Network Working Group   B. Weis
Internet-Draft   Cisco Systems
Intended status: Standards Track   U. Mangla
Expires: June 21, 2014   N. Maheshwari
                                                 Juniper Networks Inc.
                                                               T. Karl
                                                       Deutsche Telekom
                                                     December 18, 2013

IP Delivery Delay Detection Protocol
draft-weis-delay-detection-00

Abstract

   This memo describes a method that indicates to the receiver of an IP
   Internet Protocol datagram the time at which it was sent. The
   receiver of the datagram can then judge how it was recently it was
   sent and discard packets deemed to be 'too old' according to their
   policy.

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 http://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 June 21, 2014.

Table of Contents

1. Introduction

   IP datagrams can be subject to a delivery delay attack, where a host
   or gateway receives datagrams that are not fresh. A fresh datagram
   is defined as one "Recently generated; not replayed from some earlier
   interaction of the protocol." [RFC4949]. Even when IP datagrams are
   validated legitimate using an integrity transform and are protected
   from replay, delivered datagrams may have been delayed for an
   unbounded period of time.

   Delivery delay detection is useful when the value of data included in
   the IP datagram is time sensitive. Delivery delay protection is even
   more valuable when replay protection is not available. For example,
   while the IPsec [RFC4301] transforms IP Encapsulating Security
   Payload (ESP) [RFC4303] and IP Authentication Header [RFC4302]
   include replay protection, it must be disabled when multi-sender
   IPsec security associations are used to protect multicast and group
   communications. Such Many-to-Many Applications [RFC3170] often
   require the use of multi-sender security associations, where
   receivers cannot use the sequence number replay protection method.
   In such cases, a single counter cannot record responses from multiple
   senders, and no provision is made for multiple counters in the
   security association.

   This memo describes an IP Delivery Delay Detection Protocol using
   timestamps to declare the age of an IP datagram, enabling receivers
   to make a judgement whether to accept an IP datagram as "fresh"
   [RFC4949].

1.1. Requirements notation

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in [RFC2119].

2. IP Delivery Delay Detection Protocol (IP-D3P)

An IP-D3P datagram consists of a header and an IP payload. The IP-
D3P header includes a timestamp, which receivers of the packet use to
determine whether or not the packet has been recently generated.
Receivers compare the timestamp delivered in the IP packet to their
local time and make a determination as to whether it should be
accepted. Freshness does not provide replay protection, as it makes
no explicit judgement as to whether a receiver has received a
particular packet before. Rather, it allows the receiver to
determine whether the packet has been delayed in delivery beyond an
acceptable point in time. A typical policy would be to choose a time
larger than a reasonable delivery time, which delay indicates a
possible packet delay attack. When combined with replay protection,
IP-D3P provides the receiver assurance that the packet is not only
unique, but also fresh.

Varying formats of timestamps are possible. This memo declares
several formats with varying degrees of timestamp precision. It is
important that both sender and receiver interpret the time in the
same degree of accuracy, else the receiver determination of freshness
may be in error. For example, if the sender reported a timestamp in
seconds and he receiver interpreted the timestamp in microseconds it
would likely reject many datagrams due to the precision error.

2.1. IP-D3P Header

The IP-D3P header consists of a type and a fixed-size timestamp,
followed by a datagram, which may be either a full or partial IP
datagram depending on the use case.

```
 0                   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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-!
! Next Header   !     Type      !          RESERVED            !
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
!                         Timestamp                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-!
```
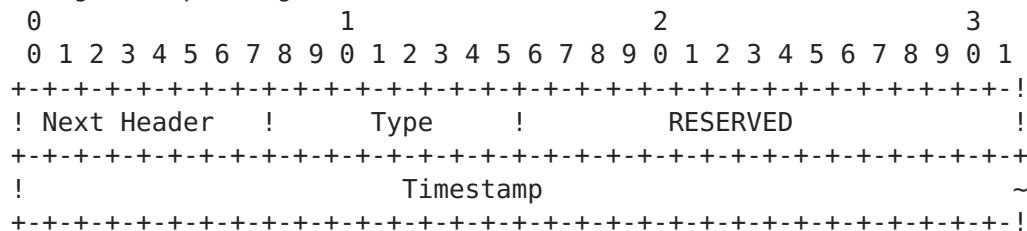
Figure 1: IP-D3P Header

