Domain Name System (DNS)

Outline

- Ecosystem and overview
  - You'll notice multiple occurrences of "$"
- DNSSEC (some of Olaf's "Block_03" slides)
  - https://www.dns-school.org/ is somewhat old but still accurate in large part
  - A number of images are copied from there (thanks NLnet!)
- DNSSEC deployment
- DNS privacy

What's the DNS?

- **The** single world-wide distributed naming system/database that essentially replaced the host file because that got too big

  - There's also the DNS protocol, which is how clients and servers interact

- Main purpose is to map names to IP addresses so that applications and humans can deal with names rather than addresses

  - tcd.ie is easier to type/remember than something within 134.226/16 or 2001:770:10::/48

  - IPv4 addresses use A resource record (RR) type

  - IPv6 addresses use AAAA resource record (RR) type

- The DNS is also used for many other purposes

  - Mail address right-hand-side to mail server name(s) via MX RR

  - DNS block lists of spam sources (and other block lists)

  - "Passive DNS" monitoring for various security purposes

  - Telling lies (RPZ) to help people avoid "bad" domains

  - State- or enterprise-level censorship

  - ...

DNS names

- We know that tcd.ie is a DNS name, right?

    - tcd.ie is a domain

- www.tcd.ie and down.dsg.cs.tcd.ie are also DNS names

    - Those are Fully Qualified Domain Names (FQDNs)

- tcd.ie is the parent of cs.tcd.ie, cs.tcd.ie is a child of tcd.ie

- TCD internally (modulo cloudiness) manage the names below tcd.ie within their "zone," often spoken about as if managed via a "zone file"

- Reality: most zone data is probably in databases, but tiny domains can keep all their data in one file and there's a standard format for zone files that's supported by most DNS tools

Zone file fragment

```
$ORIGIN .
$TTL 3600   ; 1 hour
responsible.ie      IN SOA  ns2.my-own.net. hostmaster.responsible.ie. (
          2020101984 ; serial
          86400      ; refresh (1 day)
          7200       ; retry (2 hours)
          3600000    ; expire (5 weeks 6 days 16 hours)
          3600       ; minimum (1 hour)
          )
      NS  ns2.my-own.net.
      NS  ephemera.nmugroup.se.
      A   213.108.108.218
      TXT "This host is used as the souce of zmap scans ..."
      AAAA    2a00:c6c0:0:119:7::10
      CAA 128 issue "letsencrypt.org"
```

The real zone file for this domain has keys and signatures for DNSSEC

DNS Ecosystem

- The root: "."

- Top Level Domains (TLDs)
  - Country-code TLDs (ccTLDs): .ie, .uk, .is,...
    - Each more or less do what they want
    - IEDR manage .ie zone, CZ.nic manage .cz, ...
  - Generic TLDs (gTLDs): .com, .org, .net,...
    - Run under ICANN's oversight (https://icann.org)
    - There are ~1500 of those now (because $$$)
      https://data.iana.org/TLD/tlds-alpha-by-domain.txt

- Second level domains (2LDs), or effective Top Level Domains (eTLD)
  - Comply with parental controls (to some extent)
  - Examples: example.com, tcd.ie, amazon.com
  - .com zone has ~160M names, .ie has ~330k, .org has ~10M

- Third level and below: controlled by 2LD/eTLD
  - E.g. down.dsg.cs.tcd.ie

Public Suffix List (PSL)

- Some TLDs don't have all 2LDs directly below the TLD, e.g. .co.uk, .com.au etc.

