dprive or dnsop                                          K. Henderson
Internet-Draft                                                Verisign
Intended status: Informational                               T. April
Expires: January 9, 2020                                       Akamai
                                                          J. Livingood
                                                              Comcast
                                                         July 8, 2019

### Authoritative DNS-over-TLS Operational Considerations
#### draft-hal-adot-operational-considerations-00

Abstract

   DNS over TLS (DoT) has been gaining attention, primarily as a means
   of communication between stub resolvers and recursive resolvers.
   There have also been discussions and experiments involving the use of
   DoT to communicate with authoritative nameservers (Authoritative DNS
   over TLS or "ADoT"), including communication between recursive and
   authoritative resolvers.  However, we have identified a number of
   operational concerns with ADoT.  These operational concerns need to
   be addressed prior to ADoT's deployment at scale by DNS operators in
   order to maintain the stability and resilience of the global DNS.
   The document also provides some suggested next steps to advance the
   operator community's understanding of ADoT's operational impact.

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF).  Note that other groups may also distribute
   working documents as Internet-Drafts.  The list of current Internet-
   Drafts is at https://datatracker.ietf.org/drafts/current/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time.  It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on January 9, 2020.

Table of Contents

## 1.  Introduction

   Typically, DNS communication between stub resolvers, recursive
   resolvers, and authoritative servers is not encrypted.  Some argue
   that this can pose a privacy challenge for Internet users, because
   their access to named network resources can potentially be tracked
   through their DNS communication.  In principle, any network element
   along the path between the user and resolving or authoritative
   nameservers could observe this unencrypted traffic.  Many of these
   concerns are addressed in [RFC7626].

   [RFC8310] proposes using DNS over TLS (DoT) in order to encrypt DNS
   traffic.

   Historically, much of the work on DNS encryption has focused on the

stub-to-recursive path as the recursive-to-authoritative server path
does not leak user specific information.  However, with the increased
deployment of EDNS0 Client Subnet [RFC7871], recursive-to-
authoritative encryption is becoming an area of interest.  Therefore,
this document focuses on the recursive-to-authoritative aspect of DoT
and we use the term Authoritative DNS over TLS (ADoT) in order to
differentiate it from the stub-to-recursive path.

The addition of ADoT, while providing encryption for DNS
communication, also introduces other factors that might impact the
stability and resiliency of authoritative nameserver operations which
may have been optimized for unencrypted DNS, often focusing on UDP
transport.

The objective of this document is to try to describe the problem
space, make suggestions about solutions, and propose next steps that
can help inform both recursive and authoritative operators on how to
assess and address the challenges posed by ADoT deployment.

## 1.1.  Background and Motivation

### 1.1.1.  Why operational considerations are so important for ADoT

The main concerns for most authoritative operators are the stability,
resiliency, scalability, and performance of their platforms.  These
concerns need to be weighed against the benefits, if any, offered to
the end user by encrypting DNS queries to the authoritative servers
and the associated benefit of protection from modifications in
transit.

The communication between the recursive-to-authoritative server is
less able to be associated with a particular user than information
communicated along the stub-to-recursive path given that the
originating IP address is that of the recursive resolver, not the
user's, and the QNAME is potentially minimized.  Therefore, initial
deployments of ADoT may offer an immediate expansion of the attack
surface (additional port, transport protocol, and computationally
expensive crypto operations for an attacker to exploit) while
potentially providing limited benefit to end users.

### 1.1.2.  Other considerations related to ADoT

As resolvers add encryption on the client-to-recursive path, they may
also change the way they handle data on the recursive-to-
authoritative path.  This is expressed in Mozilla Trusted Recursive
Resolver (TRR) requirements [1], for example, which require
participating resolvers to perform QNAME minimization [RFC7816], and
TRR requirement #6, which forbids the EDNS0 Client subnet (ECS) from
being propagated unless the recursive-to-authoritative path is
encrypted.

The latter requirement may have the possible unintended consequence of reducing the authoritative name servers' ability to provide a best response to DNS queries, until such time as they deploy DNS encryption.

Given that recursive resolvers should be configured to prevent ECS transmission to root, top-level, and effective top-level domain (TLD) servers [RFC7871] section 12.1 [2] - the ECS encryption requirement motivates consideration of authoritative DNS encryption below these levels.

At the higher levels, techniques such as QNAME minimization and Aggressive Use of DNSSEC-Validated Cache [RFC8198] arguably provide an alternate path toward mitigating the risk of disclosure of sensitive information without the operational risk of DNS encryption.