The header fields are defined as follows:

o Next Header (1 octet) -- Identifies the type of data in the
   Datagram following the IP-D3P header. Valid types are described
   in [PROT-REG].

o Type (1 octet) -- Type of timestamp contained in the header.
   Valid values are defined in Section 2.1.1.

o RESERVED (2 octets) -- MUST be zero, and MUST be ignored on
   receipt.

o Timestamp (variable) -- Timestamp declaring when the packet was
   originated. Length depends on type of timestamp.

2.1.1. Type

This memo defines the following types. A conforming implementation
MUST implement POSIX-TIME-SEC. Each type indicates a fixed sized
timestamp, which is valuable for packet evaluation efficiency. The
length of the header can then also be trivially determined based on
its Type.

o POSIX-TIME-SEC. Specifies that the timestamp refers to current
   time as seconds since the POSIX Epoch [POSIX.1] (i..e., time as
   the number of seconds beginning January 1, 1970 not counting leap
   seconds. The timestamp has a length of 4 bytes.

o POSIX-TIME-DECISEC. Specifies that timestamp refers to current
   time as seconds since the POSIX Epoch, specified in deciseconds
   (tenths of a second). The timestamp has a length of 6 bytes.

o POSIX-TIME-MSEC. Specifies that the timestamp refers to current
   time as seconds since the POSIX Epoch, specified in milliseconds.
   The timestamp has a length of 8 bytes.

2.2. Outbound Packet Processing

An IP-D3P header is added to the packet as described in Figure 1.
The sender inserts the type defined by policy into the Type field and
places the current time value in the Timestamp field as specified by
the Type. For example, when the Type is POSIX-TIME-SEC, the current
value of seconds from the system clock is translated into POSIX time
as defined in Section 2.1.1 and placed into the header.

When IP-D3P is added to an IP packet, the IP-D3P Next Header field
identifies the Protocol Id representing the next header (e.g., TCP).
When IP-D3P is used with a IP-IP encapsulation (e.g., IPsec Tunnel
Mode), the IP-D3P Next Header field represents the type of datagram
following (e.g., 4 for IPv4, 41 for IPv6).

2.3. Inbound Packet Processing

A receiver first validates that the Type value in the IP-D3P header
matches the expected type, as according to policy. It then extracts
the timestamp and compares it to a sliding window maintained by the
receiver that is based on the Type of time. A timestamp found within

the sliding window is accepted, and the packet is treated as a fresh
packet (e.g., is processed further). If the timestamp is outside of
the receiver's window, the packet is discarded.

Figure 2 shows the sliding window used by D3P. A receiver chooses a
window size of W, which lies between Ws and We centered around the
current time of the receiver (Tr). When a packet is received with a
Timestamp Ts that lies below Ws, it is rejected as being too old. If
Ts lies in between Ws and We it is accepted as a fresh packet. If Ts
is in advance of We it is rejected as an invalid time. However, the
reason for rejection (being too old or invalid) MUST NOT be
discernible to any entity other than the receiver who rejected the
packet.

```
         Reject  | Accept |   Reject
        (Replay) |        | (Invalid)
       ------------------------------->
        t           Ws   |   We
                         |
                         Tr
```
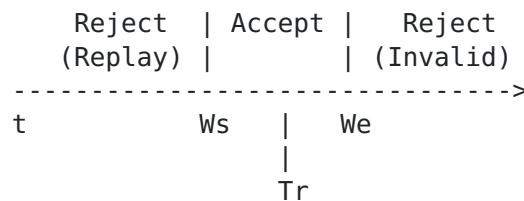
                 Figure 2: D3P Sliding Window

Decapsulation happens as follows. When IP-D3P follows an IP header
(e.g., IPsec Transport Mode) the Next Header value in the IP-D3P
header is placed in the IP header in place of TBD-1. The IP-D3P
header is then removed from the packet. New Total Length and
checksum fields of the IP header are also restored by the receiver.
In the case of an IP-IP encapsulated packet, the IP-D3P header is
simply discarded and processing continues on the encapsulated IP
packet.