- Causes a problem for browsers, when deciding whether to re-tx cookies in HTTP

- Ickky "solution" is the PSL

    - https://publicsuffix.org/ maintained by Mozilla and others

    - A text file with 15495 lines at 20240219

        - Was 14305 lines at 20230307; 13,679 lines at 20210419; 13,085 lines at 20200218

- PSL has other uses, e.g. anti-spam

- PSL ideally would be maintained via information in the DNS, but is not, and attempts to do that (IETF DBOUND wg) have all failed so far

    – Another attempt is in the offing now

- Indicative of how DNS can be messy but works despite all

Registry/Registrar/Registrant

- Top Level Domains (TLDs) are operated by registries,

    - IEDR for .ie (they rebranded themselves recently as weare.ie)

    - "Identity Digital" operate a whole bunch of ccTLDs and gTLDs (used to be donuts, and acquired Affilias)
    https://www.identity.digital/tld-portfolio

    - Public Interest Registry (PIR) operate .org (and feed $$$ to Internet Society, which feeds $$ to IETF and RFC editor – notwithstanding 2019/2020 controversy about possible sale of PIR)
    https://thenew.org/org-people/about-pir/

- Registrars are accredited by registries and deal with registration of names (and transfer and de-registration)

- Registrant is the entity that wants/has a name registered

    - Per-registry rules may apply, e.g. "connection to Ireland" for .ie

- Registries handle name conflicts, e.g. when trademark issues arise via some dispute resolution process (can involve $$$)

- Registration costs to registrants from registrars vary from "free" to ~$1000, but mostly ~$10 per year

    - Some money flows up from registrar to registry (ccTLD or gTLD) and to ICANN (for gTLDs)

- ICANN auction new gTLDs now and then

    - Costs ~$1M+ to play that game, ICANN have ~$150M resting in an account as a result

Registry/Registrar

- Registrar <-> registry protocols vary a lot

  - IEDR have a web console and an "API" that accredited registrars can use

  - Extended Provisionin Protocol (EPP)

  - Registration Data Access Protocol (RDAP)

- whois

  - "Legacy" protocol where registry publishes some registrant data
    - May contain personally identifying information (PII)

  - You can install "whois" on you machine or use via the web

  - Lots of fun with ICANN and whois and GDPR

DNS Operators/Servers

- A (logical) zone file for a domain is served by an Authoritative DNS server

- Typically 2 or 3 authoritative servers are serving each zone for redundancy

- The DNS protocol has "zone transfer" commands (e.g. AXFR) that help syncing multiple servers

- Recursive DNS servers query the Authoritatives to resolve names, e.g. starting at "." ask "where is .ie"; get answer; at ".ie" ask "where is tcd.ie"; at "tcd.ie" ask "what is the IP for www.tcd.ie"

- Clients (your laptop/phone) ask Recursive servers to resolve names e.g. "where is tcd.ie" and if it doesn't already know the answer (or if earlier answers have timed out) the Recursive will do as much of the dance above as needs (re-)doing

- The protocol spoken between clients, Recursives and Authoritatives is the DNS protocol

- On linuxes, the "`dig`" tool allows you to explore the DNS ("`nslookup`" on windoze)

Example

```
$ dig +dnssec ietf.org

; <<>> DiG 9.18.4-2ubuntu2.1-Ubuntu <<>> +dnssec ietf.org

;; global options: +cmd

;; Got answer:

;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 27100

;; flags: qr rd ra ad; QUERY: 1, ANSWER: 2, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:

; EDNS: version: 0, flags: do; udp: 1232

;; QUESTION SECTION:

;ietf.org. IN  A

;; ANSWER SECTION:

ietf.org.  1800  IN  A 50.223.129.194

ietf.org.  1800  IN  RRSIG A 5 2 1800 20240229210120 20230301200316 40452 ietf.org.
    JZN6Rk7d+wv34m728PfIW0Yq87FANjlTKmrMvBjTBKmuxW6uMM9/3bC0
    HqCOb0UR2kg+YKfs9v9zppbaZmTJR80Zd4NGXYfnrU0QcraDSeV9AGLb
    fQJMq3oDwtE7TG5d8nEFOpPN3ggv32k1Bu1r66qihEl7On8+SmofiZI0
    Nb3087leOLUTwS002sMyD1NS94gZGUSb1JcX7CUYfDjK3nm81k5GR7k3
    6cynaGJHGE2IGJ5ZWpP4Ro2OJ9g6rcYi5c+/ZqPGHB1lcdmolb417S6F
    UtIu8JGpzalWT63yAHWreA3yE86Q4lCaNDfR4Gta1q5qnHme8h9aLNEN LWs66w==

;; Query time: 1135 msec

;; SERVER: ::1#53(::1) (UDP)

;; WHEN: Wed Mar 15 16:06:15 GMT 2023

;; MSG SIZE  rcvd: 349
```

