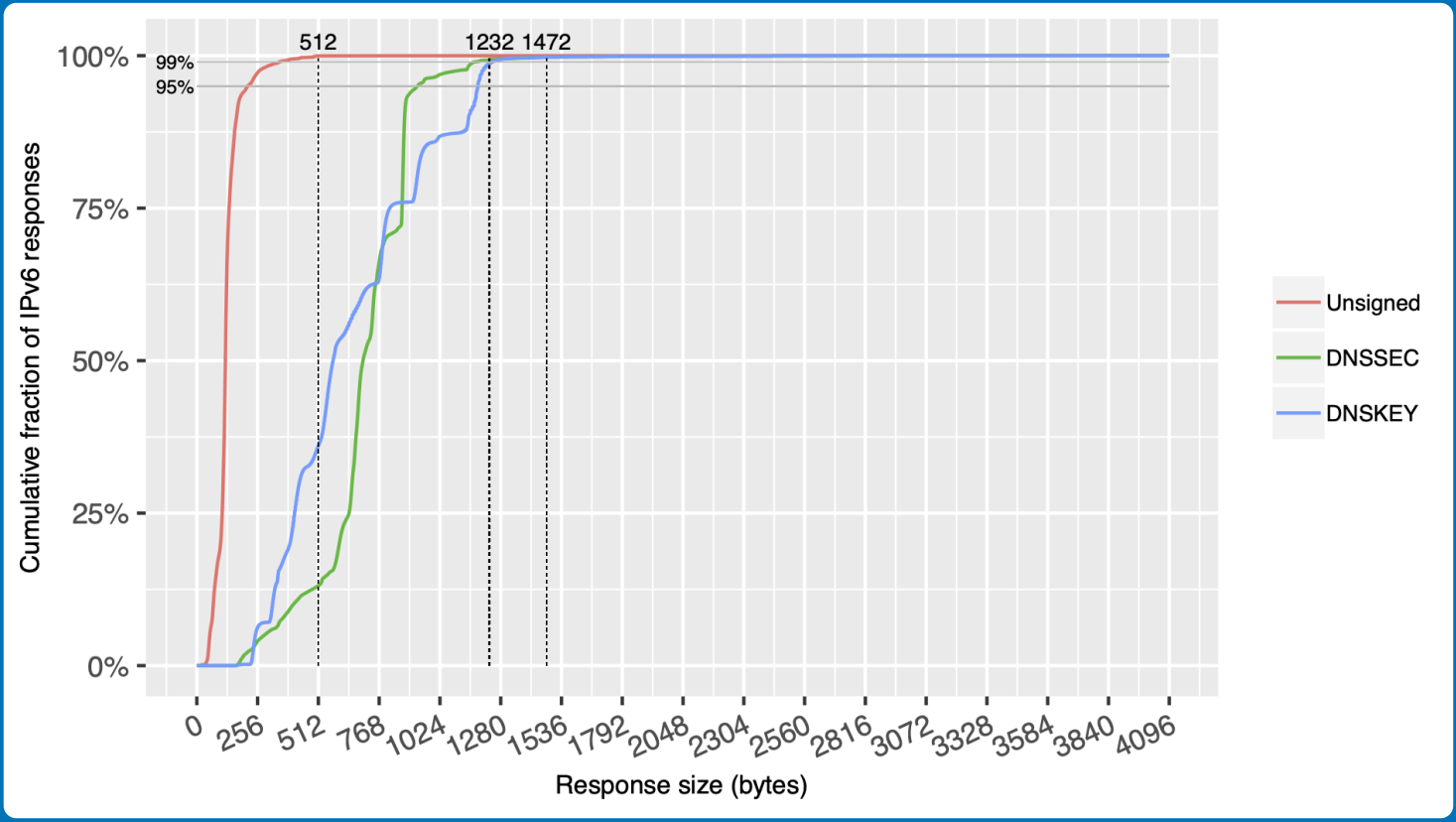
OpenINTEL

- OpenINTEL is an Internet research platform that collects DNS responses from 227 000 000 DNS domains
- OpenINTEL observes around 60% of the public Internet
- OpenINTEL processes 2.4 billion DNS datasets per day
- This study looked into the fragmentation seen in NS, A and AAAA DNS responses