3. IP-D3P and IPsec

   When IP-D3P is defined as part of the policy, it is applied before
   ESP or AH processing. Another instance of this approach is IPcomp
   [RFC3173]. Using IP-D3P with IPsec is particularly useful to protect
   group communications when the use of single-sender Security
   Associations (SAs) is not possible, and thus IPsec replay protection
   is not available (Section 2.2 of [RFC4303]). This has been a
   particular problem when applying IPsec to routing protocols using
   multicast communications (e.g., OSPF [RFC4552] and PIM [RFC4601]),
   where maintaining an IPsec SA for each sender in not always feasible.
   Additionally, for operational reasons IPsec SAs and keying material
   are often manually configured for these routing protocols, in which
   case the use of IPsec replay protection in not possible. The use of
   IP-D3P does not replace this missing replay protection, but can
   effectively bound the lifetime of a replayed packet.

   When IP-D3P is applied as part of IPsec transport mode encapsulation,
   the IP-D3P encapsulation is added as the first Protocol following the
   IP header as shown in Figure 3 and Figure 4. Several IP header
   fields need to be adjusting when adding the IP-D3P header: Total
   Length is adjusted to the original length of the payload, the
   Protocol field (IPv4 header) or Next Header field (IPv6) is set to
   TBD-1 identifying IP-D3P as the next protocol, and the header
   checksum is adjusted.

```
             BEFORE APPLYING IP-D3P
          -----------------------------
   IPv4   |orig IP hdr  |     |      |
          |(any options)| TCP | Data |
          -----------------------------

             AFTER APPLYING IP-D3P
          --------------------------------------
   IPv4   |orig IP hdr  |        |     |      |
          |(any options)| IP-D3P | TCP | Data |
          --------------------------------------


             AFTER APPLYING ESP
          -------------------------------------------------------
   IPv4   | orig IP hdr |     |        |   |     | ESP    | ESP|
          |(any options)| ESP | IP-D3P |TCP|Data|Trailer| ICV|
          -------------------------------------------------------
                              |<--------- encryption -->|
                        |<------------- integrity ----->|
```
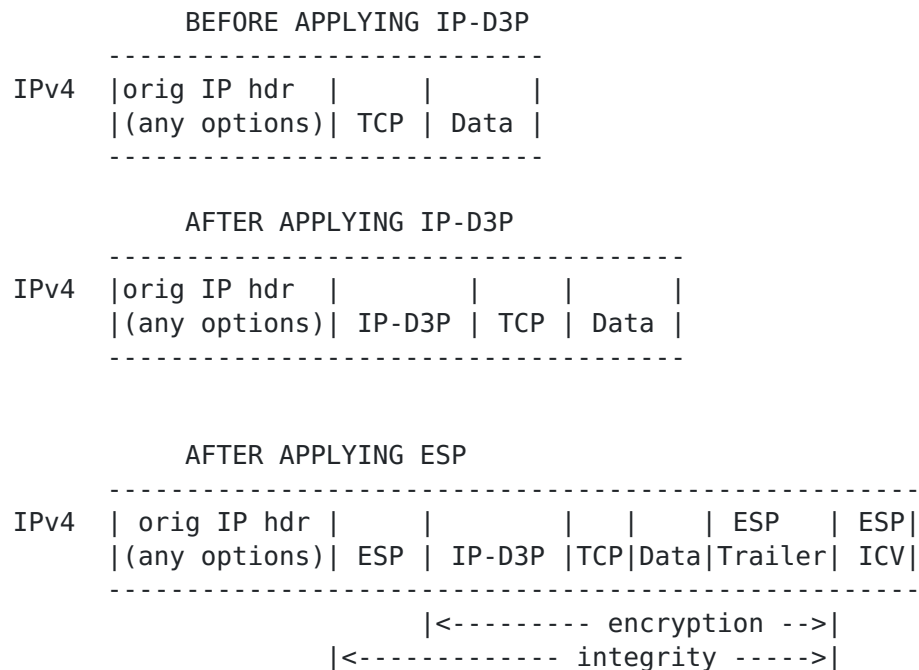
Figure 3: IPv4 Transport Mode Encapsulation

```
                    BEFORE APPLYING D3P
             ----------------------------------------
    IPv6 |                | ext hdrs |     |      |
         | orig IP hdr |if present| TCP | Data |
             ----------------------------------------

                    AFTER APPLYING D3P
             ------------------------------------------------
    IPv6 |                |  ext hdrs |     |     |     |
         | orig IP hdr | if present|IP-D3P| TCP | Data |
             ------------------------------------------------

                    AFTER APPLYING ESP
         -------------------------------------------------------
    IPv6 | orig |orig ext|    |     |   |    | ESP    | ESP|
         |IP hdr| hdrs    |ESP|IP-D3P|TCP|Data|Trailer| ICV|
         -------------------------------------------------------
                             |<------ encryption --->|
                           |<---------integrity ------>|
```