The root zone

- The root zone "." is special – it's content is (carefully) managed by IANA and handed over to the root server operators…

    - https://www.iana.org/domains/root/servers

- The root server operators serve the root zone – about 1000 instances worldwide in about 130 countries, with subsets of those managed by 12 organisations via 13 named root servers (there was one merger in the last couple of decades)

- Most root zone instances are accessed using anycast IP routing

    - Other public authoritative and even recursive servers (e.g. QuadN's such as 8.8.8.8 or 1.1.1.1 or 9.9.9.9) also use anycast for better performance

- The root zone is pretty stable – and the Internet really needs that to be the case

- Every Recursive needs at least one root server IP to start

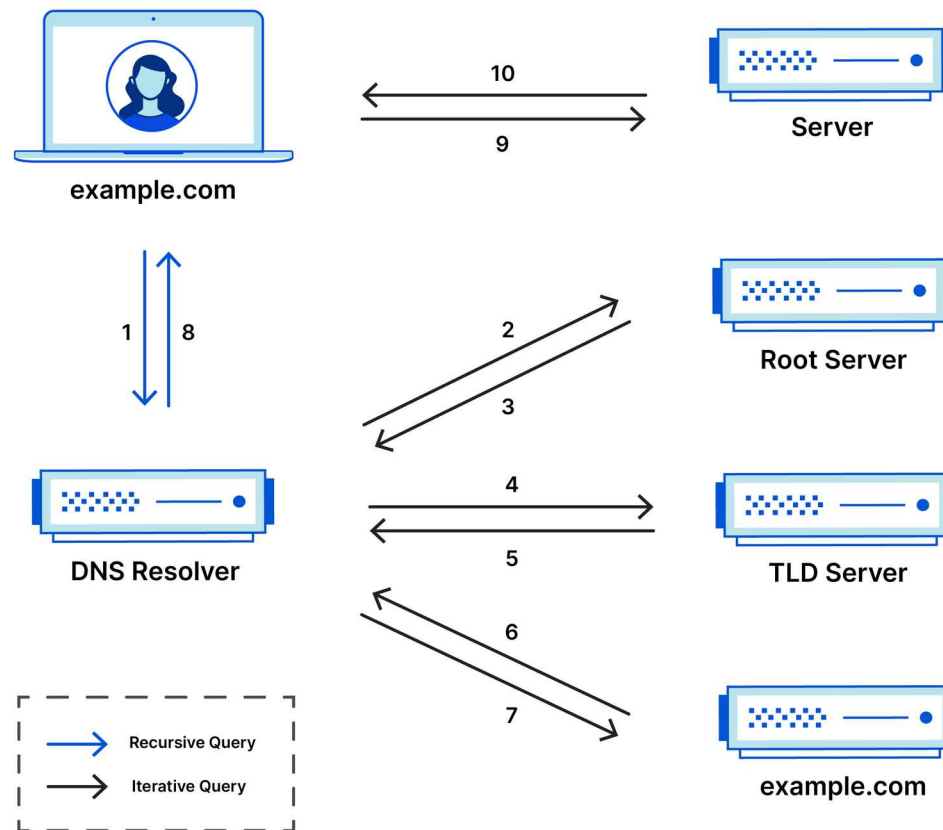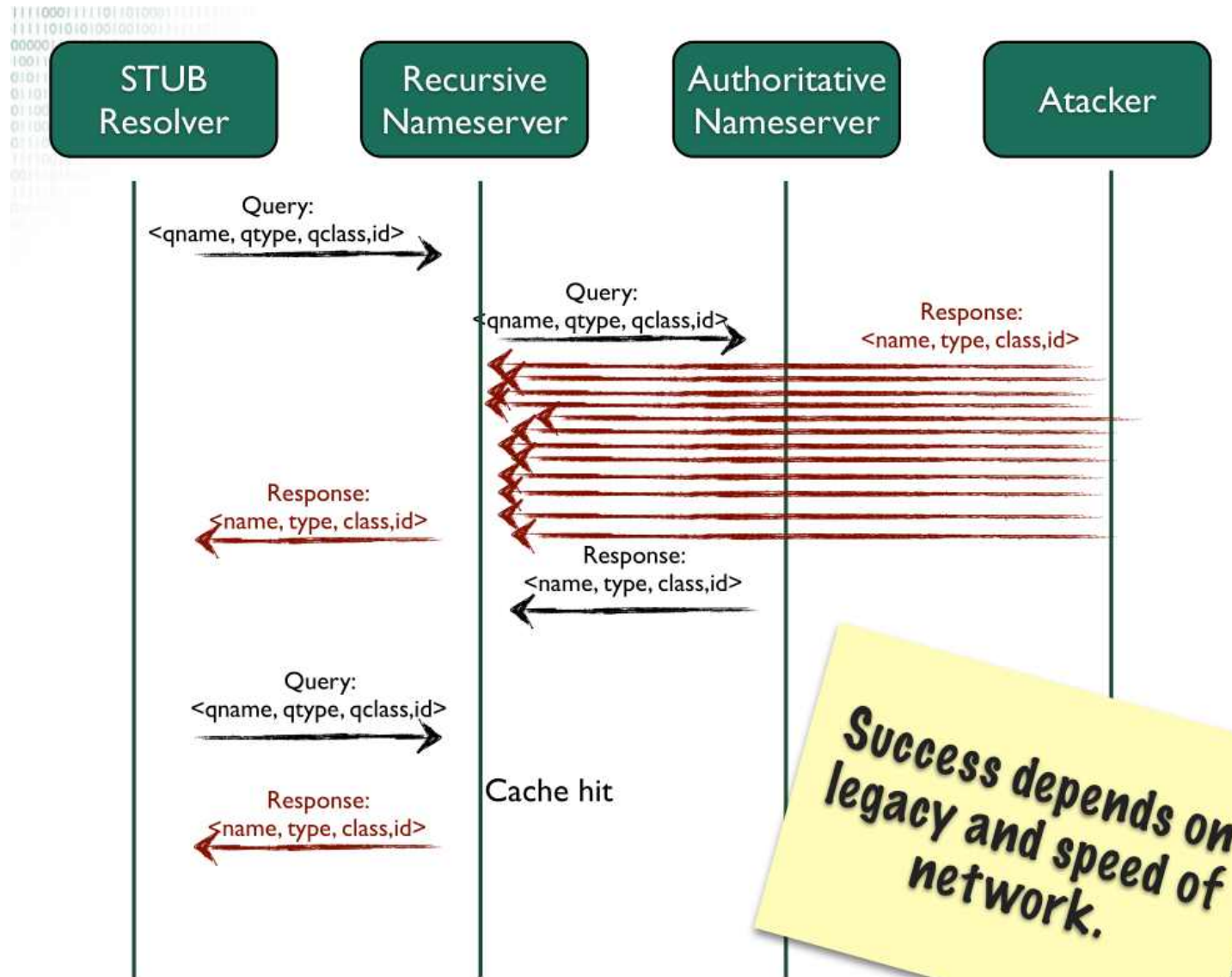# Example DNS Resolution

Complete DNS Lookup and Webpage Query

Image from: https://www.cloudflare.com/learning/dns/what-is-dns/

DNS Poisoning

- DNS queries sent by default via UDP to port 53

- Query contains:
  - QNAME, type, class, 16-bit id, source port

- Answer contains
  - QNAME, type, class, id, RRset, additional data

- Anyone can answer

DNS Poisoning

Olaf's DNSSEC Slides

DNSSEC Deployment

- Dependency on parent (for DS record) makes DNSSEC hard to deploy

    - Should registrar or registrant contact parent?

