NAT Behavioral Requirements for ICMP

draft-ietf-behave-nat-icmp-01.txt

Presented by Fernando Gont (UTN/FRH)

67th IETF Meeting, November 5-10, 2006
San Diego, CA, U.S.A.

REQ#1 – ICMP Query ID Mapping

- 1. NAT MUST permit ICMP query based applications from private hosts.

  a) ICMP Query ID mapping SHOULD be external host independent.
- Justification:
  - 1a) is useful for future P2P apps; Also, useful for debugging.
  - 1a) is not a requirement for Ping/Traceroute.
  - 1a) is listed as a SHOULD.

REQ#2 – ICMP Query Session Timeout

- 2. ICMP Query session timer MUST NOT expire in less than 60 seconds.

  a) ICMP Query session timer is RECOMMENDED to be made configurable.
- Justification:
  - □ Max. RTT is 4 minutes.
  - □ Most ICMP apps complete in a few seconds.
  - □ 60 secs is a pragmatic timeout.

REQ#3 – ICMP Error Payload Validation

- 3. NAT SHOULD do the following on ICMP error payload

  a) IP checksum: Validate IP checksum.

  b) IP options: Move past IP options to get to the transport header.

  c) Transport checksum: Validate transport checksum if the entire transport segment is embedded and the transport protocol mandates checksum validation.
- Justification:
  - Newer revisions of endhost stacks do not accept ICMP error msgs with invalid IP or transport checksums.
  - ICMP error payloads messed up when NATs assume fixed offset to locate the transport header.

REQ#4 – Inbound ICMP Error Msg

- 4: For an inbound ICMP Error Msg, the NAT MUST do the following prior to forwarding.

  a) Revert IP and transport headers of the embedded packet to original form, using the matching mapping;

  b) Leave ICMP error type and code unchanged;

  c) Dest IP of the ICMP Msg: Modify to be same as src IP of embedded packet.
- Justification:
  - ICMP payload to be modified using matching mapping .
  - Dest IP of ICMP msg to be same as Src IP in the payload.

REQ#5 – Outbound ICMP Error Msg

- 5: For an Outbound ICMP Error Msg, the NAT MUST do the following prior to forwarding.

  a) Revert IP and transport headers of the embedded packet to original form, using the matching mapping;

  b) Leave ICMP error type and code unchanged;

  c) Src IP of the ICMP Msg: If NAT is enforcing Basic NAT and has active mapping for ICMP error msg sender, use the public address mapping of the sender. Otherwise, use NAT's own public IP address.
- Justification:
  - ICMP payload to be modified using matching mapping .
  - Src IP of ICMP msg to be same as NAT's public IP or IP address mapping of the message sender (if NAT is enforcing Basic NAT)

Conflicting ICMP Payload Translation Rqmts - Problem Stmt

- UDP draft says that ICMP payload SHOULD be translated and forwarded (Ref: section 9, preceding REQ-12 of UDP draft)
- ICMP draft says that ICMP payload MUST be translated and forwarded (Ref: REQ-4 & REQ-5 of ICMP draft).
- Updated TCP draft says that ICMP payload MUST be translated and forwarded (Ref: section 7.3, preceding REQ-9 of TCP draft). This is in alignment with the ICMP requirements.
- Problem: Which of the drafts provides the normative precedence on ICMP payload translation requirement?

Conflicting ICMP Payload Translation Rqmts - Resolution

- Spencer Dawkins says – Issue UDP RFC as is, UPDATE it with ICMP; TCP draft refers to ICMP for the ICMP requirements supporting TCP;
- Christian Huitema says- Specifying how the NAT will translate ICMP packets belongs logically to the ICMP specification.
- Resolution: Adopt Spencer’s recommendation?

REQ#6 – ICMP Errors MUST NOT effect Embedded Sessions

- 6: While processing an ICMP error packet, a NAT device MUST NOT refresh or delete the NAT Session that pertains to the embedded payload.
- Justification:
  - Spoofing a NAT with ICMP error messages must not effect the liveness of NAT Sessions.

Embedded Session Impact from ICMP Error Msgs– Problem

- Req#12 of the UDP draft says - Receipt of an ICMP error message MUST NOT terminate NAT mapping.
- Req#6 of ICMP draft says – Upon Receipt of an ICMP error message, NAT MUST NOT refresh or delete the NAT Session that pertains to the embedded payload.
- Req#9 of TCP draft says - Receipt of an ICMP error message MUST NOT terminate the NAT mapping or TCP connection for which the ICMP was generated.
- Problem: Which of the drafts provides normative rqmt on the impact an ICMP error msg might have on the NAT session pertaining to embedded payload?

Embedded Session Impact from ICMP Error Msgs– Resolution

- Christian’s says - How the translation state for a protocol is affected by the receipt of ICMP belongs logically in the specification of that protocol. E.g., reaction of TCP to "soft errors" or "hard errors" belongs in TCP.
- Suresh says - The text for Req#6 in the ICMP draft is neutral to the transport protocol of the embedded packet. So, why not keep the normative text in the ICMP draft?
- Resolution: Adopt Christian’s proposal and require each of the drafts to explicitly repeat the requirement as pertaining to that protocol? Ex: The Req#6 in ICMP draft to be scoped to the ICMP error msgs pertaining to ICMP query pkts only?

Req#7- ICMP Hairpinning Support

- 7: Basic NAT MUST support hairpinned ICMP query sessions. All NATs MUST support hairpinned ICMP error messages.

  a) When forwarding a hairpinned error msg, NAT MUST translate Dest IP of the outer IP to be same as src IP of the embedded packet.
- Justification:
  - Hairpinning support for ICMP queries & error msgs
  - Extend REQ#5 for hairpinning.

Req#8 – Use of ICMP Error code 13

- 8: When a NAT is unable to establish a NAT Session for a new flow due to resource constraints or administrative restrictions, the NAT SHOULD send ICMP error with a code of 13 to the sender, and drop the original packet.
- Justification:
  - RFC 1812 recommends the use of Error code13 (Communication administratively prohibited) for administrative restrictions.

Req#9 – Conform to RFC 1812

- 9: NAT MUST conform to RFC 1812 in IP packet handling. Specifically,

  a) If DF bit is set and the NAT cannot forward without fragmentation, the NAT MUST send a "Packet too big" ICMP message (type 3, Code 4) with a suggested MTU back to the sender and drop the original IP packet.

  b) NAT MUST generate "Time Exceeded" ICMP error message when it discards a packet due to expired TTL, unless explicitly configured otherwise.
- Justification:

  - As router, NAT must conform to RFC 1812.

Some tweaks to be incorporated

- Clarification on how to recompute the TCP checksum when the packet that elicited the ICMP error message was source-routed. (As pointed by S. Guha)
- Several tweaks suggested by P. Matthews
- Several tweaks suggested by F. Gont

Next Steps

- ICMP draft Ready for WG last call.
- Any questions/comments?