Figure 4: IPv6 Transport Mode Encapsulation

When IP-D3P is applied as part of IPsec tunnel mode encapsulation,
the IP-D3P encapsulation is added immediately after the outer IP
header, such that the inner IP header is fully encapsulated. This is
shown in Figure 5 and Figure 6. The Protocol field in the outer IPv4
header or Next Header field in the outer IPv6 header is set to TBD-1,
indicating that an IP-D3P datagram follows.

```
              BEFORE APPLYING D3P
         ----------------------------
  IPv4  |orig IP hdr  |     |      |
        |(any options)| TCP | Data |
         ----------------------------

              AFTER APPLYING D3P
         ---------------------------------------
  IPv4  |        |orig IP hdr  |     |      |
        | IP-D3P |(any options)| TCP | Data |
         ---------------------------------------


              AFTER APPLYING ESP
         ----------------------------------------------------------------
  IPv4  | new IP hdr  |   |      |orig IP hdr  |   |    | ESP   | ESP|
        |(any options)|ESP|IP-D3P|(any options)|TCP|Data|Trailer| ICV|
         ----------------------------------------------------------------
                            |<--------- encryption -------------->|
                       |<-----------integrity ---------------->|


              Figure 5: IPv4 Tunnel Mode Encapsulation

              BEFORE APPLYING D3P
         ----------------------------------------
  IPv6  |             | ext hdrs |     |      |
        | orig IP hdr |if present| TCP | Data |
         ----------------------------------------

              AFTER APPLYING D3P
         -------------------------------------------------
  IPv6  |     |             | ext hdrs |     |      |
        |IP-D3P| orig IP hdr | if present| TCP | Data |
         -------------------------------------------------

              AFTER APPLYING ESP
         -----------------------------------------------------------------
  IPv6  | new   |new ext|   |      | orig |orig ext|   |    | ESP   | ESP|
        |IP hdr| hdrs  |ESP|IP-D3P|IP hdr| hdrs   |TCP|Data|Trailer| ICV|
         -----------------------------------------------------------------
                            |<--------- encryption ---------------->|
                       |<---------- integrity ------------------>|


              Figure 6: IPv6 Tunnel Mode Encapsulation
```
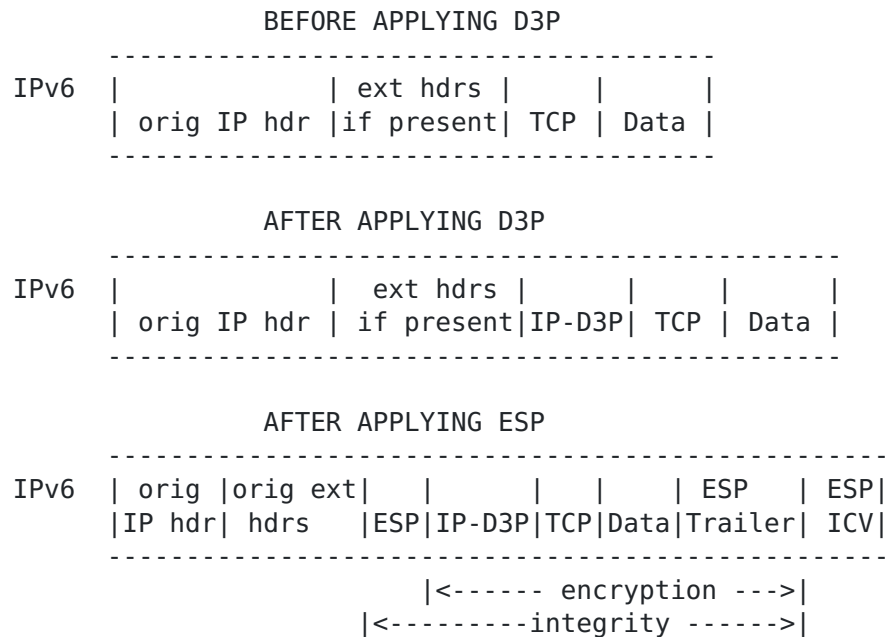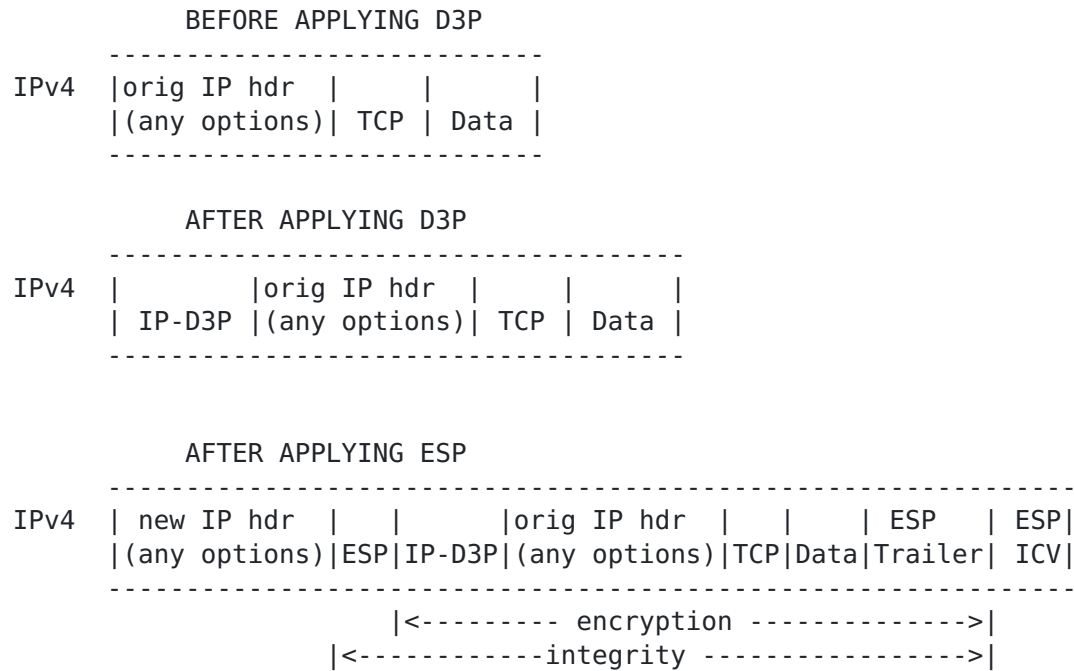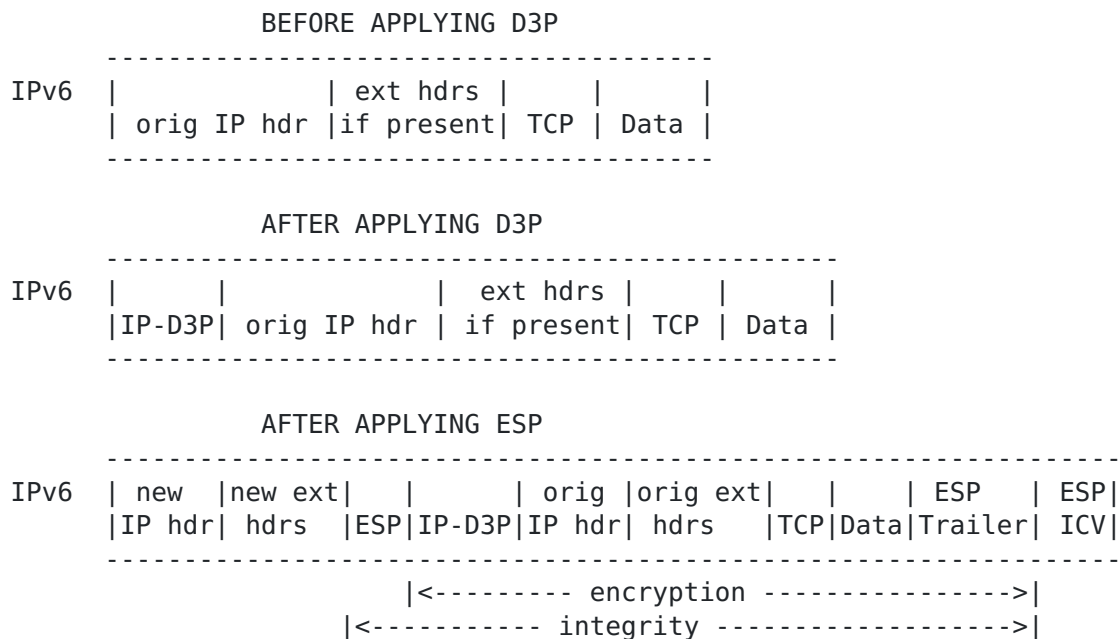
[4](#). Policy Distribution using Group Domain of Interpretation