    - If registrar, how does zone get signed, or, how does DS/KSK get to registrar? (usually via a crappy web form)

    - If registrant, how does registry know it's dealing with the right party (registrant has a/c at registrar, not registry)

- Early DNSSEC deployments broke things (and still can)

    - DNSSEC adds a new thing to manage (RRSIG expiry) that you can muck up

    - But many zones these days are likely fairly dynamic (VMs) so maybe this is lessening as a downside

- There are reported issues with stubs and recursives that don't handle DNSSEC well, or who even strip DNSSEC RRs (typical middlebox issue!)

    - Browsers are quite intolerant of >1% additional failures

- There was also lots of delay getting the root zone signed (only happened in 2010)

- Some zone maintainers (say they) cannot sign their zones due to lack of control over names

- Some zone maintainers claim that DNSSEC isn't worthwhile for them

DNSSEC Deployment

- CDS/CDNSKEY (RFC 8078) provides a way for zone maintainer to publish a "new" DS (CDS) or new KSK (CDKSKEY) in their zone

  - Parent scans children (who are known to do this) and can pick up new DS value that can be used to populate parent zone file (if various conditions met)

  - Hasn't seen much deployment yet, but should help with ongoing maintainance, allowing much easier changes to KSKs

- I know of one pentester who likes to see DNSSEC as it tells him that his customers have asserted control over their DNS

- DNSSEC does offer real protection though, e.g. the DNSpoinage attack would likely have been detected and some aspects of those attacks may have been prevented

  - https://krebsonsecurity.com/2019/02/a-deep-dive-on-the-recent-widespread-dns-hijacking-attacks/

- Some big Registrars are starting to include DNSSEC in a bundle

  - https://www.godaddy.com/en-ie/hosting/premium-dns

- Some TLDs are incentivising registrars (via discounts of maybe 10% of $) to deploy DNSSEC for new domains

  - Leads to more deployment, not clear if more security

DNSSEC Deployment

- Result: ~3-4% of 2LDs signed, similar percentage of names covered

- Some stats: https://www.statdns.com/

Shows about 3.8% of .com signed (was 3.2% a year ago and 2% a year before) but 57% of .se (was closer to 70% a year ago)

- More stats: https://stats.dnssec-tools.org/
  - Shows similar numbers, slowly increasing total of 23M DS RRs in 2023

- Some major web properties/CDNs are not signed, others are
  - Not sure why – maybe something to do with internal systems

- "Economic incentives on DNSSEC deployment: time to move from quantity to quality"
  - https://ieeexplore.ieee.org/abstract/document/8406223/
  - https://research.tue.nl/en/publications/economic-incentives-on-dnssec-deployment-time-to-move-from-quanti

KeyTrap (discovered: Nov '23)

- Implementation flaw present for >10 years; arguably also a protocol flaw

- Basic idea: publish "attack" zone with colliding key IDs and *many* (bad) signatures and keys; resolver will spend CPU trying to find a good sig

- To execute: submit query to target resolver, authoritative forces query to TCP, encodes hundreds of sigs/keys in one 64K DNS answer message

- Result: up to hours of DoS with CPU @ 100%

- Most DNS servers have by now issued new releases with mitigations

- Blog: https://labs.ripe.net/author/haya-shulman/keytrap-algorithmic-complexity-attacks-exploit-fundamental-design-flaw-in-dnssec/