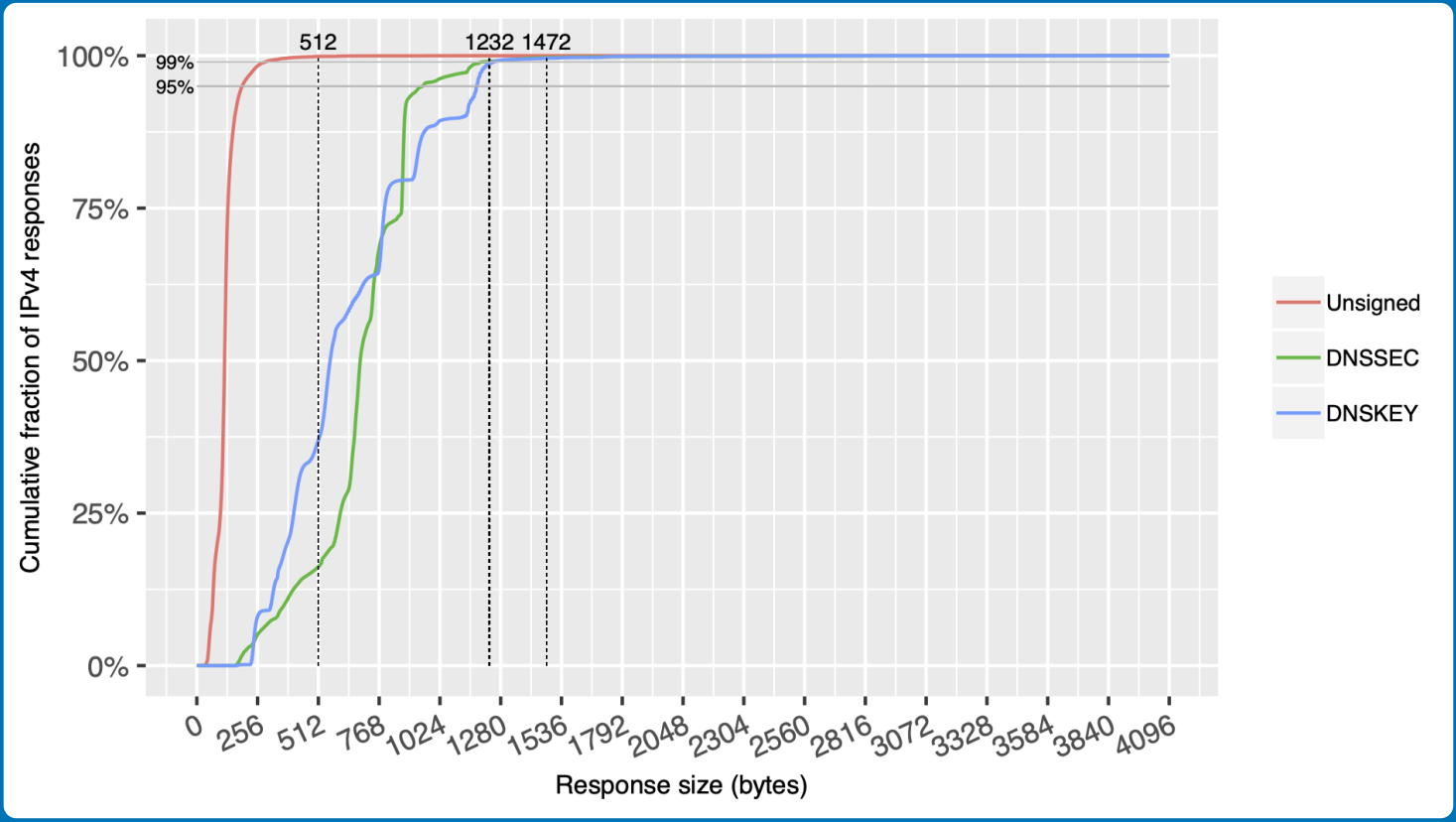
How much DNS fragmentation is seen by OpenINTEL?

