                    Split DNS Configuration for IKEv2
                     draft-ietf-ipsecme-split-dns-10

Abstract

   This document defines two Configuration Payload Attribute Types for
   the IKEv2 protocol that add support for private DNS domains. These
   domains are intended to be resolved using DNS servers reachable
   through an IPsec connection, while leaving all other DNS resolution
   unchanged. This approach of resolving a subset of domains using non-
   public DNS servers is referred to as "Split DNS".

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 19, 2019.

Table of Contents

1. Introduction

Split DNS is a common configuration for secure tunnels, such as
Virtual Private Networks in which host machines private to an
organization can only be resolved using internal DNS resolvers
[RFC2775]. In such configurations, it is often desirable to only
resolve hosts within a set of private domains using the tunnel, while
letting resolutions for public hosts be handled by a device's default
DNS configuration.

The Internet Key Exchange protocol version 2 [RFC7296] negotiates
configuration parameters using Configuration Payload Attribute Types.
This document defines two Configuration Payload Attribute Types that
add support for trusted Split DNS domains.

The INTERNAL_DNS_DOMAIN attribute type is used to convey one or more
DNS domains that SHOULD be resolved only using the provided DNS
nameserver IP addresses, causing these requests to use the IPsec
connection.

The INTERNAL_DNSSEC_TA attribute type is used to convey DNSSEC trust
anchors for those domains.

When only a subset of traffic is routed into a private network using
an IPsec SA, these Configuration Payload options can be used to
define which private domains are intended to be resolved through the
IPsec connection without affecting the client's global DNS
resolution.

For the purposes of this document, DNS resolution servers accessible
through an IPsec connection will be referred to as "internal DNS
servers", and other DNS servers will be referred to as "external DNS
servers".

A client using these configuration payloads will be able to request
and receive Split DNS configurations using the INTERNAL_DNS_DOMAIN
and INTERNAL_DNSSEC_TA configuration attributes. The client device
can use the internal DNS server(s) for any DNS queries within the
assigned domains. DNS queries for other domains SHOULD be sent to
the regular external DNS server.

1.1. 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
captials, as shown here.

2. Background

Split DNS is a common configuration for enterprise VPN deployments,
in which only one or a few private DNS domains are accessible and
resolvable via an IPsec based VPN connection.

Other tunnel-establishment protocols already support the assignment
of Split DNS domains. For example, there are proprietary extensions
to IKEv1 that allow a server to assign Split DNS domains to a client.
However, the IKEv2 standard does not include a method to configure
this option. This document defines a standard way to negotiate this
option for IKEv2.

3. Protocol Exchange

In order to negotiate which domains are considered internal to an
IKEv2 tunnel, initiators indicate support for Split DNS in their
CFG_REQUEST payloads, and responders assign internal domains (and
DNSSEC trust anchors) in their CFG_REPLY payloads. When Split DNS

has been negotiated, the existing DNS server configuration attributes
will be interpreted as internal DNS servers that can resolve
hostnames within the internal domains.

3.1. Configuration Request

To indicate support for Split DNS, an initiator includes one more
INTERNAL_DNS_DOMAIN attributes as defined in Section 4 as part of the
CFG_REQUEST payload. If an INTERNAL_DNS_DOMAIN attribute is included
in the CFG_REQUEST, the initiator SHOULD also include one or more
INTERNAL_IP4_DNS and INTERNAL_IP6_DNS attributes in the CFG_REQUEST.

The INTERNAL_DNS_DOMAIN attribute sent by the initiator is usually
empty but MAY contain a suggested domain name.

The absence of INTERNAL_DNS_DOMAIN attributes in the CFG_REQUEST
payload indicates that the initiator does not support or is unwilling
to accept Split DNS configuration.