- Tech report (paper to follow): https://www.athene-center.de/fileadmin/content/PDF/Keytrap_2401.pdf

- Credit: Elias Heftrig, Haya Schulmann, Niklas Vogel, Michael Waidner

DNS Privacy

- All data published in DNS is public, so historically there was little/no interest in confidentiality when DNSSEC was defined

- But the fact of access to DNS data can be sensitive, e.g. if you access https://www.aa.org/ that may say something about your life

- RFC9076 documents DNS privacy considerations

  - Names, timing, IP addresses (e.g. if local recursive), client-subnet

- Mitigations:

  - Use Tor browser

  - QNAME minimisation (RFC9156)

  - Define ways to provide confidentiality for DNS traffic (DoT/DoH/ADoX)

  - Don't (always) send EDNS(0) client subnet

DNS over TLS (DoT)

- IETF "DPRIVE" working group has defined how to run DNS over TLS (DoT, RFC7858)

- DoT is usable today between stub and recursive

- Generally, you replace your system stub resolver (e.g. systemd, dnsmasq) with something that can do DoT (e.g. stubby+unbound – I do that)

  - https://dnsprivacy.org/wiki/display/DP/DNS+Privacy+Daemon+-+Stubby

- There are public recursives now who offer that kind of "DNS privacy service", e.g. 9.9.9.9, 1.1.1.1, ...

DoT with padding

- DNS message lengths may leak information about names

- RFC 7830 describes an EDNS(0) padding option

- Responders MUST pad if requesters do (and MAY in any case)

- RFC 8467 describes ways in which one might use padding and recommends:

    - Pad queries to block lengths of N x 128 octets

    - Pad responses to block lengths of N x 468 octets

    - Don't do random stuff (signal leaks), maximal-length is wasteful (esp if we go > MTU)

Recursive <->Authoritative

- Today, DoT is usable for stub <-> recusive

- Would like to also secure recursive <-> authoritative

- Can't amortise TLS state so much so needs lots of performance testing, esp., if done near root

- Not clear if/how to authenticate authoritative (various proposal being considered)

- Might get deployed in medium term, but not clear

- Despite uncertainty, we sometimes talk about Authoritative DNS over TLS, (ADoT) 'cause that's what we'll most likely end up doing

- Opportunistic ADoT can be deployed - RFC9539

DNS over HTTPS (DoH)

- Browsers and some JS code however can't easily tell if DoT is being used as there's no portable OS API to use for that today (some OSes may add such a thing, not sure)

- So DNS over HTTPS (DoH, RFC 8484) describes how to encapsulate DNS traffic in HTTPS

- Supported today in FF with their "Trusted Recursive Resolver" (TRR) concept, with a set of "built-in" TRR instances (Cloudflare, NextDNS,...)

- Google, Microsoft and Apple have all made announcements about their plans for DoH. They differ in various ways.

- DoH lead to the usual **major** fuss – the move from a system/OS stub, to an in-browser stub causes many changes and people fear/dislike such changes

Anti-DoH!

- Around 2019, various operator-like folks described their problems with DoH (or more correctly with the mozilla/CF deployment they feared might happen)

    - https://tools.ietf.org/html/draft-bertola-bcp-doh-clients

    - https://datatracker.ietf.org/doc/html/draft-doh-reid-operator

    - https://tools.ietf.org/html/draft-livingood-doh-implementation-risks-issues

- Some of the above folks even started an industry association, with an apparent goal of encouraging deployment of DNS privacy, but possibly in a flavour that better suits ISPs who traditionally operated the DNS recursives: https://www.encrypted-dns.org/

- None of the above are objective analyses, but work will likely happen to do that analysis, because there are some real issues (if DoH gets widely deployed in applications):

    - Split horizon

    - Loss of enterprise control for BYOD and/or parental control @ home

    - Passive DNS

IETF has chartered a working group to look at some of the less controversial aspects of all this: https://datatracker.ietf.org/wg/add/about/

DNS Privacy enables ECH

- Once/if we get deployment of DNS privacy (whether via DoT or DoH) then we can try to tackle SNI encryption as part of the TLS handshake

  - https://tools.ietf.org/html/draft-ietf-tls-esni

- Idea: publish a new DH public share in DNS and use that to encrypt SNI in the TLS ClientHello

- Still in-flux, but it works!