- 3 893 453 582 DNS responses
  - IPv4: 2 837 177 438 [72.870%]
  - IPv6: 1 056 276 144 [27.130%]
- fragmented responses
  - IPv4: 1 334 549 [0.047%]
  - IPv6: 1 008 894 [0.096%]

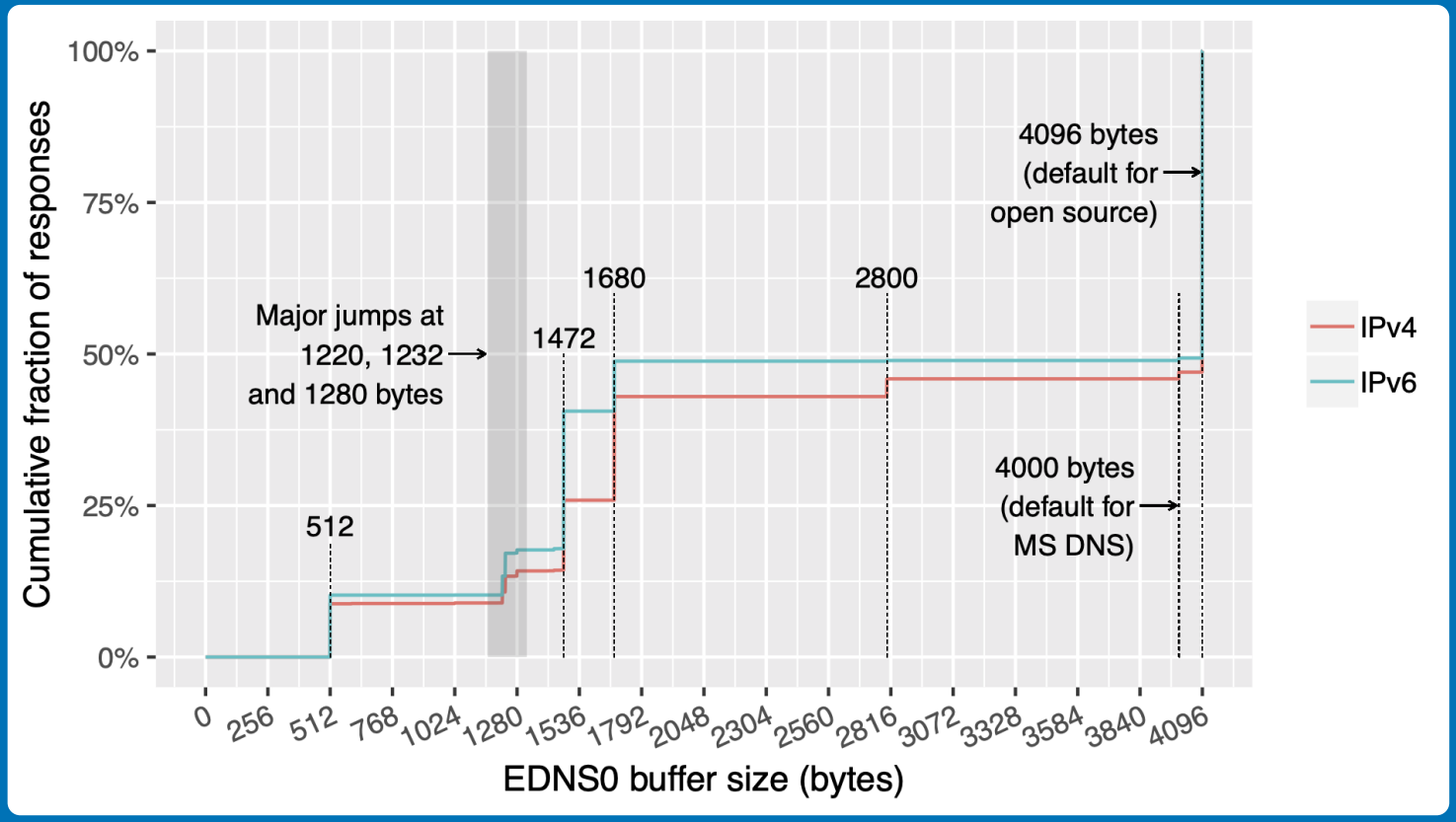
OpenINTEL: Size of DNS datagrams over IPv6