Resolver requirements may change as the understanding of DNS encryption options evolve, but in the meantime, they provide motivation for authoritative name server operators to weigh the risks and benefits of DNS encryption, hence the importance of understanding these operational considerations.

## 2. Terminology

### 2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

### 2.2. Definitions

ADoT: Authoritative DNS over TLS - the use of DNS over TLS (DoT) when communicating with an authoritative DNS server. eg: between the recursive resolver and the authoritative DNS server (recursive-to-authoritative).  We use DoT when discussing generic DNS over TLS or when referring to encrypted communication between stub and recursive resolver.

Attack Surface: The sum of attack vectors where an unauthorized user (attacker) can try to enter or extract data from the environment or compromise a service via resource starvation.

Authoritative Operator: An operator of an authoritative DNS server.

CDN: Content Delivery Network - distributed network of servers which proxy traffic between content providers and end users in order to provide high availability and high performance.

ECS: EDNS0 Client Subnet [RFC7871] - an extension to EDNS0 where the

client's subnet is included in the DNS query, intended to provide a
hint to authoritative servers who may wish to provide different
answers in an attempt to provide higher performance for end users
based on their network location.

EPSK: External Pre-Shared Key - TLS 1.3 [RFC8446] uses the same PSK
extension for keys established both during handshake (resumption PSK)
and keys established externally.  The EPSK acronym was introduced in
draft [I-D.draft-wood-tls-external-psk-importer-00] in order to
disambiguate External vs Resumption PSKs.

Performance:

o  QRTT: Query Round-Trip Time - the time it takes between sending a
   query and receiving a response.

o  Best Response - whether or not the authoritative server, if
   dynamic responses are used as they are in CDNs, are able to
   determine or infer location and provide the most local response.
   It is a key part of the end-to-end performance for end users to
   get not just _an_ answer quickly but to get the best and most
   local answer.

o  System Performance - the cost in system resources such as CPU/IO/
   Memory.

o  Queries Per Second (QPS) - the maximum number of simultaneous
   queries that a DNS server can handle, on a per second basis.

Authoritative Server - see [RFC8499]

Recursive Resolver - see [RFC8499]

Stub Resolver - see [RFC8499]

TLS: - see [RFC8446]

SUDN: Special-Use Domain Names - see [RFC6761]

## 3.  Key Issues and Questions

## 3.1.  Signaling Support for ADoT

[RFC8310] does not define a method for a nameserver to advertise its
support of DoT other than to have the client make a connection
attempt to the default port of 853.  The extra round-trip to check
for ADoT support imposes a penalty for clients and resolvers that
either do not remember the nameserver or have not communicated with
that nameserver before.  The extra round-trip required may lead some
implementers down a similar path to happy eyeballs [RFC8305] which,
in the case of DNS, would send the same query over both encrypted and
un-encrypted channels at the same time.  A happy eyeballs type

approach, which we'll call "leaky resolvers", would defeat the
purpose of the encryption protection for the testing query, but may
enable subsequent queries to be sent over a private channel with the
first query being subject to on-path adversaries.  An implementation
could use some constant query string as a test query.  However any
query included in the set of queries comprising the iterative
resolution for a QNAME first sent over an encrypted channel that
leaks the original stub QNAME, SHOULD NOT be used.

## 3.2.  Port number

[RFC7858] section 3 [3] indicates that port 853 MUST be used for
session establishment unless otherwise negotiated and configured by
both the client and server.  In the stub-to-recursive connection,
changing the port is something that can be done at stub configuration
time however, managing this negotiation between the recursive-to-
authoritative server is not scalable or standardized.   The
scalability problem is due to the fact that recursive resolvers
communicate with thousands of authoritative servers, therefore port/
service discovery for each of these authoritatives becomes difficult.

Static use of a pre-defined port provides on-path adversaries the
ability to more easily drop or manipulate traffic intended for that
port, possibly triggering resolvers to downgrade a connection back to
a traditional DNS query, eliminating the encryption protections.
This attack is more likely to happen on the stub-to-recursive
connection but is also a possible threat for recursive-to-
authoritative connections.