To indicate support for DNSSEC, an initiator includes one or more
INTERNAL_DNSSEC_TA attributes as defined in Section 4 as part of the
CFG_REQUEST payload. If an INTERNAL_DNSSEC_TA attribute is included
in the CFG_REQUEST, the initiator SHOULD also include one or more
INTERNAL_DNS_DOMAIN attributes in the CFG_REQUEST. If the initiator
includes an INTERNAL_DNSSEC_TA attribute, but does not inclue an
INTERNAL_DNS_DOMAIN attribute, the responder MAY still respond with
both INTERNAL_DNSSEC_TA and INTERNAL_DNS_DOMAIN attributes.

An initiator MAY convey its current DNSSEC trust anchors for the
domain specified in the INTERNAL_DNS_DOMAIN attribute. If it does
not wish to convey this information, it MUST use a length of 0.

The absence of INTERNAL_DNSSEC_TA attributes in the CFG_REQUEST
payload indicates that the initiator does not support or is unwilling
to accept DNSSEC trust anchor configuration.

3.2. Configuration Reply

Responders MAY send one or more INTERNAL_DNS_DOMAIN attributes in
their CFG_REPLY payload. If an INTERNAL_DNS_DOMAIN attribute is
included in the CFG_REPLY, the responder MUST also include one or
both of the INTERNAL_IP4_DNS and INTERNAL_IP6_DNS attributes in the
CFG_REPLY. These DNS server configurations are necessary to define
which servers can receive queries for hostnames in internal domains.
If the CFG_REQUEST included an INTERNAL_DNS_DOMAIN attribute, but the
CFG_REPLY does not include an INTERNAL_DNS_DOMAIN attribute, the
initiator SHOULD behave as if Split DNS configurations are not
supported by the server.

Each INTERNAL_DNS_DOMAIN represents a domain that the DNS servers
address listed in INTERNAL_IP4_DNS and INTERNAL_IP6_DNS can resolve.

If the CFG_REQUEST included INTERNAL_DNS_DOMAIN attributes with non-
zero lengths, the content MAY be ignored or be interpreted as a
suggestion by the responder.

For each DNS domain specified in an INTERNAL_DNS_DOMAIN attribute,
one or more INTERNAL_DNSSEC_TA attributes MAY be included by the
responder. This attribute lists the corresponding internal DNSSEC
trust anchor in the DNS presentation format of a DS record as
specified in [RFC4034]. The INTERNAL_DNSSEC_TA attribute MUST
immediately follow the INTERNAL_DNS_DOMAIN attribute that it applies
to.

3.3. Mapping DNS Servers to Domains

All DNS servers provided in the CFG_REPLY MUST support resolving
hostnames within all INTERNAL_DNS_DOMAIN domains. In other words,
the INTERNAL_DNS_DOMAIN attributes in a CFG_REPLY payload form a
single list of Split DNS domains that applies to the entire list of
INTERNAL_IP4_DNS and INTERNAL_IP6_DNS attributes.

3.4. Example Exchanges

3.4.1. Simple Case

In this example exchange, the initiator requests INTERNAL_IP4_DNS and
INTERNAL_DNS_DOMAIN attributes in the CFG_REQUEST, but does not
specify any value for either. This indicates that it supports Split
DNS, but has no preference for which DNS requests will be routed
through the tunnel.

The responder replies with two DNS server addresses, and two internal
domains, "example.com" and "city.other.com".

Any subsequent DNS queries from the initiator for domains such as
"www.example.com" SHOULD use 198.51.100.2 or 198.51.100.4 to resolve.

```
CP(CFG_REQUEST) =
  INTERNAL_IP4_ADDRESS()
  INTERNAL_IP4_DNS()
  INTERNAL_DNS_DOMAIN()

CP(CFG_REPLY) =
  INTERNAL_IP4_ADDRESS(198.51.100.234)
  INTERNAL_IP4_DNS(198.51.100.2)
  INTERNAL_IP4_DNS(198.51.100.4)
  INTERNAL_DNS_DOMAIN(example.com)
  INTERNAL_DNS_DOMAIN(city.other.com)
```

3.4.2. Requesting Domains and DNSSEC trust anchors

In this example exchange, the initiator requests INTERNAL_IP4_DNS,
INTERNAL_DNS_DOMAIN and INTERNAL_DNSSEC_TA attributes in the
CFG_REQUEST.

