IKEv2 IPv4 Downstream Fragmentation Notification Extension

draft-liu-ipsecme-ikev2-mtu-dect

Liu, Zhang. Migault
The goal of the document is to limit the reassembly operations being performed by the egress security gateway. It defines:

- The IKEv2 Link Maximum Atomic Packet
- The Packet Too Big Extension
Illustrative Example:

1. Mid-tunnel (performed by a router on N) (only for IPv4 DF=0 TLP)

```
+---+---+---+---+---+--+
|IPi|IPe|ESP|IPs|IPd|Da| (TLP)
+---+---+---+---+---+--+
+---+---+--+
|IPi|IPe|ta| (TLP)
+---+---+--+
```
2. Egress node detects fragmentation

- a) it collects IPVersion the IP version of the first fragment as well as FragLen, the fragment length
- b1) If all segments can be reassembled and the reassembled packet is properly decrypted a Link Maximum Atomic Packet Notification (LMAP) is sent on the IKEv2 channel.

```
[IKEv2]
<-- N( LMAP [ IPVersion, FragLen] )
```
b2) If the packet is too big and cannot be fully processed PTB indicates LMTU of the router component of the egress node.

```latex
[IKEv2]
\text{\textbackslash--- N( LMAP [ IPVersion, FragLen] )}
N( PTB [LMTU, EMTU_R] )
```
3. Upon receiving the LMAP or optionally the ingress node
   • a) Update the TMTU so that the Source performs source fragmentation with TTP packets that are not fragmented.

```
Source fragmentation
(IPv6 or IPv4)
   +-----------+
   | IPs|IPd|Da|   (TTP)
   +-----------+
   +-----------+
   | IPs|IPd|ta|
   +-----------+
```
b) Performs inner fragmentation TTP packets that exceeds the TMTU and will generate some fragments.

Inner fragmentation (performed by the Ingress node) (only for IPv4 DF=0 TTP)

```
+---+---+---+---+---+--+
|IPi|IPe|ESP|IPs|IPd|Da| (TLP)
+---+---+---+---+---+--+
+---+---+---+---+---+--+
|IPi|IPe|ESP|IPs|IPd|ta| (TLP)
+---+---+---+---+---+--+
```
In both cases the egress node does not proceed to reassembly operations:

<table>
<thead>
<tr>
<th>IPs</th>
<th>IPd</th>
<th>Da</th>
<th>(TTP)</th>
</tr>
</thead>
</table>

+---+---+--+
| IPs | IPd | Da |
+---+---+--+
| IPs | IPd | ta |
+---+---+--+
The draft has been presented and reviewed several times. We took all comments into consideration.

1. PTB discussion: The IKE PTB, in our view, is largely motivated by enabling the egress interface to provide the EMTU_R (see ietf-intarea-tunnels section 4.2.2.1.

2. TMTU:

- ietf-intarea-tunnels considers the router component - carrying the TTP - and the interface component - handling LTP - independent. Such independence between the Tunnel MTU (for TTP) and link layer MTU for (LTP) is provided by performing outer fragmentation when needed.

- RFC4301 considers the router component can adapt to the specific needs of the interface component. **This is what we do here.**
Differentiated Services Field Codepoints Internet Key Exchange version 2 Notification

draft-mglt-ipsecme-dscp-np

Migault, Halpern, Parkholm, Liu
This document specifies the DSCP Notification Payload, which, in a CREATE_CHILD_SA Exchange, explicitly mentions which DSCP code points will be tunneled in the newly created tunnel.
Illustrative Example

<table>
<thead>
<tr>
<th>Initiator</th>
<th>Responder</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDR, SK {IDi, [CERT,] [CERTREQ,] [IDr,] AUTH, SAi2, TSi, TSr} --&gt;</td>
<td>HDR, SK {IDr, [CERT,] AUTH, SAr2, TSi, TSr} --&gt;</td>
</tr>
<tr>
<td>HDR, SK {SA, Ni, KEi, N(DSCP, AF11, AF3)} --&gt;</td>
<td>HDR, SK {SA, Nr, KEr, N(DSCP, AF11, AF3)} --&gt;</td>
</tr>
<tr>
<td>HDR, SK {SA, Ni, KEi, N(DSCP, EE)} --&gt;</td>
<td>HDR, SK {SA, Nr, KEr} --&gt;</td>
</tr>
</tbody>
</table>
Thanks!