IPv6 Packet Truncation
(draft-leddy-6man-truncate)

Ron Bonica
John Leddy
Igor Lubashev
IETF 103

The Status Quo

- Source nodes can avoid PMTU black holes by doing one of the following:
  - Never sending packets much larger than the IPv6 minimum link MTU
    - Limit between 1280 and 1440 bytes
  - **RELIABLY** discovering the PMTU and never sending packets larger than the discovered limit
- Currently, many hosts never send packets much larger than the IPv6 minimum link MTU
  - Because PMTUD isn't reliable

What Network Service Is Required?

- Network must provide instant feedback to the source node when it sends a packet larger than the PMTU
- Feedback mechanism must be activated on a per-session basis at the request of an upper-layer protocol
  - Default is deactivated
- Feedback mechanism must detect PMTU decreases
  - But is not required to detect PMTU increases
- Feedback mechanism must be more reliable than PMTUD is today
  - But is not required to be 100% reliable
  - Because that goal is not achievable

Approach: Make PMTUD More Reliable

- Today, when the source node sends a packet larger than the MTU, the network sends one ICMP PTB message
  - From an intermediate node
- In the proposed approach, when the source node sends a packet larger than the MTU, the network sends two ICMP PTB messages
  - One from an intermediate node
  - One from the destination node
- If the network fails to deliver the ICMP PTB from the intermediate node, it may deliver the ICMP PTB from the destination node
  - Because the destination node is very likely to have a viable path to the source node

Draft-leddy-6man-truncate (Current Version)

- Source node marks packet with Truncation Eligible Destination Option
- If an intermediate node cannot forward the packet because of MTU:
  - Sends an ICMP PTB message to the source
  - Overwrites the Truncation Eligible option with a Truncated Packet option
  - Truncates the packet
  - Recalculates the upper-layer checksum, if possible
  - Forwards the truncated packet to the destination node
- If the destination recognizes the Truncated Packet Option
  - Sends an ICMP PTB message to the source and discards the packet
- If the destination does not recognizes the Truncated Packet Option
  - Sends an ICMP Parameter Problem message to the source and discards the packet

Status of Current Version

- Significant Objections
  - Updating checksum in flight
  - Overwriting Destination Option type in flight
- Refinements
  - Sunday night meeting with three of the people raising objections
  - Proposed modifications follow

Do Not Update The Checksum

- Benefit
  - Reduces risk of truncated packet being introduced to the data stream
- Cost
  - Reduces feedback mechanism reliability
  - Reduces the probability that the truncated packet will be delivered to the destination node
    - Many stateful middle boxes (e.g., firewalls) silently discard packets that have incorrect checksums

Collapse Destination Options (Part I)

- Name: Truncation
- Type: TBD
  - Act bits (10) indicate the required action if processing node does not recognize the option. The required action is to discard packet and send ICMP Parameter Problem message to the source node.
  - Chg bit (0) indicates that option data cannot be modified in flight
- Option Data Length: Must be equal to 2 bytes
- Option Data: Set by source. Must be equal to Payload Length field in IPv6 header

Collapse Destination Options (Part II)

- Source node marks packet with Truncation Destination Option
- If an intermediate node cannot forward the packet because of MTU:
  - Sends an ICMP PTB message to the source
  - Truncates the packet
  - Forwards the truncated packet to the destination node
- If the destination node does not recognize the Truncation Option
  - Sends an ICMP Parameter Problem message to the source node
  - Discards the packet
- If the destination node recognizes the Truncation Option and the Option Data does not equal the IPv6 Payload Length
  - Sends an ICMP PTB message to the source node
  - Discards the packet
- Otherwise, the destination node continues to process the packet

Collapse Destination Options (Part III)

- Benefit
  - Avoids overwriting Option Type in flight
  - Makes feedback mechanism compatible with AH integrity check
- Cost
  - Additional cost incurred when destination node does not recognize the option and no packets larger than the PMTU are sent
  - Limits applicability of mechanism to select applications that are only deployed on nodes that recognize the Truncation option

Reliability

- No feedback mechanism is 100% reliable
  - Because the network is not 100% reliable
- When an application detects a session stall
  - The root cause may be a PMTU black hole that was not detected by the feedback mechanism
  - The root cause may be something else (e.g., congestion)
  - Fallback procedures are required
- Probing
  - Draft-hinden-6man-mtu-option
  - PLPMTUD

Next Steps

- Discuss proposed changes here
- Next version reflects outcome of the discussion
- Post before the end of IETF 103

Questions