  – I'm working on code for this:
    https://github.com/sftcd/openssl or via https://defo.ie

My DNS client setups

- Laptop OS: unbound+stubby
    - Unbound caches, stubby does DoT; config prefers my DNS recursive
    - QNAME minimisation/padding turned on
    - Stubby config has to include DoT servers on port 443 because of port blocking
    - Occasional glitches require "sudo service stubby restart" now and then
- DoT testing: kdig tool
- Within FF: DoH to NextDNS TRR (other browsers: I forget:-)
- De-googled android uses DoT to 1.1.1.1 (Cloudflare)
- On-phone browser (brave) uses DoH to 8.8.8.8 (Google)
- At home: my home router (Turris) does DoT to Cloudflare using knot-resolver (as stub)
    - Clients at home use Turris as their DNS recursive via Do53
    - Recursive distribution via DHCP/RA as normal
- Anyone could do all that (but you'd wonder why they'd bother;-) client stuff ought be much more out-of-the-box and likely will become so
- I used to operate my own DNS recursive resolver and use it in the above – I moved that VM though and haven't gotten around to setting that up again yet

My (ex-)DNS recursive setup

- Runs on a small VPS rented from a local hoster (as of 20200225)

- DoH: Apache handles port 443 DoH queries, hands those to dnsdist which en/decpasulates DNS in/from HTTP, and uses a local unbound as the real caching, validating recursive

- DoT: knot resolver (as recursive) directly handles port 853 traffic

  - dnsdist could/can also do that, just not today's setup

- QNAME minimisation turned on in upstream queries (note to self: check that!)

- Padding turned on for stubs

- Server side problem: above means me, those in my home, and ~two buddies "hiding" in a very tiny crowd – would be better with a bigger (local) crowd but without involving mega-company services like Google or Cloudflare

- A setup like this is never going to be as fast/reliable as a mega-company one

- Net effect: moves DNS surveillance/poisoning risk from any of ISP, TCD or visited-networks mostly to my chosen hoster or someone attacking his network – but he's a chap who lives in Sligo and is hence much closer to me accountability-wise, compared to Cloudflare or Google

SVCB/HTTPS RR

- Browsers like things being fast but DNS and TLS add RTTs

- HTTPS RR (RFC 9460) provides a way to put more things needed for TLS into a single RR and maybe save RTTs

```
$ dig +short https rte.ie

1 . alpn="h3,h2"
  ipv4hint=104.16.239.220,104.17.160.147
  ipv6hint=2606:4700::6810:efdc,2606:4700::6811:a093
```

- Can also "solve" the "CNAME at apex" issue for CDNs

- SVCB is a generalisation of HTTPS

- HTTPS and SVCB RR definition extensible allowing new tag/value pairs to be defined so even more connection and TLS related info can be found via DNS before connecting to (TLS) server

DELEG RR

- A current topic under IETF discussion: could we benefit if we had an SVCB-like RR higher up the tree?

- Possibly make it easier to handle NSes, DS RRs, ADoT and other things needed at a "zone-cut"

- Maybe also make DNS operators into first-class citizens (like registrant/registrar/registry) as more people now outsource DNS ops

- To be/was discussed at IETF-119 in March 2024

DNS Conclusions

- DNS is sort-of criticial infrastructure that (sometimes amazingly) works well

- DNSSEC deployment is still woeful

- DNS privacy is starting to be addressed, and represents a real change, but one that will happen in the presence of significant tussles