## 3.3.  TLS version

Implementers of ADoT should read, understand, and follow the guidance
provided in BCP195 [4], also known as [RFC7525], when deploying DoT
on their platforms.  At the time of writing, [RFC7525] did not
include coverage for TLS 1.3.  However, TLS 1.3 should be included in
the document that obsoletes this BCP.  Until this happens, TLS 1.3
SHOULD be preferred over TLS 1.2, as 1.3 offers both security and
performance enhancements.  Additionally, operators should monitor TLS
version issues and cipher suite vulnerabilities for the version of
TLS that their platforms offer.

In the absence of any widespread ADoT deployments, it is easier to
limit TLS version 1.3 or greater.  The absence of widespread adoption
also allows the IETF to create and enforce standards/policies that
ensure TLS versions are kept current going forward.

## 3.4.  Resumptions

TLS resumption allows clients and servers to use information from a
previously established session in order to bootstrap the
cryptographic state while avoiding a full handshake.  The resumption

mechanism is redesigned in TLS 1.3 [RFC8446] section 2.2 [5] and
section 2.3 [6], eliminating both [RFC5077] session tickets and
session ID resumption.

Resumption improves both connection and resource (socket and CPU)
efficiency, therefore operators SHOULD allow for TLS resumption.
However, special consideration should be given to 0-RTT resumption as
it is vulnerable to replay attacks [RFC3552] see Section 3.3.1 [7].
The replay attack may not be as important for DNS, as DNS queries are
generally idempotent.  However consideration should be given to
possible side-channel attacks [8].

## 3.5.  Operational Monitoring

Many operators use external passive monitors in order to understand
the health and performance of their infrastructure.  Infrastructure
monitoring is also often done to retain a copy of traffic for
forensic purposes - such as the BIND "packet of death" [9] scenario.
These legacy monitoring systems may break with the use of TLS 1.3.
Therefore alternatives may need to be deployed/developed in order to
maintain effective operational performance and security monitoring
functionality.

A number of solutions have been suggested:

o  Data Center use of Static Diffie-Hellman in TLS 1.3
   [I-D.draft-green-tls-static-dh-in-tls13-01]

o  TLS Security and Data Center Monitoring: Searching for a Path
   Forward [10]

o  TLS 1.3: Will Your Network Monitoring Go Blind? [11]

## 3.6.  Architecture

Operators often reconfigure their architectural designs to best
deliver a new product offering or service.  Operators should consider
the following design alternatives for the new ADoT service:

o  Operators should consider segregating ADoT addresses from
   traditional DNS over UDP/TCP to enable better attack mitigation,
   better service monitoring, less service interference, and more
   stability.

o  Operators should weigh the pros/cons of using a TLS proxy vs
   direct client-to-host connection.  In case of ADoT, the client is
   most likely a recursive resolver and the host is the authoritative
   host server.

## 3.7.  Socket efficiency/tuning considerations

Operators can realize substantial gains in client session

establishment and improve overall RTT by tuning sockets setting for
best use-case efficiency.

For the ADoT use case, operators should consult [RFC7766] section 6.2
[12] and minimally consider the following:

o  Optimal number of persistent connections - consideration should be
   given to the number of persistent connections maintained for both
   the recursive resolvers and authoritative servers

o  Optimal read/write buffer size

o  Optimal session timeout

o  Optimal close wait state time

o  Optimal connection time/timeout

## 3.8.  Post-Quantum Security

Given that ADoT deployments will likely have a long lifetime and are
being introduced in an era where post-quantum security is now an
important design consideration, it is prudent to consider how
protections against quantum computers might be integrated into the
deployments.

[I-D.draft-hoffman-c2pq-05] outlines the threat quantum computing
presents to classical cryptographic algorithms.

External Pre-Shared Keys (EPSKs) may be less vulnerable to quantum
attacks.  A proposed approach to combining EPSKs and certificates in
TLS is described in
[I-D.draft-housley-tls-tls13-cert-with-extern-psk-03].

## 4.  Suggestions for further research and development

## 4.1.  Required studies and analysis

Unlike stub-to-recursive DNS communication, authoritative nameservers
affect users in ways that end users cannot avoid or work around.  In
the event that all authoritative servers for a zone are unreachable,
the zone becomes globally unavailable.  Hence, in order to preserve
stability and resiliency of authoritative nameservers when deploying
ADoT, more empirical studies and analysis MUST be conducted.  The
following list is a minimal set of studies and considerations that
need to be conducted/addressed in order to maintain authoritative
stability and resilience.

o  Attack vectors and mitigation: consider the new adversarial powers
   enabled by ADoT - types of attacks and denial of service, or other
   security challenges that are created with the addition of ADoT to
   authoritative nameservers.