The Group Domain of Interpretation (GDOI) [[RFC6407](#)] is a group key
management protocol used to distribute group policy and keying
material to a set of Group Members (GMs). When groups using GDOI key
management services require the use of IP-D3P, GDOI distributes this
policy in a Group Associated Policy (GAP) payload using the D3P-TYPE
attribute. This attribute indicates the use of IP-D3P for all SA
TEKs distributed within the SA payload, and defines the Type of
timestamp that is to be placed into datagrams matching those SA TEKs.
Value for the D3P-TYPE attribute are taken from the IP-D3P Type
Registry.

The GDOI KS policy can also define the size of the time window using
the GAP payload D3P-WINDOWSIZE attribute. The value of the attribute
is in the units defined by the Type. For example, for the POSIX-
TIME-SEC attribute the value is in seconds. If the D3P-TYPE is sent
without an accompanying D3P-WINDOWSIZE attribute then window size is
chosen by the receiver.

5. **Security Considerations**

    IP-D3P provides an indication of freshness an IP packet rather than
    an absolute detection of all replays.

    In networks where active attackers are anticipated, packets including
    an IP-D3P header SHOULD be protected by an IPsec transform. IPsec
    integrity protection defends against active man in the middle
    attackers changing the timestamp, and making it appear to be stale or
    invalid.

    The Types defined in this memo all use the system clock. In many
    cases, the system clock is set from an external protocol (e.g., NTP
    [RFC5905]), and indeed this will maximize the likelihood that the
    system clocks of both sender and receiver are synchronized. However,
    care should be taken that external protocol is resistant to a man in
    the middle attack. For example, NTP packets could be distributed
    within an independent well-protected network believed not available
    to active man in the middle attackers, or the NTP Message Digest
    could be used to provide integrity protection.