Any subsequent DNS queries from the initiator for domains such as
"www.example.com" or "city.other.com" would be DNSSEC validated using
the DNSSEC trust anchor received in the CFG_REPLY.

In this example, the initiator has no existing DNSSEC trust anchors
would the requested domain. the "example.com" dommain has DNSSEC
trust anchors that are returned, while the "other.com" domain has no
DNSSEC trust anchors.

```
CP(CFG_REQUEST) =
  INTERNAL_IP4_ADDRESS()
  INTERNAL_IP4_DNS()
  INTERNAL_DNS_DOMAIN()
  INTERNAL_DNSSEC_TA()

CP(CFG_REPLY) =
  INTERNAL_IP4_ADDRESS(198.51.100.234)
  INTERNAL_IP4_DNS(198.51.100.2)
  INTERNAL_IP4_DNS(198.51.100.4)
  INTERNAL_DNS_DOMAIN(example.com)
  INTERNAL_DNSSEC_TA(43547,8,1,B6225AB2CC613E0DCA7962BDC2342EA4...)
  INTERNAL_DNSSEC_TA(31406,8,2,F78CF3344F72137235098ECBBD08947C...)
  INTERNAL_DNS_DOMAIN(city.other.com)
```

4. Payload Formats

All multi-octet fields representing integers are laid out in big
endian order (also known as "most significant byte first", or
"network byte order").

4.1. INTERNAL_DNS_DOMAIN Configuration Attribute Type Request and Reply

```
                    1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-----------------------------+-----------------------------+
|R|        Attribute Type       |            Length           |
+-+-----------------------------+-----------------------------+
|                                                             |
~           Domain Name in DNS presentation format            ~
|                                                             |
+-------------------------------------------------------------+
```
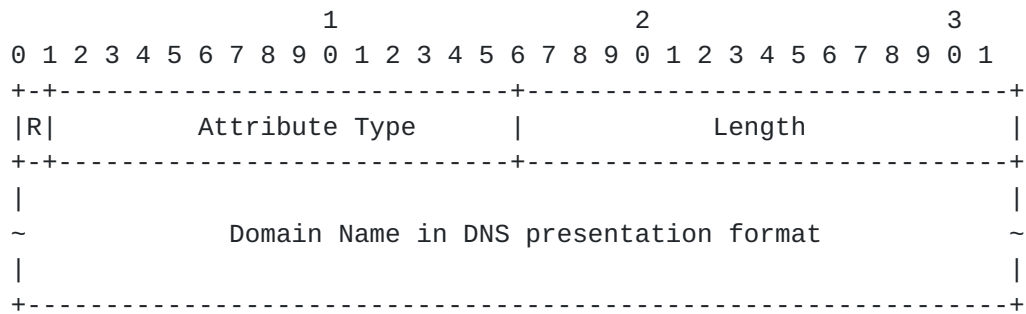
o Reserved (1 bit) - Defined in IKEv2 RFC [RFC7296].

o Attribute Type (15 bits) set to value 25 for INTERNAL_DNS_DOMAIN.

o Length (2 octets) - Length of domain name.

o Domain Name (0 or more octets) - A Fully Qualified Domain Name
   used for Split DNS rules, such as "example.com", in DNS
   presentation format and optionally using IDNA [RFC5890] for
   Internationalized Domain Names. Implementors need to be careful
   that this value is not null-terminated.

4.2. INTERNAL_DNSSEC_TA Configuration Attribute

An INTERNAL_DNSSEC_TA Configuration Attribute can either be empty, or
it can contain one Trust Anchor by containing a non-zero Length with
a DNSKEY Key Tag, DNSKEY Algorithm, Digest Type and Digest Data
fields.