o  Traffic: consider how traffic patterns to authoritative
   nameservers change with the introduction of ADoT and how these
   traffic patterns change when the parameters of the service are
   changed; e.g. persistent connection lifetime, TLS connection
   parameters, use of TLS session tickets [RFC5077] or Pre-Shared Key
   extension in TLS 1.3 [RFC8446] section 2.2 [13].  Consider how
   these traffic pattern changes will affect the architecture and
   infrastructure for authoritative operators.

o  ADoT capacity and footprint expansion: consider how common scaling
   techniques impact authoritative operators; e.g. anycast, load
   balancing, custom hardware.

o  DTLS/UDP - consider if there is any reason to implement DTLS given
   that we lose the benefit of pipelining requests and must drop back
   to TLS/TCP in the case of fragmentation.

It is critical to conduct large-scale measurements of DNS
infrastructure in order to quantify some of the scalability issues.
While these tests may be performed initially in a controlled lab
environment, the public Internet is fundamentally more variable.
Therefore, global testing at scale on the Internet MUST also be
conducted in order to understand and measure potential issues which
must be overcome before full global deployment can occur.

## 4.2.  Authoritative DNS over TLS (ADoT) Profile

Profiles can be used as a mechanism to help mitigate operational
concerns over increased attack surface by restricting features such
as computationally expensive processes, insecure ciphers, general
starvation vectors, or other features that may limit operational
performance.

Therefore, an ADoT application profile draft, taking into account the
conclusions of required studies and analysis, may help assuage some
of the concerns raised in this document.

## 5.  Security Considerations

In addition to the applicable security considerations described in
RFCs [RFC7626] and [RFC8310], considerations focused on future
deployment of quantum computers are described in Post-Quantum
Security (Section 3.8).  Additional considerations associated with
ADoT are TBD based on working group discussions.

## 6.  References

## 6.1.  Informative References

[I-D.draft-green-tls-static-dh-in-tls13-01]
          Green, M., Droms, R., Housley, R., Turner, P., and S.

               Fenter, "Data Center use of Static Diffie-Hellman in TLS
               1.3", draft-green-tls-static-dh-in-tls13-01 (work in
               progress), July 2017.

     [I-D.draft-hoffman-c2pq-05]
               Hoffman, P., "The Transition from Classical to Post-
               Quantum Cryptography", draft-hoffman-c2pq-05 (work in
               progress), May 2019.

     [I-D.draft-housley-tls-tls13-cert-with-extern-psk-03]
               Housley, R., "TLS 1.3 Extension for Certificate-based
               Authentication with an External Pre-Shared Key", draft-
               housley-tls-tls13-cert-with-extern-psk-03 (work in
               progress), November 2018.

     [I-D.draft-wood-tls-external-psk-importer-00]
               Wood, C., "Importing External PSKs for TLS 1.3", draft-
               wood-tls-external-psk-importer-00 (work in progress),
               October 2018.

     [RFC2119]  Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119,
               DOI 10.17487/RFC2119, March 1997,
               <https://www.rfc-editor.org/info/rfc2119>.

     [RFC3552]  Rescorla, E. and B. Korver, "Guidelines for Writing RFC
               Text on Security Considerations", BCP 72, RFC 3552,
               DOI 10.17487/RFC3552, July 2003,
               <https://www.rfc-editor.org/info/rfc3552>.

     [RFC5077]  Salowey, J., Zhou, H., Eronen, P., and H. Tschofenig,
               "Transport Layer Security (TLS) Session Resumption without
               Server-Side State", RFC 5077, DOI 10.17487/RFC5077,
               January 2008, <https://www.rfc-editor.org/info/rfc5077>.

     [RFC6761]  Cheshire, S. and M. Krochmal, "Special-Use Domain Names",
               RFC 6761, DOI 10.17487/RFC6761, February 2013,
               <https://www.rfc-editor.org/info/rfc6761>.

     [RFC7525]  Sheffer, Y., Holz, R., and P. Saint-Andre,
               "Recommendations for Secure Use of Transport Layer
               Security (TLS) and Datagram Transport Layer Security
               (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May
               2015, <https://www.rfc-editor.org/info/rfc7525>.

     [RFC7626]  Bortzmeyer, S., "DNS Privacy Considerations", RFC 7626,
               DOI 10.17487/RFC7626, August 2015,
               <https://www.rfc-editor.org/info/rfc7626>.

