IPv6 Minimum Path MTU Hop-by-Hop Option

<draft-hinden-6man-mtu-option-00>

Bob Hinden
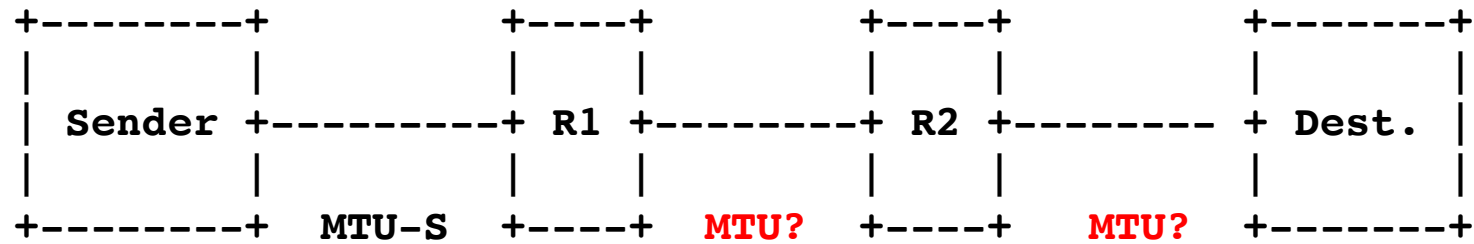
Gorry Fairhurst

IETF103 Bangkok

Background

- Current RFC8201 PMTUD isn't working well

- This hop-by-hop option came from the idea that it will be more reliable for the Destination to send Path MTU feedback to the Source
  - Better trust relationship than RFC8201 PMTUD

Goals

- Adapt to varying Path MTU over connection life time.

- Avoid complicated probing when path has multiple successive bottlenecks.
  - Like to detect Path MTU in single round trip

- Needs to work with network devices that "read" transport headers.

The Problem

```
+--------+          +----+          +----+          +-------+
|        |          |    |          |    |          |       |
| Sender +----------+ R1 +----------+ R2 +--------- + Dest. |
|        |          |    |          |    |          |       |
+--------+  MTU-S   +----+   MTU?   +----+   MTU?   +-------+
```

The Hop-by-Hop Option

- Send an IPv6 packet with a hop-by-hop option with a minimum PMTU.

- Payload carries a transport header to associate the packet with the PL flow.

```
    Option        Option        Option
     Type        Data Len        Data
   +--------+--------+--------+--------+
   |BBCTTTTT|00000010|   2 octet value  |
   +--------+--------+--------+--------+
```
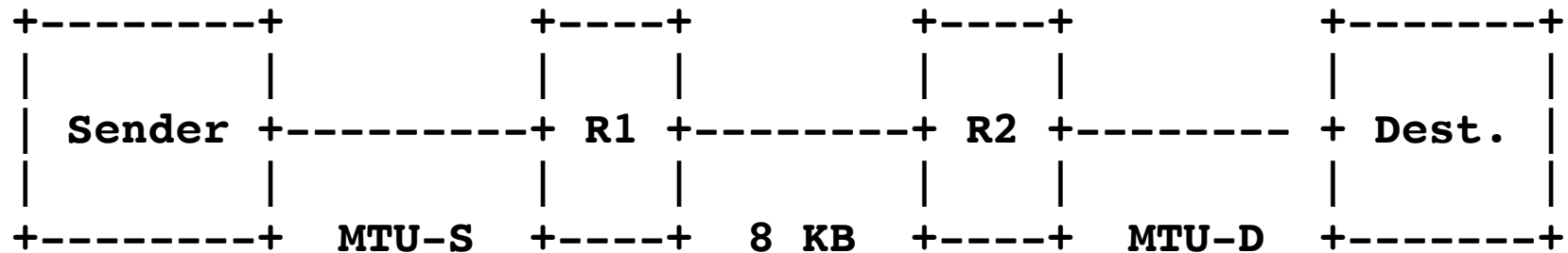
- RFC8200 allows devices to skip over HBH Option.

- Destination sends ICMP message to Source with Minimum Path MTU value.

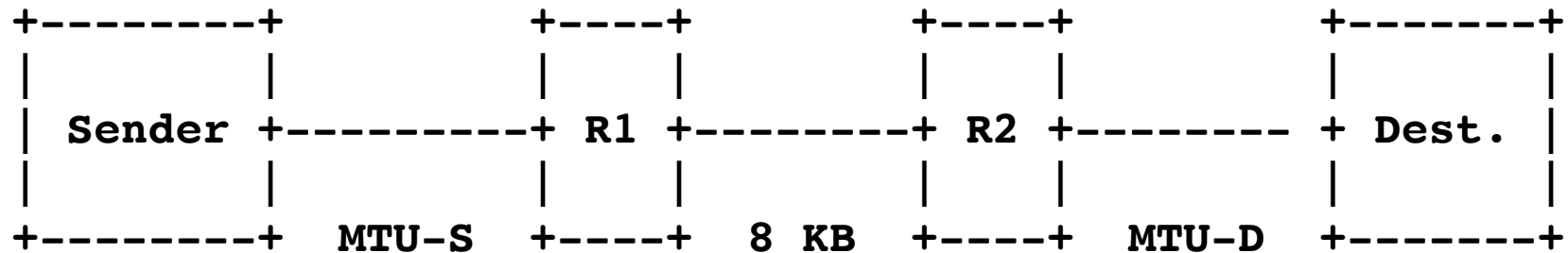- Source becomes aware of a potential target to probe the PLPMTU.