OpenINTEL: Size of DNS datagrams over IPv4



OpenINTEL: Size of the advertised EDNS buffer



Authoritative DNS servers supporting TCP

Authoritative DNS servers supporting TCP

- A DNS response that does not fit into an UDP response must be sent over TCP
- Response size limits of DNS UDP messages:
  - 512 Byte: classic DNS RFC 1034/1035 (1987)
  - 4096 Byte: EDNS RFC 2671 (1999)
  - 1232 Byte: popular recommendation to prevent DNS fragmentation
- Question: How many authoritative DNS servers support DNS/TCP?
  - How *popular* are the domains that are hosted on DNS sever that **do not** support DNS/TCP?

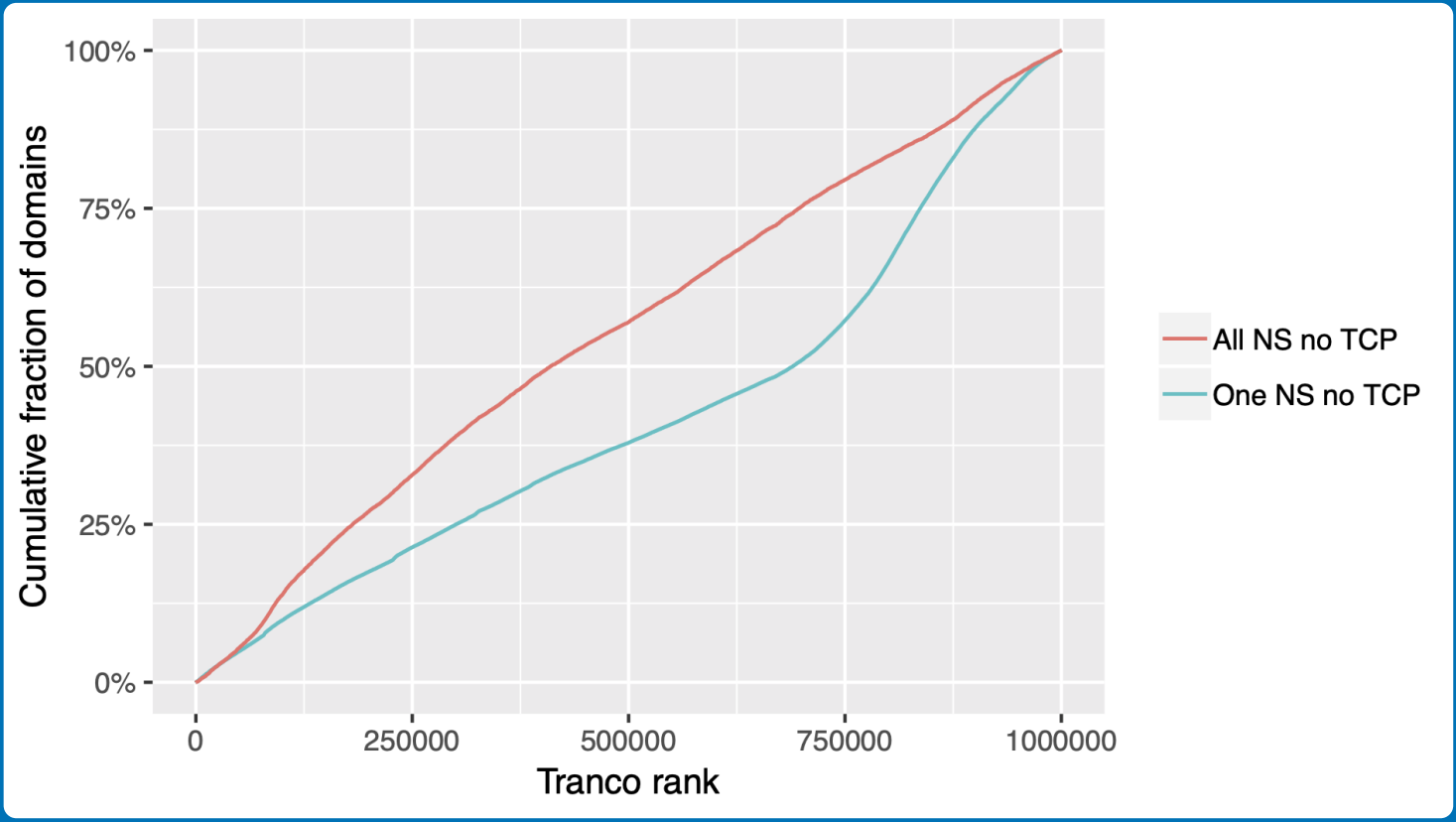
TCP Support

- 879 345 IPv4/IPv6 addresses of authoritative DNS server
  - This DNS server are authoritative for 202 765 149 domains
  - 197 773 383 (97.57%) of these domains have at least one DNS server offering DNS/TCP
  - From 183 549 827 (90.55%) domains all announced DNS servers (NS Record) offer DNS/TCP
  - 4 925 715 (2.43%) of the surveyed domains have **no** DNS server supporting DNS/TCP

TCP Support

- Domains where at least one DNS server does not support DNS/TCP contain popular Internet destinations such as `live.com`, `office.com` (Microsoft) and `yahoo.com` (Yahoo)
- 1.5% of all domains of the Tranco 1M list (list of the 1 million most popular Internet domains) have **no** DNS server with TCP support

Rank of Tranco 1M domains lacking TCP support



DNS/TCP Support - Conclusion

- Few, but also some *popular* domains do not support DNS over TCP
- Usage of DNS over TCP to mitigation DNS fragmentation attacks is therefore not recommended

ICMP Spoofing Vulnerabilities

Which Operating Systems are vulnerable to ICMP PathMTU Spoofing?

- To increase the success of a DNS attack via fragmentation, a attacker would try to lower the Path-MTU between the DNS resolver and the authoritative DNS server