An empty INTERNAL_DNSSEC_TA CFG attribute:

```
                      1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-----------------------------+-------------------------------+
 |R|        Attribute Type       |        Length (set to 0)      |
 +-+-----------------------------+-------------------------------+
```

A non-empty INTERNAL_DNSSEC_TA CFG attribute:

```
                      1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-----------------------------+-------------------------------+
 |R|        Attribute Type       |             Length            |
 +-+-----------------------------+---------------+---------------+
 |          DNSKEY Key Tag       |   DNSKEY Alg  |  Digest Type  |
 +-------------------------------+---------------+---------------+
 |                                                               |
 ~                          Digest Data                          ~
 |                                                               |
 +---------------------------------------------------------------+
```

o Reserved (1 bit) - Defined in IKEv2 RFC [RFC7296].

o Attribute Type (15 bits) set to value 26 for INTERNAL_DNSSEC_TA.

o Length (0 or 2 octets) - Length of DNSSEC Trust Anchor data (4
   octets plus the length of the Digest Data).

o DNSKEY Key Tag value (0 or 2 octets) - Delegation Signer (DS) Key
   Tag as specified in [RFC4034] Section 5.1.

o DNSKEY Algorithm (0 or 1 octet) - DNSKEY algorithm value from the
   IANA DNS Security Algorithm Numbers Registry.

o Digest Type (0 or 1 octet) - DS algorithm value from the IANA
   Delegation Signer (DS) Resource Record (RR) Type Digest Algorithms
   Registry.

o Digest Data (0 or more octets) - The DNSKEY digest as specified in
   [RFC4034] Section 5.1 in presentation format.

INTERNAL_DNSSEC_TA payloads MUST immediately follow an
INTERNAL_DNS_DOMAIN payload. As the INTERNAL_DNSSEC_TA format itself
does not contain the domain name, it relies on the preceding
INTERNAL_DNS_DOMAIN to provide the domain for which it specifies the
trust anchor.

5. **INTERNAL_DNS_DOMAIN Usage Guidelines**

   If a CFG_REPLY payload contains no INTERNAL_DNS_DOMAIN attributes,
   the client MAY use the provided INTERNAL_IP4_DNS or INTERNAL_IP6_DNS
   servers as the default DNS server(s) for all queries.

   If a client is configured by local policy to only accept a limited
   number of INTERNAL_DNS_DOMAIN values, the client MUST ignore any
   other INTERNAL_DNS_DOMAIN values.

   For each INTERNAL_DNS_DOMAIN entry in a CFG_REPLY payload that is not
   prohibited by local policy, the client MUST use the provided
   INTERNAL_IP4_DNS or INTERNAL_IP6_DNS DNS servers as the only
   resolvers for the listed domains and its sub-domains and it MUST NOT
   attempt to resolve the provided DNS domains using its external DNS
   servers.

   If the initiator host is configured to block DNS answers containing
   IP addresses from special IP address ranges such as those of
   [RFC1918], the initiator SHOULD allow the DNS domains listed in the
   INTERNAL_DNS_DOMAIN attributes to contain those Special IP addresses.

   If a CFG_REPLY contains one or more INTERNAL_DNS_DOMAIN attributes
   and its local policy does not forbid these values, the client MUST
   configure its DNS resolver to resolve those domains and all their
   subdomains using only the DNS resolver(s) listed in that CFG_REPLY
   message. If those resolvers fail, those names MUST NOT be resolved
   using any other DNS resolvers. Other domain names SHOULD be resolved
   using some other external DNS resolver(s), configured independently
   from IKE. Queries for these other domains MAY be sent to the
   internal DNS resolver(s) listed in that CFG_REPLY message, but have
   no guarantee of being answered. For example, if the
   INTERNAL_DNS_DOMAIN attribute specifies "example.com", then
   "example.com", "www.example.com" and "mail.eng.example.com" MUST be
   resolved using the internal DNS resolver(s), but "anotherexample.com"
   and "ample.com" SHOULD NOT be resolved using the internal resolver
   and SHOULD use the system's external DNS resolver(s).

