DNS(SEC) & MTU & IPv6

S me P k t Drops Incl ded.

Tobias Fiebig

Max-Planck Institut für Informatik

,

What this is about

- PMTUD is a mess
- DNS is a Camel
- DNSSEC makes a bigger Camel
- PMTUD for DNS given MTU breaks is a messy camel
- RFC3901 does leave IPv6 for the DNS as ’optional’

Recap: PMTUD, IPv4, IPv6

- PMTUD breaks
  - BOGON sources & IX networks dropped
  - ICMP dropped
  - ...
- IPv4
  - Kind of meh, but in the end the on-path nodes can always fragment
- IPv6
  - Only end-hosts can fragment
  - A Tier 1 is dropping IPv6 fragments on transit

Recap: DNS & MTU

- The times of 512b max DNS packets are long gone
- DNSSEC makes packets (even) bigger
- Not everyone follows RFC9715 (Frag avoidance)
- DNS over TCP is ‘not fun’ either
- **Claim:** This means that IPv6 breaks DNS resolution[1]

[1] This discussion is ongoing for some time now: `https://insinuator.net/2015/11/some-notes-on-the-drop-ipv6-fragments-vs-this-will-break-dnssec-debate/`

Authoritative Side Measurements

- Make a lot of clients resolve names
- Make alterations to the resolution to figure out how they react to that
- Log their queries

Measuring the Resolver Side

- Challenging to get a sufficiently large sample
- Each parameter increases the number of packets needed for the measurements by a factor
- Needs a lot of parameters

Parameters

- MTU (1500, 1280 on-link, 1280 on-path)
- PMTUD (working, not working)
- v4 only, v6 only, dual-stack
- EDNS0 512, 1232, 4096

Doing Science: Forming Hypotheses

- Hypotheses: Testable predictions
- Here: *DNS resolution is negatively impacted by IPv6 under MTU breaks; DNSSEC makes this even worse.*
  - The NOERROR % should be notably higher on IPv4 vs. IPv6
  - The SERVFAIL/TIMEOUT % should be notably lower on IPv4 vs. IPv6
  - There should be notably more TCP fallback and too-big-packets for v6 as PMTUD is more broken there
  - There should be more EDNS0 fallback to 1232 for v6 as PMTUD is more broken there
  - Forced MTU breaks notably impact v6 more than v4
  - Effects are more pronounced for DNSSEC enabled zones

How and when to measure

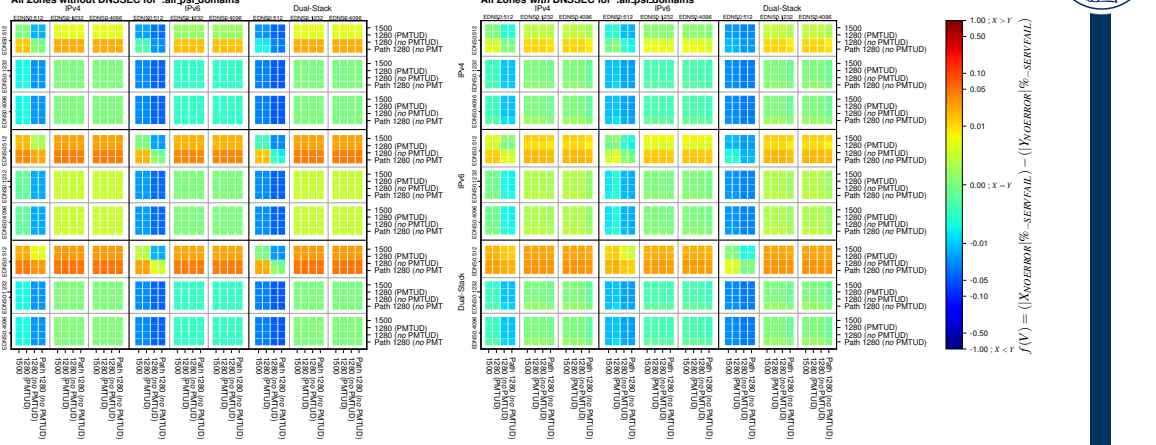
- We only need to ask each NS-Set once
- Regularly get the Google CrUX Top 10M and identify unique NS-Sets
- Pick one domain for each NS-Set
- For each NS-Set that has at least one DNSSEC enabled zone, also pick one of those
- Running once per day