  - This can be done by sending spoofed ICMP error messages
- Question: Which popular operating systems are vulnerable to ICMP Path-MTU spoofing?

Operating-Systems and ICMP Path-MTU Spoofing

- Tested the vulnerability of popular operating systems for ICMP Spoofing in a lab environment
- Question: Would an authoritative DNS server send fragmented DNS responses after a successful Path-MTU spoofing attack?

Operating Systems and ICMP PathMTU Spoofing

Operating System	minMTU IPv4	minMTU IPv6	success IPv4	success IPv6
Debian 6 / Kernel 2.6.32-5-amd64	552	1.280	X	X
Ubuntu 14.04.1 / Kernel 3.13.0-45-generic (12/2014)	552	1.280	X	X
Ubuntu 14.04.1 LTS / Kernel 3.13.0-170-generic (05/2019)	552	1.280	X	X
Ubuntu 16.04.6 LTS / Kernel 4.4.0-184-generic (06/2020)	552	1.280	X	X
Ubuntu 18.04.4 LTS / Kernel 4.15.0-106-generic (06/2020)	1.500	1.280	-	X
CentOS 6 / Kernel 2.6.32-504.3.3.el6.x86_64 (12/2014)	552	1.280	X	X
CentOS 7 / Kernel 3.10.0-1127.10.1.el7.x86_64 (06/2020)	1.500	1.280	-	X
CentOS 8 / Kernel 4.18.0-147.8.1.el8_1.x86_64 (04/2020)	1.500	1.280	-	X
SUSE EL 15SP1 / Kernel 4.12.14-197.45-default (06/2020)	1.500	1.280	-	X
FreeBSD 12.1 / Kernel 12.1-RELEASE r354233 GENERIC amd64	1.500	1.280	-	X
OpenBSD 6.7 / Kernel 6.7 GENERIC#234 i386	1.500	1.280	-	X
Windows Server 2008R2	1.500	1.280	-	X
Windows Server 2012R2	1.500	1.280	-	X
Windows Server 2016	1.500	1.280	-	X
Windows Server 2019	1.500	1.280	-	X

Operating Systems used for DNS Server

- The popular BIND 9 DNS server software responds with it's version number over DNS on request
  - This version number often contains the version of the Linux-Kernel and the version of the Linux distribution
  - We've used OpenINTEL to query for the versions used on authoritative DNS server

Operating Systems used for DNS Server (Summer 2020)