   When an IKE SA is terminated, the DNS forwarding MUST be
   unconfigured. This includes deleting the DNS forwarding rules;
   flushing all cached data for DNS domains provided by the
   INTERNAL_DNS_DOMAIN attribute, including negative cache entries;
   removing any obtained DNSSEC trust anchors from the list of trust
   anchors; and clearing the outstanding DNS request queue.

   INTERNAL_DNS_DOMAIN attributes SHOULD only be used on split tunnel
   configurations where only a subset of traffic is routed into a
   private remote network using the IPsec connection. If all traffic is

   routed over the IPsec connection, the existing global
   INTERNAL_IP4_DNS and INTERNAL_IP6_DNS can be used without creating
   specific DNS exemptions.

[6](#). **INTERNAL_DNSSEC_TA Usage Guidelines**

   DNS records can be used to publish specific records containing trust
   anchors for applications. The most common record type is the TLSA
   record specified in [[RFC6698](#)]. This DNS record type publishes which
   CA certificate or EE certificate to expect for a certain host name.
   These records are protected by DNSSEC and thus can be trusted by the
   application. Whether to trust TLSA records instead of the
   traditional WebPKI depends on the local policy of the client. By
   accepting an INTERNAL_DNSSEC_TA trust anchor via IKE from the remote
   IKE server, the IPsec client might be allowing the remote IKE server
   to override the trusted certificates for TLS. Similar override
   concerns apply to other public key or fingerprint based DNS records,
   such as OPENPGPKEY, SMIMEA or IPSECKEY records.

   Thus, installing an INTERNAL_DNSSEC_TA trust anchor can be seen as
   the equivalent of installing an Enterprise Certificate Agency (CA)
   certificate. It allows the remote IKE/IPsec server to modify DNS
   answers including its DNSSEC cryptographic signatures by overriding
   existing DNS information with trust anchor conveyed via IKE and
   (temporarilly) installed on the IKE client. Of specific concern is
   the overriding of [[RFC6698](#)] based TLSA records, which represent a
   confirmation or override of an existing WebPKI TLS certificate.
   Other DNS record types that convey cryptographic materials (public
   keys or fingerprints) are OPENPGPKEY, SMIMEA, SSHP and IPSECKEY
   records.

   IKE clients MUST use a preconfigured whitelist of one or more domain
   names for which it will allow INTERNAL_DNSSEC_TA updates. This list
   may be sent in the CFG_REQUEST payload, or may be applied after
   reception of the CFG_REPLY payload.

   IKE clients should take care to only whitelist domains that apply to
   internal or managed domains, rather than to generic Internet traffic.
   The DNS root zone (".") MUST NOT be whitelisted. Other generic or
   public domains, such as top-level domains, similarly SHOULD NOT be
   whitelisted.

   Any updates to this whitelist of domain names MUST happen via
   explicit human interaction to prevent invisible installation of trust
   anchors.

   IKE clients SHOULD accept any INTERNAL_DNSSEC_TA updates for
   subdomain names of the whitelisted domain names. For example, if

"example.net" is whitelisted, then INTERNAL_DNSSEC_TA received for
"antartica.example.net" SHOULD be accepted.

IKE clients MAY interpret an INTERNAL_DNSSEC_TA for domain that was
not preconfigured as an indication that it needs to update its IKE
configuration (out of band). The client MUST NOT use such a
INTERNAL_DNSSEC_TA to reconfigure its local DNS settings.

IKE clients MUST ignore any received INTERNAL_DNSSEC_TA requests for
a FDQN for which it did not receive and accept an INTERNAL_DNS_DOMAIN
Configuration Payload.

In most deployment scenario's, the IKE client has an expectation that
it is connecting, using a split-network setup, to a specific
organisation or enterprise. A recommended policy would be to only
accept INTERNAL_DNSSEC_TA directives from that organization's DNS
names. However, this might not be possible in all deployment
scenarios, such as one where the IKE server is handing out a number
of domains that are not within one parent domain.