Open Data

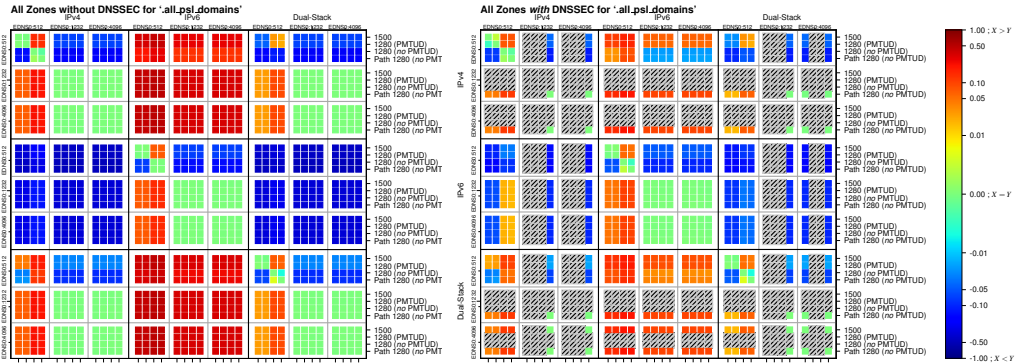
- ZDNS resolution logs
- ZDNS resolution results
- PCAPs
- PCAPs of (intentionally) dropped packet-too-big messages
- All intermediate data
- Pre-generated plots
- `https://data.measurement.network/dns-mtu-msmt/out_data/`

Note: Currently still in a rather rough shape, and more of a ‘first glimpse’