Linux OS	Number of Server	Percent from total
RedHat Linux	240.481	28.2%
RedHat EL5	7.876	0.9%
Redhat EL6	98.443	11.5%
RedHat EL7	121.103	14.2%
RedHat EL8	1.594	0.2%
Ubuntu Linux	25.034	2.9%
Ubuntu 14.04	5.110	0.6%
Ubuntu 16.04	9.314	1.1%
Ubuntu 18.04	9.467	1.1%

Operating-Systems and ICMP Path-MTU spoofing - conclusion

- Windows operating systems are not vulnerable (to Path-MTU spoofing)
- Older Linux-Kernel are vulnerable
  - These older Linux-Kernel are still in use in long-term support Enterprise-Linux systems!
  - The vulnerable Linux versions are used for authoritative DNS server on the Internet

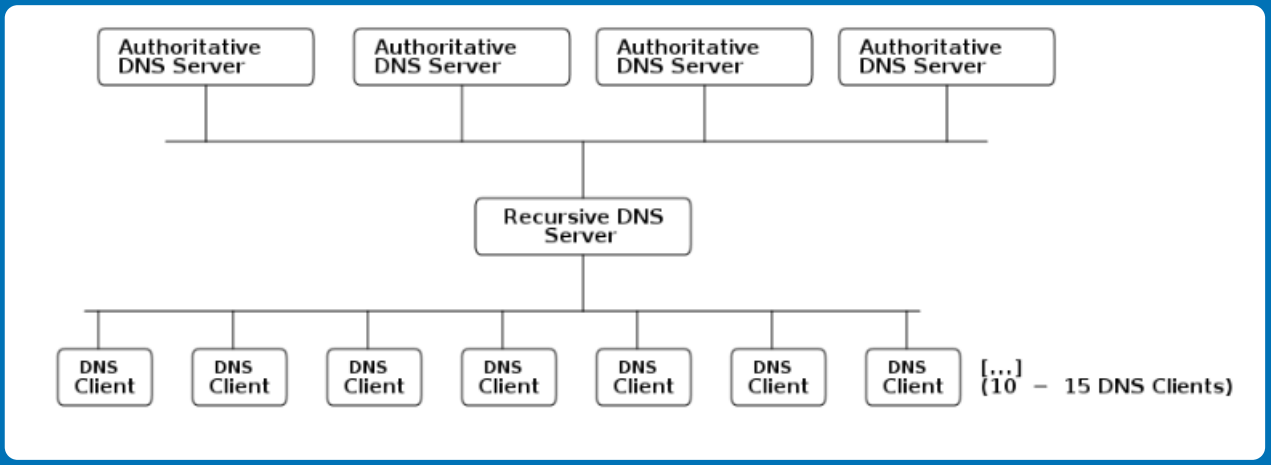
Mitigations

- Several mitigations were tested during the study
  - Impact on DNS name resolution (does it break the DNS?)
  - Impact on performance
  - Support in popular DNS server software

The Testbed

- In order to test the mitigations in a reproducable way we created a scale model of the DNS part of the Internet.
  - It allowed a simulation of realistic scenarios and ensure conclusive results
  - The test-bed setup used in this study had already been used in commercial projects with the goal of evaluating the load impact of different DNS query patterns on large DNS resolver implementations. In these previous endeavors, the test-bed had already demonstrated that its results allowed to predict real-world scenarios with high accuracy.

The Testbed


Authoritative DNS Server
Authoritative DNS Server
Authoritative DNS Server
Authoritative DNS Server
Recursive DNS Server
DNS Client
DNS Client
DNS Client
DNS Client
DNS Client
DNS Client
DNS Client
[...]
(10 – 15 DNS Clients)

Mitigation: TCP-ONLY DNS SERVICE

- TCP, in contrast to UDP, hardly suffers from IP fragmentation.
- Queries over TCP involve more work for both the DNS resolver and authoritative DNS server, and more work implies it will take longer to perform the DNS resolution.
- When communicating only with TCP towards the authoritative server, we observed a 40-44% performance drop compared to normal UDP-based DNS queries