6. IANA Considerations

    A new IP Protocol is defined for IP-D3P (value TBD-1).

    A new IP-D3P Type registry is created. Values for this attribute are
    shown in the following table. The terms Reserved and Unassigned are
    to be applied as defined in [RFC5226].

                    Value   Type
                    ------   ----
                      0   Reserved
                      1   POSIX-TIME-SEC
                      2   POSIX-TIME-DECISEC
                      3   POSIX-TIME-MSEC
                    4-128   Unassigned
                    129-255   Private Use

    The following additions are made to the GDOI Payloads [GDOI-REG]
    registry.

    Two new attributes are added to the GAP Payload Policy Attributes
    registry.

    o Attribute type D3P-TYPE is a Basic attribute and takes the value
       of TBD-2. Valid values come from the IP-D3P Type registry.

    o Attribute type D3P-WINDOWSIZE is a Variable attribute and takes
       the value of TBD-3. There are no described set of valid values.

7. Acknowledgements

   Some diagrams were adapted from similar diagrams in [RFC4303].

8. References

8.1. Normative References

[POSIX.1] IEEE Std 1003.1, "Standard for Information Technology--
           Portable Operating System Interface (POSIX) Base
           Specifications, Issue 7", 2008.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
           Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the
           Internet Protocol", RFC 4301, December 2005.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an
           IANA Considerations Section in RFCs", BCP 26, RFC 5226,
           May 2008.

8.2. Informative References

[GDOI-REG]
           Internet Assigned Numbers Authority, "Group Domain of
           Interpretation (GDOI) Payload Type Values", IANA Registry,
           <http://www.iana.org/assignments/gdoi-payloads/
           gdoi-payloads.xml>.

[PROT-REG]
           Internet Assigned Numbers Authority, "Protocol Numbers",
           IANA Registry, <http://www.iana.org/assignments/
           protocol-numbers/protocol-numbers.xhtml>.

[RFC3170] Quinn, B. and K. Almeroth, "IP Multicast Applications:
           Challenges and Solutions", RFC 3170, September 2001.

[RFC3173] Shacham, A., Monsour, B., Pereira, R., and M. Thomas, "IP
           Payload Compression Protocol (IPComp)", RFC 3173,
           September 2001.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302,
           December 2005.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)",
           RFC 4303, December 2005.

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality
           for OSPFv3", RFC 4552, June 2006.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas,

                  "Protocol Independent Multicast - Sparse Mode (PIM-SM):
                  Protocol Specification (Revised)", RFC 4601, August 2006.

     [RFC4949]   Shirey, R., "Internet Security Glossary, Version 2",
                  RFC 4949, August 2007.

     [RFC5905]   Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network
                  Time Protocol Version 4: Protocol and Algorithms
                  Specification", RFC 5905, June 2010.

     [RFC6407]   Weis, B., Rowles, S., and T. Hardjono, "The Group Domain
                  of Interpretation", RFC 6407, October 2011.

Authors' Addresses

    Brian Weis
    Cisco Systems
    170 W. Tasman Drive
    San Jose, California 95134-1706
    USA

    Phone: +1-408-526-4796
    Email: bew@cisco.com

    Umesh Mangla
    Juniper Networks Inc.
    1133 Innovation Way
    Sunnyvale, California 94089
    USA

    Phone: +1-408-936-1022
    Email: umangla@juniper.net

    Nilesh Maheshwari
    Juniper Networks Inc.
    1133 Innovation Way
    Sunnyvale, California 94089
    USA

    Phone: +1-408-936-7570
    Email: nileshm@juniper.net

    Thomas Karl
    Deutsche Telekom
    Landgrabenweg 151
    Bonn,   53227
    Germany

    Phone: +49 221 91611582
    Email: thomas.karl@detecon.com