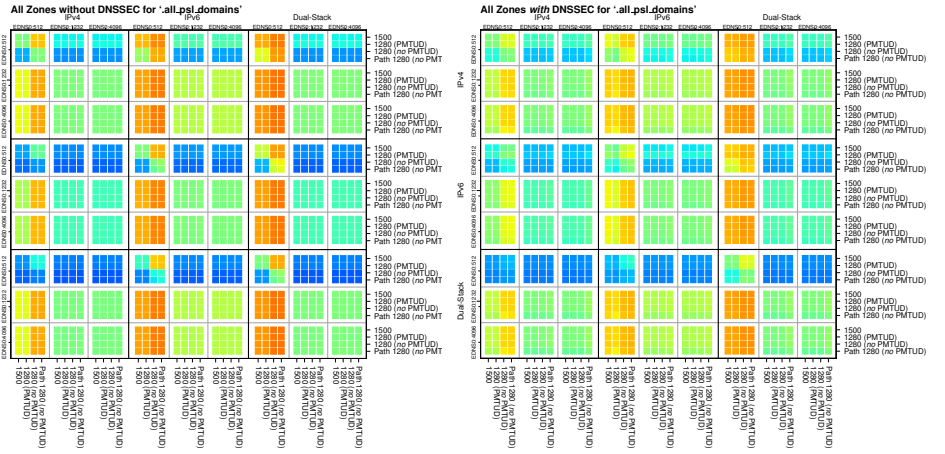
Results: NOERROR%



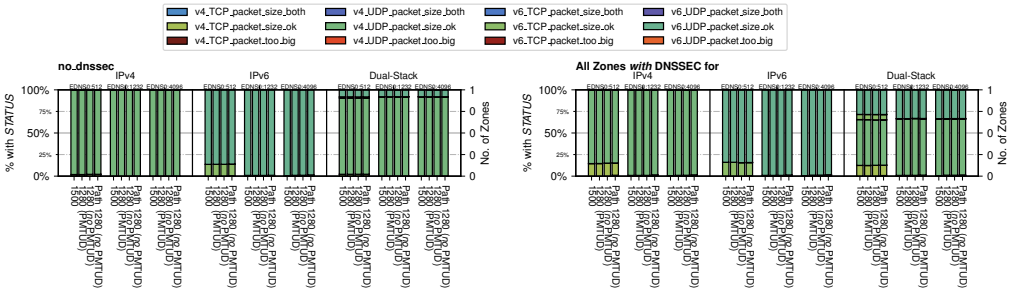
Results: SERVFAIL%



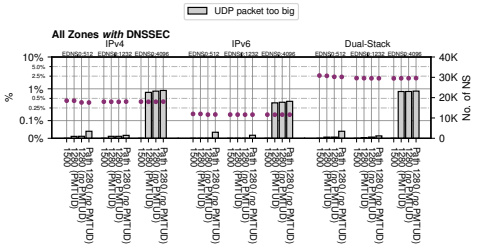
Results: Timeout Comp

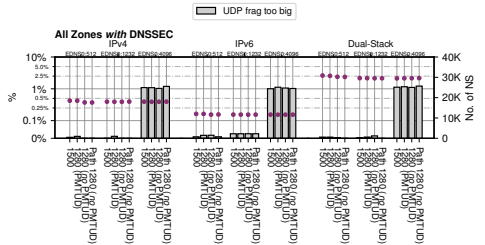


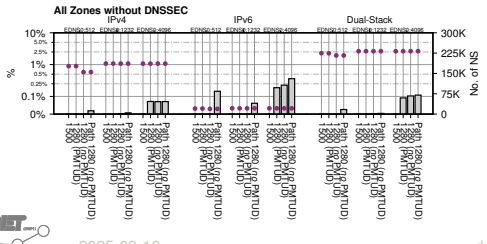
Results: Packets per NS

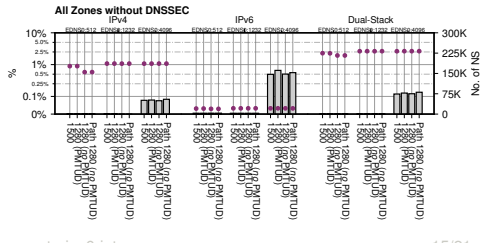


UDP Packet Too Big / Frag Too Big

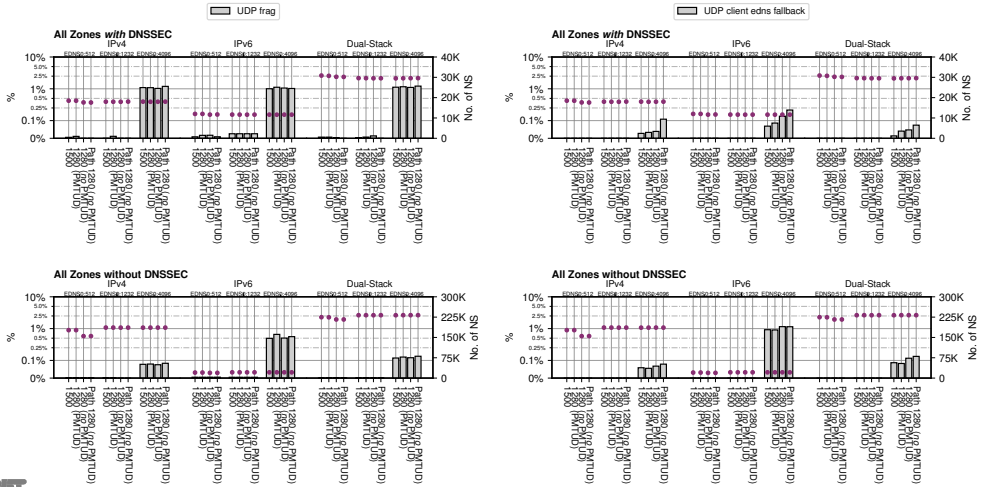




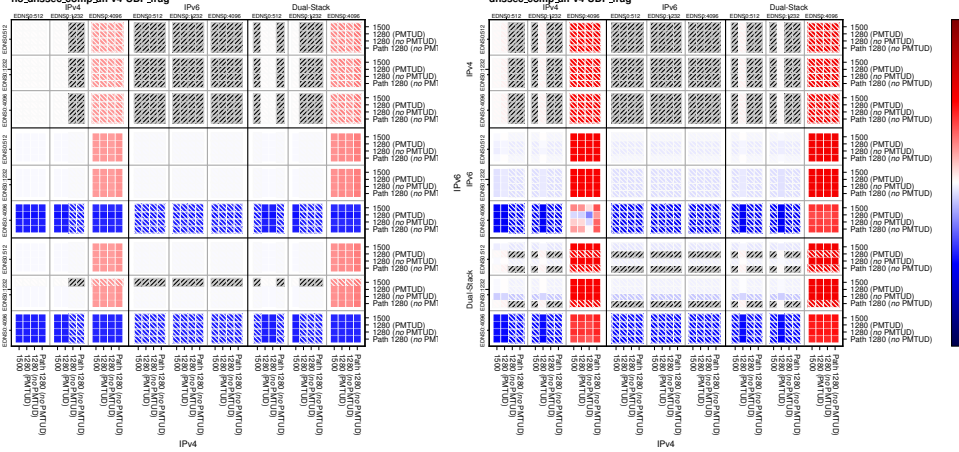




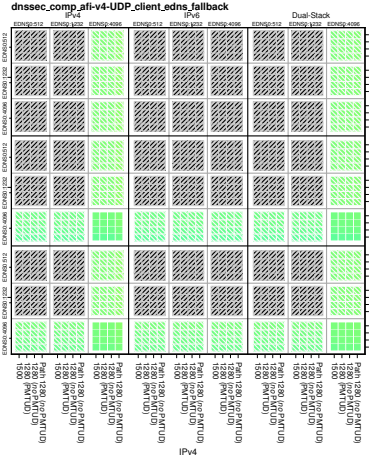
Results Fragments / EDNS0 Fallback

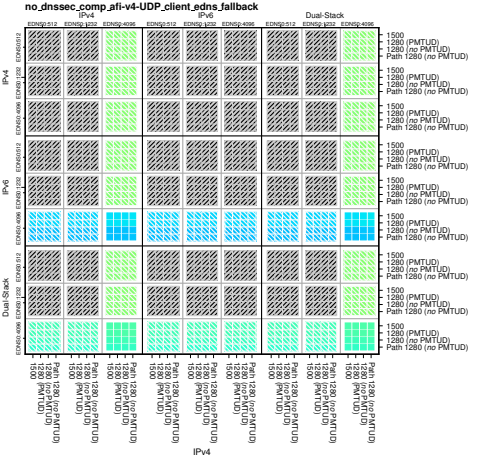


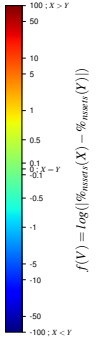
Recap: AFI Frag Comp



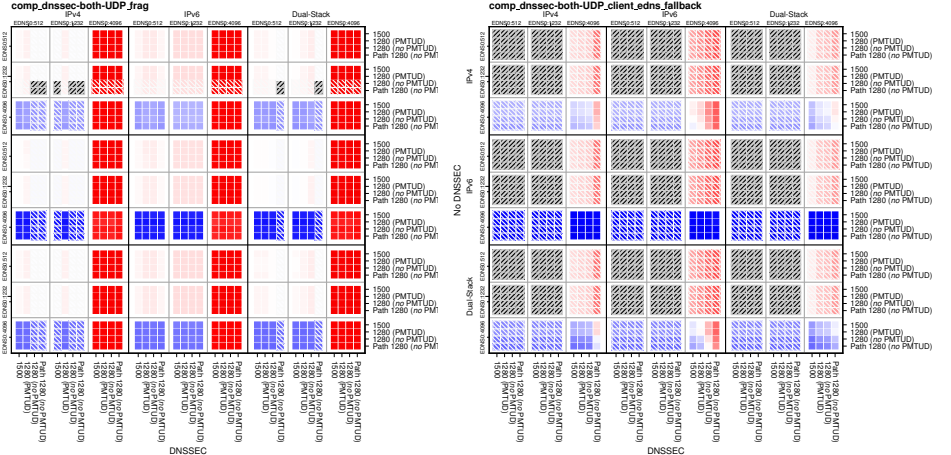
Recap: EDNS Fallback AFI Comp







Recap: DNSSEC Comp



Conclusion

- The NOERROR % should be notably higher on IPv4 vs. IPv6
  - $< 1\%$
- The SERVFAIL/TIMEOUT % should be notably lower on IPv4 vs. IPv6
  - For SERVFAIL: $< 1\%$; Less SERVFAIL for DNSSEC serving NS-Sets
  - For TIMEOUT: $< 1\%$; Notably less for DNSSEC serving NS-Sets across AFIs
- There should be more TCP fallback and too-big-packets for v6 as PMTUD is more broken there.
  - TCP fallback is mostly an EDNS512 thing and comparable between v4 and v6 for NS-Sets hosting DNSSEC enabled zones
  - For non-DNSSEC zones, TCP Fallbacks are mostly an IPv6 thing (unreachable cases)
  - There is *less* UDP packet-too-big for DNSSEC hosting NS-Sets than for v4

Conclusion

- There should be notably more EDNS0 fallback to 1232 for v6 as PMTUD is more broken there
  - There is around $0.05 - 0.1\%$ more NS-Sets with an EDNS0-Fallback for v6 vs. v4 for NS-Sets hosting DNSSEC enabled zones
  - EDNS0 fallback is more pronounced for v6 for NS-Sets not hosting DNSSEC enabled zones
- ~ Forced MTU breaks notably impact v6 more than v4
  - Mild ($< .25\%$) impact for individual metrics, more pronounced for NS-Sets not hosting DNSSEC enabled zones
- Effects are more pronounced for DNSSEC enabled zones
  - Effects are *less* pronounced for DNSSEC enabled zones