Mitigation: TCP USING TLS 1.3

- TLSv1.3 has been heavily optimized for speed
- It might be slower than plain TCP, but the security gain using an encrypted layer might justify the performance loss
  - DNS-over-TLS (DoT) between DNS resolver and authoritative server is still experimental and not widely adopted
- The performance seen in this measurement differs by 4% from plain DNS-over-TCP, which is within the measurement variance

Mitigation: DNS USING OPPORTUNISTIC TCP (1/2)

- A fraction of authoritative DNS servers still offer their service only over UDP
  - A complete switch to a TCP-only DNS stack would have an operative impact, as domains on the UDP-only servers would become unreachable
- This test measures a setup where DNS resolvers were to use TCP whenever possible and only downgrade to UDP if the destination was unreachable over TCP
  - Experimental implementation for this study in Unbound 1.9.7

Mitigation: DNS USING OPPORTUNISTIC TCP (2/2)

- The test has seen a significant performance drop of over 70% in a scenario where some authoritative DNS server do not support DNS-over-TCP
  - The code implemented was not optimized. Further optimization could improve the performance of opportunistic TCP, but we do not expect TCP based DNS queries to achieve performance similar to that of UDP

Mitigation: DNS USING UDP FOR SMALL RESPONSES ONLY (1/3)

- An unfragmented DNS response cannot be used to poison a resolver's queue.
  - If authoritative DNS servers were to use UDP for small responses only, there would be no attack vector for UDP fragmentation-based cache poisoning
  - Measurements showed there is almost no "natural" (non-attack) fragmentation below the IPv6 minimal MTU of 1 280 Bytes

Mitigation: DNS USING UDP FOR SMALL RESPONSES ONLY (2/3)

- The test used EDNS0 to restrict the UDP message size to 1 232 Bytes (1 280 - IP-Header - UDP-Header)
  - Authoritative DNS server that need to send a larger response will trigger a query over TCP from the DNS resolver via the TC-Flag in the DNS response message
    - This is widely implemented in current DNS server software

Mitigation: DNS USING UDP FOR SMALL RESPONSES ONLY (3/3)

- This mitigation has been tested with popular DNS server products
  - BIND 9
  - Knot Resolver
  - PowerDNS Recursor
  - Unbound
  - Windows DNS

Performance Impact using EDNS 1.232 Bytes

- DNS Resolver Performance loss/win

DNS Server Product	Performance difference from baseline
Unbound	+5.2%
BIND 9	-0.8%
PowerDNS Recursor	+0.5%
Knot Resolver	-3.2%
Windows DNS	-1.0%

Mitigation: DNS DISCARDING FRAGMENTED PACKETS

- Dropping fragmented DNS traffic in general would solve the problem once and for all
  - If a firewall dropped fragmented IP packets, would a DNS resolver be able to compensate for the loss of data so that consumers would not have to suffer from that loss?
- This tests showed a slight increase of query performance for some of the tested DNS resolver.
  - This is likely due to the build-in fragmentation avoidance workarounds build into modern DNS resolvers, where when a query does not return from an authoritative DNS server, the DNS resolver will automatically lower the EDNS message size for these targets to avoid fragmentation

Mitigation: DISCARDING SMALL FRAGMENTS ONLY (1/2)

- "Natural" fragmentation below the Ethernet MTU of 1 500 Bytes is very rare
  - Discarding the first fragment only, with a size smaller than that typical to Ethernet networks with an MTU of 1 500 octet (Bytes), might turn out to be a viable tactic when it comes to mitigating IP fragmentation attacks
  - a firewall rule was set which would discard any fragmented DNS packets (first fragment) smaller than the MTU size of 1 500 Bytes