     [RFC7766]  Dickinson, J., Dickinson, S., Bellis, R., Mankin, A., and
               D. Wessels, "DNS Transport over TCP - Implementation
               Requirements", RFC 7766, DOI 10.17487/RFC7766, March 2016,

                   <https://www.rfc-editor.org/info/rfc7766>.

   [RFC7816]  Bortzmeyer, S., "DNS Query Name Minimisation to Improve
              Privacy", RFC 7816, DOI 10.17487/RFC7816, March 2016,
              <https://www.rfc-editor.org/info/rfc7816>.

   [RFC7858]  Hu, Z., Zhu, L., Heidemann, J., Mankin, A., Wessels, D.,
              and P. Hoffman, "Specification for DNS over Transport
              Layer Security (TLS)", RFC 7858, DOI 10.17487/RFC7858, May
              2016, <https://www.rfc-editor.org/info/rfc7858>.

   [RFC7871]  Contavalli, C., van der Gaast, W., Lawrence, D., and W.
              Kumari, "Client Subnet in DNS Queries", RFC 7871,
              DOI 10.17487/RFC7871, May 2016,
              <https://www.rfc-editor.org/info/rfc7871>.

   [RFC8174]  Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [RFC8198]  Fujiwara, K., Kato, A., and W. Kumari, "Aggressive Use of
              DNSSEC-Validated Cache", RFC 8198, DOI 10.17487/RFC8198,
              July 2017, <https://www.rfc-editor.org/info/rfc8198>.

   [RFC8305]  Schinazi, D. and T. Pauly, "Happy Eyeballs Version 2:
              Better Connectivity Using Concurrency", RFC 8305,
              DOI 10.17487/RFC8305, December 2017,
              <https://www.rfc-editor.org/info/rfc8305>.

   [RFC8310]  Dickinson, S., Gillmor, D., and T. Reddy, "Usage Profiles
              for DNS over TLS and DNS over DTLS", RFC 8310,
              DOI 10.17487/RFC8310, March 2018,
              <https://www.rfc-editor.org/info/rfc8310>.

   [RFC8446]  Rescorla, E., "The Transport Layer Security (TLS) Protocol
              Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018,
              <https://www.rfc-editor.org/info/rfc8446>.

   [RFC8499]  Hoffman, P., Sullivan, A., and K. Fujiwara, "DNS
              Terminology", BCP 219, RFC 8499, DOI 10.17487/RFC8499,
              January 2019, <https://www.rfc-editor.org/info/rfc8499>.

## 6.2.  URIs

   [1] https://wiki.mozilla.org/Security/DOH-resolver-
       policy#Privacy_Requirements

   [2] https://tools.ietf.org/html/rfc7871#section-12.1

   [3] https://tools.ietf.org/html/rfc7858#section-3

   [4] https://tools.ietf.org/html/bcp195

[5]    https://tools.ietf.org/html/rfc8446#section-2.2

[6]    https://tools.ietf.org/html/rfc8446#section-2.3

[7]    https://tools.ietf.org/html/rfc3552#section-3.3.1

[8]    https://eprint.iacr.org/2005/388.pdf

[9]    https://www.nominet.uk/the-packet-of-death/

[10]   https://www.rsa.com/en-us/blog/2017-08/tls-security-and-data-
       center-monitoring-searching-for-a-path-forward

[11]   https://www.extrahop.com/company/blog/2018/maintain-visibility-
       with-tls-1.3/

[12]   https://tools.ietf.org/html/rfc7766#section-6.2

[13]   https://tools.ietf.org/html/rfc8446#section-2.2

## Appendix A.  Acknowledgements

Thanks to those that provided usage data, reviewed and/or improved
this document, including: Piet Barber, Michael Bentkofsky, David
Blacka, Florent Guiliani, Scott Hollenbeck, Burt Kaliski, Glen Wiley,
and Richard Wilhelm.

## Appendix B.  Change Log

RFC EDITOR:  PLEASE REMOVE THE THIS SECTION PRIOR TO PUBLICATION.

TODO:  Zero this change log out when -00 is submitted to IETF.

pre-00

o  Initial draft.

~~~ 01234567890123456789012345678901234567890123456789012345678901234
5678912

Authors' Addresses

Karl Henderson
Verisign

Email: khenderson@verisign.com


Tim April
Akamai

Email: ietf@tapril.net

Jason Livingood
Comcast

Email: jason_livingood@comcast.com