7. Security Considerations

The use of Split DNS configurations assigned by an IKEv2 responder is
predicated on the trust established during IKE SA authentication.
However, if IKEv2 is being negotiated with an anonymous or unknown
endpoint (such as for Opportunistic Security [RFC7435]), the
initiator MUST ignore Split DNS configurations assigned by the
responder.

If a host connected to an authenticated IKE peer is connecting to
another IKE peer that attempts to claim the same domain via the
INTERNAL_DNS_DOMAIN attribute, the IKE connection SHOULD only process
the DNS information if the two connections are part of the same
logical entity. Otherwise, the client SHOULD refuse the DNS
information and potentially warn the end-user.

If the initiator is using DNSSEC validation for a domain in its
public DNS view, and it requests and receives an INTERNAL_DNS_DOMAIN
attribute without an INTERNAL_DNSSEC_TA, it will need to reconfigure
its DNS resolver to allow for an insecure delegation. It SHOULD NOT
accept insecure delegations for domains that are DNSSEC signed in the
public DNS view, for which it has not explicitly requested such
deletation by specifying the domain specifically using a
INTERNAL_DNS_DOMAIN(domain) request.

Deployments that configure INTERNAL_DNS_DOMAIN domains should pay
close attention to their use of indirect reference RRtypes such as

CNAME, DNAME, MX or SRV records so that resolving works as intended
when all, some, or none of the IPsec connections are established.

The content of INTERNAL_DNS_DOMAIN and INTERNAL_DNSSEC_TA may be
passed to another (DNS) program for processing. As with any network
input, the content SHOULD be considered untrusted and handled
accordingly.

8. IANA Considerations

This document defines two new IKEv2 Configuration Payload Attribute
Types, which are allocated from the "IKEv2 Configuration Payload
Attribute Types" namespace.

		Multi-		
Value	Attribute Type	Valued	Length	Reference
25	INTERNAL_DNS_DOMAIN	YES	0 or more	[this document]
26	INTERNAL_DNSSEC_TA	YES	0 or more	[this document]

Figure 1

9. References

9.1. Normative References

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G.,
           and E. Lear, "Address Allocation for Private Internets",
           BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996,
           <https://www.rfc-editor.org/info/rfc1918>.

[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>.

[RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S.
           Rose, "Resource Records for the DNS Security Extensions",
           RFC 4034, DOI 10.17487/RFC4034, March 2005,
           <https://www.rfc-editor.org/info/rfc4034>.

[RFC5890] Klensin, J., "Internationalized Domain Names for
           Applications (IDNA): Definitions and Document Framework",
           RFC 5890, DOI 10.17487/RFC5890, August 2010,
           <https://www.rfc-editor.org/info/rfc5890>.

   [RFC6698] Hoffman, P. and J. Schlyter, "The DNS-Based Authentication
              of Named Entities (DANE) Transport Layer Security (TLS)
              Protocol: TLSA", RFC 6698, DOI 10.17487/RFC6698, August
              2012, <https://www.rfc-editor.org/info/rfc6698>.

   [RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T.
              Kivinen, "Internet Key Exchange Protocol Version 2
              (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October
              2014, <https://www.rfc-editor.org/info/rfc7296>.

   [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>.

9.2. Informative References

   [RFC2775] Carpenter, B., "Internet Transparency", RFC 2775,
              DOI 10.17487/RFC2775, February 2000,
              <https://www.rfc-editor.org/info/rfc2775>.

   [RFC7435] Dukhovni, V., "Opportunistic Security: Some Protection
              Most of the Time", RFC 7435, DOI 10.17487/RFC7435,
              December 2014, <https://www.rfc-editor.org/info/rfc7435>.

Authors' Addresses

   Tommy Pauly
   Apple Inc.
   One Apple Park Way
   Cupertino, California 95014
   US

   Email: tpauly@apple.com

   Paul Wouters
   Red Hat

   Email: pwouters@redhat.com