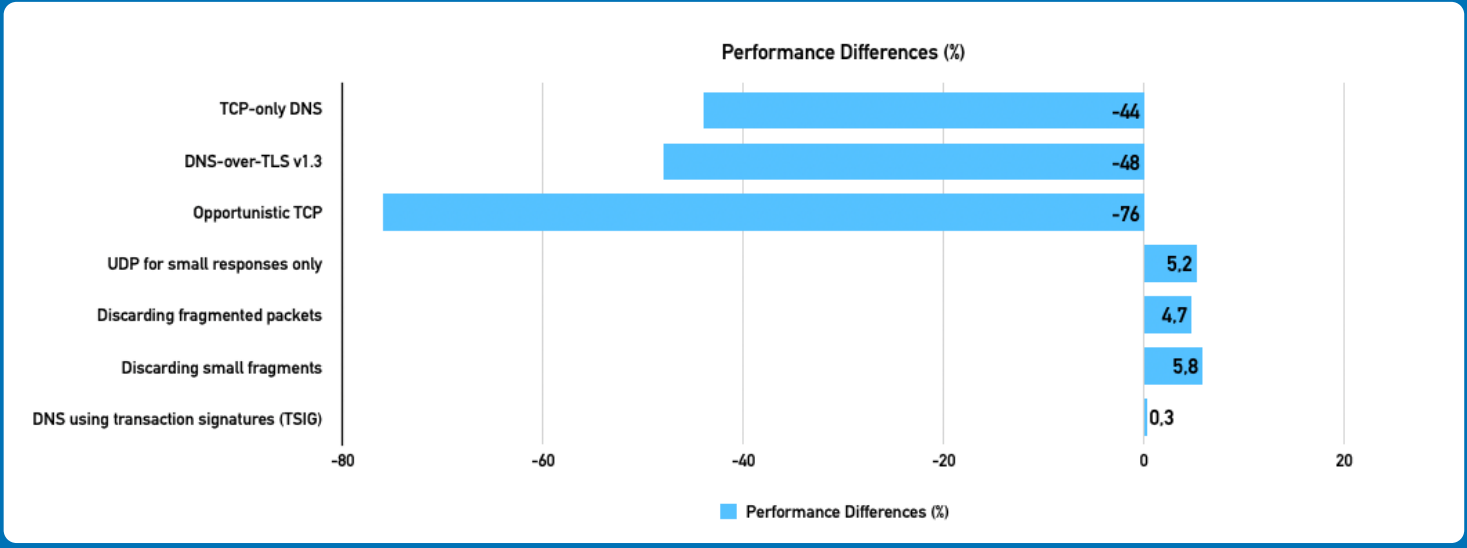
Mitigation: DISCARDING SMALL FRAGMENTS ONLY (2/2)

- Similar to the previous mitigation (dropping all fragmented packets), a slight increase of query performance has been found
  - firewall rules to filter small fragmented packets are more complex than simply dropping any fragmented packets

Mitigation: DNS USING TRANSACTION SIGNATURES

- Cryptographically-signed DNS responses (via TSIG) would secure the DNS response as a whole, making it impossible to replace any part of the response without this going unnoticed
  - This mitigation approach has been described in the Internet Draft "Defeating DNS/UDP Fragmentation Attacks"
- Our measurements show that TSIG signing a DNS response from an authoritative DNS server does not add any significant performance overhead to the DNS transaction.

Comparing the mitigations



Performance change compared to baseline DNS-over-UDP

Comparing the mitigations

Mitigation	Change from baseline (classic DNS)
TCP-only DNS	-44%
DNS-over-TLSv1.3	-48%
Opportunistic TCP	-76%
UDP for small responses only(*)	-3.2% - +5.2%
Discarding fragmented packets(*)	-3.7% - +4.7%
Discarding small fragments	+5.8%
DNS using transaction signatures TSIG	+0.3%

(*) depending on DNS resolver product

Conclusion

Conclusion (1/2)

- It is possibe to attack DNS content by means of DNS fragmentation
- The amount of *natural* (non attack) DNS fragmentation in the Internet is minimal yet still significant
- Popular domains are vulnerable
- Fragmentation of DNS responses should be avoided
  - Among the tested mitigations, lowering the EDNS buffer is the most effective one
    - once the EDNS buffer is lowered, no *natural* fragmentation should occur. All remaining fragmentation can be dropped at (host-)firewall level

Conclusion (2/2)

- The mitigations against DNS fragmentation focus on the effect and do not eliminate the cause
  - DNS cache poisoning and many other attacks on DNS infrastructure would cease to exist if operators began to DNSSEC-sign their DNS zones and DNS resolvers would DNSSEC-verify DNS responses by default

Questions?