Using the Option

- Relation to PLPMTUD:

  - Max Packet Size can only be checked at the Packetization Layer.

  - The PLPMTU method MUST already be robust to failure, and path changes.

  - PLPMTUD method does NOT rely on this option.

  - It makes little sense to send this option with a probe > PLPMTU!!

Example Scenarios (1 of 3)

```
+--------+              +----+           +----+           +-------+
|        |              |    |           |    |           |       |
| Sender +--------------+ R1 +-----------+ R2 +-------- + Dest. |
|        |              |    |           |    |           |       |
+--------+   MTU-S    +----+   8 KB   +----+   MTU-D   +-------+
```

```
+-+-----+-----+----+----+----------+------------------------+
| |MTU-S|MTU-D| R1 | R2 | Rec PMTU | Note                   |
+-+-----+-----+----+----+----------+------------------------+
| | 8KB | 8KB | H  | H  |  8 KB    | Endpoints attempt to   |
| |     |     |    |    |          | use an 8 KB PMTU.      |
+-+-----+-----+----+----+----------+------------------------+
```

Example Scenarios (2 of 3)

```
+--------+           +----+           +----+           +-------+
|        |           |    |           |    |           |       |
| Sender +-----------+ R1 +-----------+ R2 +-------- + Dest. |
|        |           |    |           |    |           |       |
+--------+   MTU-S   +----+   8 KB    +----+   MTU-D   +-------+
```

```
+-+-----+-----+----+----+----------+----------------------+
| |MTU-S|MTU-D| R1 | R2 | Rec PMTU | Note                 |
+-+-----+-----+----+----+----------+----------------------+
| | 8KB |1500B| H  | H  | 1500B    | Endpoints attempt to |
| |     |     |    |    |          | use an 1500 B PMTU.  |
+-+-----+-----+----+----+----------+----------------------+
```

Example Scenarios (3 of 3)

```
+---------+           +----+            +----+           +--------+
|         |           |    |            |    |           |        |
| Sender  +-----------+ R1 +------------+ R2 +-------- + Dest. |
|         |           |    |            |    |           |        |
+---------+   MTU-S   +----+    8 KB    +----+   MTU-D   +--------+
```
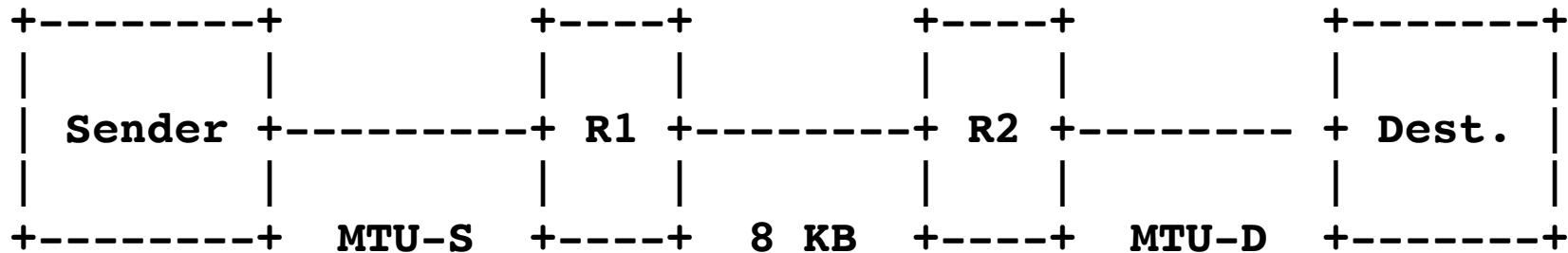
```
+-+-----+-----+----+----+----------+------------------------+
| |MTU-S|MTU-D| R1 | R2 | Rec PMTU | Note                   |
+-+-----+-----+----+----+----------+------------------------+
| | 8KB |1500B|  H |  - |   8 KB   | Endpoints attempt to   |
| |     |     |    |    |          | use an 8 KB PMTU, but  |
| |     |     |    |    |          | need to implement a    |
| |     |     |    |    |          | method to fall back    |
| |     |     |    |    |          | to use an 1500B PMTU   |
+-+-----+-----+----+----+----------+------------------------+
```

Other Work

- RFC7872 published data in 2016, reporting high drop rates.
  - This care needed about which packets are marked.
  - What are the pathologies for a destination PTB message?
  - Do we have new data?
- Related drafts:
  - draft-troan-6man-pmtu-solution-space-00
  - draft-ietf-tsvwg-datagram-plpmtud
  - draft-leddy-6man-truncate
- Also being discussed in TSVWG

Next Steps

- Adopt as a 6man draft?

QUESTIONS / COMMENTS?