Path MTU discovery improvements at the network layer?

ot@cisco.com

draft-troan-6man-pmtu-solution-space

Why?

- Current RFC1981/RFC8201 ICMP based PMTUD:
  - Signalling back to sender not robust:
    - ICMP error sending is throttled in routers
    - ICMP is filtered in firewalls
    - Ignored by hosts
    - Doesn’t work well with anycast / load-balancers
- PMTUD requires multiple round-trips to detect path MTU
- Detects MTU per destination not per flow

Existing work

- RFC1063: IP MTU discovery options
- RFC1191: Path MTU discovery
- RFC1435: IESG Advice from Experience with Path MTU Discovery
- RFC1981: Path MTU Discovery for IP version 6
- RFC2923: TCP Problems with Path MTU Discovery
- RFC4459: MTU and Fragmentation Issues with In-the-Network Tunneling
- RFC4638: Accommodating a Maximum Transit Unit/Maximum Receive Unit (MTU/MRU) Greater Than 1492 in the Point-to-Point Protocol over Ethernet (PPPoE)
- RFC4821: Packetization Layer Path MTU Discovery
- RFC7588: A Widely Deployed Solution to the Generic Routing Encapsulation (GRE) Fragmentation Problem
- RFC7690: Close Encounters of the ICMP Type 2 Kind (Near Misses with ICMPv6 Packet Too Big (PTB))
- RFC8201: Path MTU Discovery for IP version 6
- RFC8249: Transparent Interconnection of Lots of Links (TRILL): MTU Negotiation

Routing area

- Support for Path MTU (PMTU) in the Path Computation Element Communication Protocol (PCEP).
  draft-dhody-pce-pcep-pmtu-00

- IS-IS Extensions for Path MTU
  draft-hu-lsr-isis-path-mtu-00

- Segment Routing Path MTU in BGP
  draft-li-idr-sr-policy-path-mtu-00

- BGP-LS Extensions for Advertising Path MTU
  draft-zhu-idr-bgp-ls-path-mtu-00

Multicast

- Path Maximum Transmission Unit Discovery (PMTUD) for Bit Index Explicit Replication (BIER) Layer
  draft-ietf-bier-path-mtu-discovery-04

- BIER MTU Discovery
  draft-venaas-bier-mtud-02

Goals / Requirements

- Robust
- Deployable
- Detect path MTU in a single RTT
- Not easy target for filtering
  (signalling back should be in transport/application stream?)
- Per-flow MTU
- Support per-neighbour path MTU discovery

Solutions

- Think of detection of MTU separated from signalling the MTU back to sender

- Signalling can be done with existing ICMP PTB message, L3, L4 or L7 options.

Solution #1
Punt to transport

- RFC4821: Packetization Layer Path MTU Discovery
- Packetization Layer Path MTU Discovery for Datagram Transports
  draft-ietf-tsvwg-datagram-plpmtud-05

- Cons: Can’t always distinguish between congestion and MTU failure.

Solution #2
Do nothing

Solution #3
In-path fragmentation

Solution #4
Fixed MTU of 1280

Solution #5
Truncation

- Sender sends a packet sized to the outgoing interface MTU. And sets a Truncation Eligible flag (TE),

- Intermediate routers for where the outgoing packet is larger than the outgoing interface MTU truncates the packet and forwards it.

    - It also records the fact that the packet is truncated by setting a Truncated Flag (TC)

- Could be done with normal data packets, if padding was applied, and the original packet length was recorded.

Solution #6
Recording

- Sender sends < 1280 packet with new HBH option. Sets MTU value in HBH option to outgoing interface MTU.

- Intermediate routers compare the HBH MTU value, and rewrites if router’s outgoing interface MTU is smaller.

- Receiver signals learnt MTU back to sender.

- Packet can be a separate probe or the HBH option is attached to a normal data packet (e.g. TCP SYN)

Conclusion